Biogeophysical Effects of Land-Use and Land-Cover Change Not Detectable in Warmest Month

Luke Grant, Lukas Gudmundsson, Edouard L Davin, David M Lawrence, Nicolas Vuichard, Eddy Robertson, Roland Séférian, Aurélien Ribes, Annette L Hirsch, Wim Thiery

Biogeophysical Effects of Land-Use and Land-Cover Change Not Detectable in Warmest Month

LUKE GRANT,[a] LUKAS GUDMUNDSSON,[b] EDOUARD L. DAVIN,[c,d,e] DAVID M. LAWRENCE,[f] NICOLAS VUICHARD,[g] EDDY ROBERTSON,[h] ROLAND SÉFÉRIAN,[i] AURÉLIEN RIBES,[i] ANNETTE L. HIRSCH,[j] AND WIM THIERY[a]

[a] *Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussels, Belgium*
[b] *Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland*
[c] *Wyss Academy for Nature at the University of Bern, Bern, Switzerland*
[d] *Climate and Environmental Physics, Physics Institute, University of Bern, Bern, Switzerland*
[e] *Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland*
[f] *National Center for Atmospheric Research, Boulder, Colorado*
[g] *Laboratoire des Sciences du Climat et de l'Environnement, Gif sur Yvette Cedex, France*
[h] *Met Office Hadley Centre, Exeter, United Kingdom*
[i] *CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France*
[j] *Australian Research Council Centre of Excellence for Climate Extremes, University of New South Wales, New South Wales, Australia*

(Manuscript received 24 May 2022, in final form 7 October 2022)

ABSTRACT: Land-use and land-cover changes (hereafter simply "land use") alter climates biogeophysically by affecting surface fluxes of energy and water. Yet, near-surface temperature responses to land use across observational versus model-based studies and spatial-temporal scales can be inconsistent. Here we assess the prevalence of the historical land use signal of daily maximum temperatures averaged over the warmest month of the year (t_{LU}) using regularized optimal fingerprinting for detection and attribution. We use observations from the Climatic Research Unit and Berkeley Earth alongside historical simulations with and without land use from phase 6 of the Coupled Model Intercomparison Project to reconstruct an experiment representing the effects of land use on climate. To assess the signal of land use at spatially resolved continental and global scales, we aggregate all input data across reference regions and continents, respectively. At both scales, land use does not comprise a significantly detectable set of forcings for two of four Earth system models and their multimodel mean. Furthermore, using a principal component analysis, we find that t_{LU} is mostly composed of the nonlocal effects of land use rather than its local effects. These findings show that, at scales relevant for climate attribution, uncertainties in Earth system model representations of land use are too high relative to the effects of internal variability to confidently assess land use.

KEYWORDS: Principal components analysis; Regression analysis; General circulation models; Land surface model; Climate variability; Trends

1. Introduction

Land-use and land-cover changes (referred to simply as "land use") affect climate biogeochemically and biogeophysically. The biogeochemical effects of land use on climate occur through interactions in the carbon cycle, mostly moving carbon from land reserves to the atmosphere. Land use also influences climate biogeophysically by modulating energy and water fluxes that vary between different vegetation types and surface characteristics.

The biogeophysical effects of deforestation on near-surface temperatures are controlled by evapotranspiration, surface roughness and albedo (Jia et al. 2019; Pongratz et al. 2021). As the sign of these effects varies across latitudinal bands and time scales, they may counteract the biogeochemical cooling effects of re/afforestation efforts. At an annual mean time scale, deforestation may cool temperatures at high latitudes (de Noblet-Ducoudré et al. 2012) and warm them in mid-to-equatorial latitudes (Duveiller et al. 2018), yet for the summer season deforestation could ubiquitously lead to warming (Lejeune et al. 2018). In the northern latitudes, this cooling is understood to be caused by the increased albedo and reduced available energy of deforested regions relative to their darker, forested origins. Importantly, in the Northern Hemisphere winter, this albedo effect is amplified. At this time, existing forest cover traps outgoing radiation through multiple reflection within the canopy, and a reduced canopy exposes larger areas of the land surface to snowfall (de Noblet-Ducoudré et al. 2012). In lower latitudes, which are less energy-limited, the temperature response of deforestation is dominated by the reduced capacity of the land surface to lose heat through turbulent fluxes. With losing vegetative cover, the lowered canopy conductance of water from soil moisture reserves decreases evapotranspiration and the latent heat flux. As well, surface roughness is reduced after deforestation, thereby reducing the dissipation of energy from the surface through turbulent sensible and latent heat fluxes (Davin and de Noblet-Ducoudre 2010; Winckler et al. 2019b). The latter of these warming mechanisms

Denotes content that is immediately available upon publication as open access.

Corresponding author: Luke Grant, luke.grant@vub.be

DOI: 10.1175/JCLI-D-22-0391.1

may also explain the boreal summer warming response to deforestation. In this case, the reduced turbulent heat fluxes from a decrease in surface roughness forces a temperature increase, compensating the albedo effect's cooling (Bright et al. 2017; Duveiller et al. 2018).

These biogeophysical temperature responses to deforestation, however, are inconsistent across modeling studies using Earth system models (ESMs) (Lejeune et al. 2017, 2018). In idealized and historical model simulations, modeled temperature responses differ in magnitude (Winckler et al. 2019c; De Hertog et al. 2022) and in the latitudes where they switch from a warming to a cooling signal (Boysen et al. 2020). This is owed to uncertainties in our understanding of historical land cover change, how land use change is implemented in the models and resulting model-dependent physics for land–atmosphere interactions (Pitman et al. 2009; Meier et al. 2019, 2022). These differences are evident in model discrepancies in the simulated partitioning of net radiation into sensible and latent heat flux in response to deforestation, for example, even when they agree on the reduction of available energy and the sign of change in temperature responses to deforestation (de Noblet-Ducoudré et al. 2012; Boysen et al. 2020).

Comparisons between modeled and observed biogeophysical temperature responses to deforestation are challenging. Often, models may simulate cooling responses to deforestation while observations show warming (Winckler et al. 2019a). This difference stems from how model simulations at the grid scale and observations taken at single points in the landscape are characterizing the local and nonlocal responses of deforestation at different spatial scales. Local effects refer to atmospheric changes at the site of a land disturbance. Yet, the atmospheric conditions over a given plot of land may also be affected by the atmospheric feedbacks of land use occurring in other regions, i.e., the nonlocal effects. Some have used the "space-for-time" concept to describe the trade-offs in the assessment of this signal between models and observations. This is because observational studies, by construction, sometimes only include the local signal. For example, in situ and satellite observations may compare a local plot of deforested land to neighboring forests to assess this signal at a single point in time. In this context, the effects of internal climate variability, including their influence on the nonlocal atmospheric interactions of land use, are constant across neighboring sites and are therefore assumed to cancel (Lee et al. 2011). These assumptions no longer hold in many studies assessing the temperature response of deforestation from ESM simulations with grid scales greater than 100 km. Comparisons of historical simulations without land use against those with time-evolving land use scenarios ingest both the local effects and the nonlocal feedbacks of deforestation at the grid cell level. Further observation-model discrepancies in the case of satellite assessments stem from an inappropriate comparison of variables. Satellites measure surface temperatures, while most model studies assess near-surface or 2-m air temperatures (Bright et al. 2017; Duveiller et al. 2018). Indeed, comparisons of modeled surface and near-surface temperature responses to deforestation reveal a disagreement in their magnitudes and sometimes in their sign of change (Winckler et al. 2019c; Breil et al. 2020; Pongratz et al. 2021). Finally, Chen and Dirmeyer (2020) reconcile these differences between observations and a single model. In running offline simulations to avoid feedbacks and by evaluating them to mimic the space-for-time conditions of observational studies, they replicate the observed warming in response to deforestation. It is thus likely that the observation–model discrepancies in the temperature response to land use are not only due to model imperfections, but also because the nonlocal effects of land use are less accounted for in the observations.

The local and nonlocal effects of deforestation have only been separated in the model world. This is achieved by reconstructing historical deforestation signals and by performing idealized simulations. Reconstruction methods for assessing the signals of historical deforestation have been designed in cases where counterfactual historical simulations—those including all historical forcings but which hold land use constant—were not available to derive the land use signal. With modeled temperature and tree cover information, these applications use either regressions (Winckler et al. 2019c; Lejeune et al. 2018; Thiery et al. 2020) or mean differences between high and low tree cover change grid cells (Kumar et al. 2013; Lejeune et al. 2017) in a rolling window of grid cells to parse the temperature changes owing to varying tree cover. Importantly, this also filters some nonlocal effects under the assumption that they are shared in each window. This is evidenced by the reduction of uncertainty in the reconstructed temperature signal of deforestation when compared to factorial derivations of the signal (i.e., a lowered influence of internal variability, which modulates nonlocal effects) (Lejeune et al. 2018). Idealized "checkerboard"-type deforestation analyses apply a methodology to extract local against nonlocal effects of deforestation by using composite fields of forested and deforested grid cells (Winckler et al. 2017; De Hertog et al. 2022). These studies show that models simulate summertime local warming and nonlocal cooling. Similarly, in idealized deforestation studies that apply global deforestation rates not expected in any reasonable future land use scenario, it is thought that the nonlocal cooling masks local warming effects (Boysen et al. 2020). While evapotranspiration efficiency and surface roughness changes drive the local response (Winckler et al. 2017), albedo changes are thought to be the causal factor behind this nonlocal cooling. Albedo changes affect the temperature and humidity of the whole troposphere, allowing for advection of these perturbed conditions to other regions (Davin and de Noblet-Ducoudre 2010).

In addition to improving our process understanding of the biogeophysical effects of land use, resolving the significance of its contribution to historical climate change is of key importance. Yet this question of detection and attribution of the effects of land use on climate has received relatively little attention (Christidis et al. 2013). Detection and attribution involve assessing the significance of a signal in some aspect of the climate, and then evaluating the roles of different climate forcings that might affect it (Hegerl et al. 2006). This has been undertaken extensively, for example, to evaluate the role of anthropogenic greenhouse gas emissions in affecting target variables like temperature and precipitation relative to natural forcings, such as volcanoes and solar activity (Zwiers et al.

TABLE 1. Observational and model-based dataset characteristics.

Observations and models	Nominal resolution	t_{HIST} samples	t_{HISTNL} samples	Dynamic vegetation	Land use frequency	Plant functional types
BE	360 × 180	—	—	—	—	—
CRU	720 × 360	—	—	—	—	—
CanESM5	500 km (128 × 64)	35	7	No	Annual	9
CNRM-ESM2-1	250 km (256 × 128)	11	3	No	Annual	16
IPSL-CM6A-LR	250 km (144 × 143)	32	4	No	Annual	15
UKESM1-0-LL	250 km (192 × 144)	16	4	Yes	Annual	9

2011; Wan et al. 2015). Detection is defined as first demonstrating that a target variable has changed in a statistically significant way. This significance is defined relative to the distribution of possible changes in the target variable as forced by internal climate variability. Therefore, a signal is detected if it can be confidently determined that its probability of occurring under internal climate variability is sufficiently small. Upon detecting a signal, questions remain over which climate forcings were responsible for its significant shift. Attribution is then the process of determining the suspect forcings responsible for this shift against other plausible explanations (Hegerl et al. 2007).

One challenge in the context of detecting and attributing the effects of land use on climate is the potential nonlocal cooling response to deforestation. For locations and time scales relevant to the nonlocal cooling response, this conflicts with the spatially uniform carbon sequestration benefits of re/afforestation mitigation efforts. Existing evidence implies that this cooling may be forced by nonlocal processes linked to internal climate variability. Therefore, exploring the significance of the historical temperature response to land use in a way that takes into account internal variability is helpful. Methodologically, detection and attribution involves ensemble sampling and optimization steps that reduce the influence of internal climate variability on the historical land use signal.

Recent advances in land use scenarios (Hurtt et al. 2020) and the availability of historical simulations with and without land use change (Lawrence et al. 2016), which can dissect the biogeophysical effects of land use in phase 6 of the Coupled Model Intercomparison Project (CMIP6), merit a revisiting of the detection and attribution of the overall land use signal. Previously, a single-model study found an attributable cooling of historical warmest day temperatures to land use (Christidis et al. 2013). The researchers used an earlier generation of UKESM that included a dynamic vegetation model that ingested agricultural disturbance data from early land use harmonization datasets (Hurtt et al. 2011). However, considering the discrepancies in modeled temperature responses to land use, a more complete estimate of the land use signal's detectability may only be achieved in a multimodel framework. Here we revisit the detection and attribution of land use effects on climate by focusing on 2-m warmest month temperatures (the locally highest annual value of monthly means of daily maximum temperatures) for four ESMs simulated under the updated protocol of CMIP6 (Eyring et al. 2016).

2. Data and methods

a. Models and observations

For detection and attribution, we require observations and ESM simulations of historical and preindustrial climates (Ribes et al. 2013). Model simulations come from the following CMIP6 concentration-driven experiments: all historical forcings ("historical," hereafter HIST), all historical forcings without land use ("hist-noLu," hereafter HISTNL), and a preindustrial control representing the effects of internal variability ("piControl," hereafter PIC) (Eyring et al. 2016). Here, concentration-driven means that models are forced by scenario-prescribed CO_2 concentrations instead of calculating atmospheric CO_2 explicitly within an ESM's carbon cycle. The use of simulations from concentration-driven experiments isolates the biogeophysical effects of land use (Lawrence et al. 2016). For observations, we use Berkeley Earth monthly land temperature fields (BE; Rohde and Hausfather 2020) and Climatic Research Unit gridded Time Series version 4.04 (CRU; Harris et al. 2020), which span the historical period and provide gridded temperature fields for monthly means of daily maximum temperatures at higher resolutions than the model fields (Table 1).

ESMs include the Canadian Earth System model version 5 (CanESM5; Swart et al. 2019), the second generation Earth System model from CNRM-CERFACS (CNRM-ESM2-1; Séférian et al. 2019), the Earth system model from the Institut-Pierre-Simon Laplace (IPSL-CM6A-LR; Boucher et al. 2020) and the U.K. Earth System model (UKESM1-0-LL; Sellar et al. 2020). These models were selected on the basis of data availability on the public Earth System Grid Federation repository. Each model must have at least three realizations available in both HIST and HISTNL experiments for monthly means of daily maximum 2-m temperature. Hereafter, this variable is referred to as tasmax, and the response of tasmax to the forcings included in HIST and HISTNL experiments is denoted t_{HIST} and t_{HISTNL}, respectively. Multiple realizations are required for computing individual model means to reduce the effects of internal climate variability. Per model, the maximum number of realizations of t_{HIST} and t_{HISTNL} were included (Table 1). For details on the guidance of land use application in participating CMIP6 models, see Lawrence et al. (2016), Hurtt et al. (2020). Many studies focus on warm rather than annual mean or cold temperature responses to land use. Typically, surface property changes drive a stronger radiative response in the surface energy balance with more solar

insolation (Bright et al. 2017; Duveiller et al. 2018; Hirsch et al. 2015; Lejeune et al. 2018). We therefore opt for tasmax over other temperature indices, as this expresses a stronger signal and enriches existing daily warm temperature attribution to land use (Christidis et al. 2013). Using the monthly frequency is necessary due to data availability constraints. Note that tasmax should respond to land use the strongest at the daily frequency, yet few ESMs had sufficient outputs at daily frequency for both HIST and HISTNL to consider within this analysis (see section 4).

b. Preprocessing

We compute the mean representation of individual ESM experiments before space-time aggregation. For each ESM, we compute the mean t_{HIST} and t_{HISTNL} time series across available realizations (Table 1). Reconstructing an experiment for the effects of land use (LU) on tasmax (t_{LU}) is then derived by subtracting the mean t_{HIST} by the mean t_{HISTNL}. For capturing internal climate variability, we gather as many nonoverlapping PIC realizations as possible that match the length of our study period. These PIC realizations are taken from all CMIP6 models with simulations of our target variable (t_{PIC}). From 39 models and 25 893 years of PIC data, this yielded 498 realizations of t_{PIC}. This developed from originally confining the sampling of t_{PIC} from only the ESMs in Table 1. However, this limited the quality of internal variability estimates in the regression due to small sample sizes. We therefore opt for a compromise; assuming that all CMIP6 models adequately simulate preindustrial internal variability and that the sample across many models suitably represents the internal variability in the simulations of our four ESMs, we improve the quality of internal variability estimates in our analysis.

We select the most recent 50 years of the CMIP6 historical period (1965–2014) from all observed and modeled input data. This invokes a trade-off in our analysis; although 1915–64 would more strongly express the effects of land use relative to other historical forcings, selecting the more recent period improves the reliability of the observations and land use trajectories in the Land-Use Harmonization project (LUH2) (Hurtt et al. 2020). Time series for 1965–2014 are then resampled to 5-yr nonoverlapping block means.

In anticipation of the different ESM responses to land use and in hope of preserving the local features of these responses, we avoid regridding all models and observations to a common resolution. Instead, we regrid observations to match the spatial resolution of the individual models. This avoids otherwise (i) promoting redundancy in ESM t_{LU} patterns by interpolating coarser model fields to the finer observational resolutions, or (ii) dampening the ESM signal expression of t_{LU} through resampling all model and observational fields to a coarse resolution (Lejeune et al. 2017). This means that for individual ESM detection and attribution results, the observations have been regridded to match the resolution of the ESM's simulations. For the multimodel mean results and for pooling t_{PIC} data across many models, space–time aggregation is done before computing ensemble averages, thereby neglecting regridding.

Spatial preprocessing of all model and observational data is undertaken for coarsening inputs to detection and attribution to best express the signal to noise ratios of land use forced temperature changes. All input data points to the regression are confined to land regions, excluding Greenland, the Sahara, and Antarctica. The data points are area-weighted spatial averages that are bounded depending on the spatial scale of the detection and attribution result. A first set of weights takes into account different grid cell areas by scaling their contributions to the regional average by the cosine of their latitude. At the continental scale—where separate detection results are achieved for North America, South America, Europe, Asia, Africa, and Australia—inputs are aggregated according to the Sixth IPCC Assessment Report reference land regions (AR6; Iturbide et al. 2020). This contains the local effects of land use to regions of consistent climatological characteristics. Note that we include some regions that have undergone no land use on the basis that they may experience the nonlocal effects of land use. For assessing global detectability, to lower the dimensionality of the regression, spatial averages are computed per continent, excluding Antarctica. At both the continental- and global-scale regressions, we therefore use input vectors or fingerprints that capture space–time characteristics at a finer spatial resolution than the detection result. This allows for the estimated inverse variance–covariance matrix to optimize input vectors to the regression by emphasizing regions of low variability. As these subregions have different land areas, we additionally weight their data points (the area-weighted spatial averages) based on the areas of the subcontinental AR6 or continental regional domains (Fig. 1). We confine our study to land regions to avoid the influence of less certain nonlocal effects via atmospheric feedbacks over the oceans (Davin and de Noblet-Ducoudre 2010).

The space-time dimensionality reduction yields observed and model-based input vectors of size $n = 10$ (time steps) $\times$ 6 (continents) for performing the regression analysis on a global basis. For the continental-scale detection results, n varies according to the continent being assessed. Here, n is therefore of size $10 \times \mathrm{AR6}_c$. Before the regression, we temporally center each region's time series by removing its mean across time.

c. Regularized optimal fingerprinting

We perform detection and attribution analyses on the biogeophysical effects of land use on tasmax (t_{LU}) using using regularized optimal fingerprinting (ROF) with an ordinary least squares regression type (Ribes et al. 2013). With a generalized linear regression model, this involves decomposing observed patterns of tasmax as a linear combination of tasmax response patterns simulated under different groups of forcings, taking into account internal variability. In this way, the presence of modeled response patterns, otherwise known as "fingerprints," is evaluated in the observations by making inferences on their resulting regression parameters. These regression parameters—which are slopes in the 1D case—are

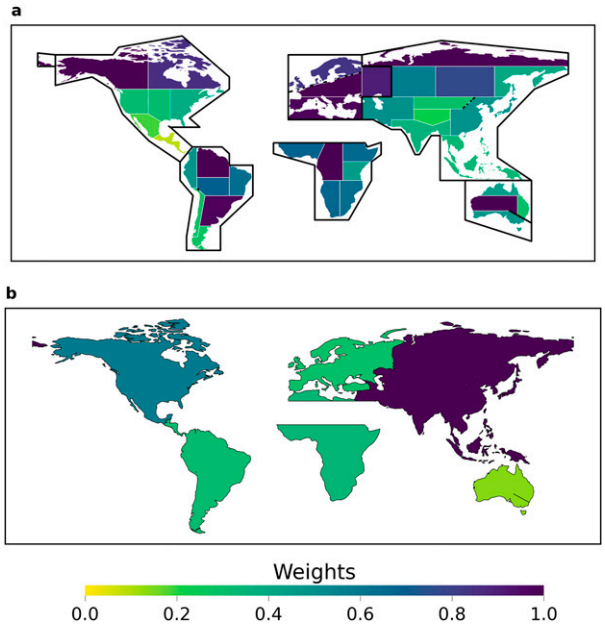

FIG. 1. Regions used for spatial aggregation and their relative weights. (a) Continental detection analyses, yielding results separately for North America, South America, Europe, Asia, Africa, and Australia, use aggregated data for land-based grid cells based on the IPCC AR6 reference regions within each continent (continental domains are outlined additionally in black), wherein each region is given a relative weight. This splits Africa north of the Sahara, as the Mediterranean AR6 region belonging to Europe encompasses land regions of both Europe and Africa. (b) For global detection analyses, yielding a single result, spatial aggregation is contained according to ESRI definitions of continental bounds, while also excluding Greenland, the Sahara and Antarctica. For the global scale, the northern tip of Africa is not considered a part of Europe but is included in the area-weighted average for data over Africa.

hereby referred to as "scaling factors" as they scale response patterns to fit the observations. This hinges on the following assumptions: (i) that models correctly represent the response patterns to different forcings, (ii) that estimates of internal variability from the models are representative of the observations, and (iii) that the fingerprints of different forcings can be reliably summed to achieve a net effect of all climate forcings in assessment of an observed change (Allen and Stott 2003). We perform this analysis for global and continental scales in both one-way and two-way regressions. The two-way regression decomposes observations as a linear combination of t_{HISTNL} and t_{LU}, yielding insights on the significance of t_{LU} in the observations. Analyses for the one-way regression case are performed separately for t_{HIST} and t_{HISTNL}, communicating whether the presence of t_{LU} changes the overall detectability of the historical climate change signal for this variable.

ROF follows a generalized linear regression model of the form, $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$. Here, $\mathbf{y}$ is a vector of observations (BE or CRU), $\mathbf{X}$ is a matrix of m column vectors of model mean simulated response patterns matching the length of $\mathbf{y}$ ($m = 1$; one-way for t_{HIST} or t_{HISTNL}, $m = 2$; two-way for t_{HISTNL} and t_{LU}), $\boldsymbol{\beta}$ is a vector of scaling factors of length m and $\boldsymbol{\varepsilon}$ is the regression residual, representing the internal variability in $\mathbf{y}$. Solving this equation for $\boldsymbol{\beta}$ requires separate covariance matrix estimates $\mathbf{C}_1$, from which a Ledoit and Wolf (2004) regularized estimate is computed for prewhitening (optimizing) all input vectors to unit variance and for computing $\boldsymbol{\beta}$, and $\mathbf{C}_2$, for estimating the variance in $\boldsymbol{\beta}$ and its 90% confidence intervals (CI; Ribes et al. 2013). For each regression, the collection of processed t_{PIC} series described above is split into two independent samples for calculating $\mathbf{C}_1$ and $\mathbf{C}_2$. The reliability of this analysis is largely dependent on the quality of t_{PIC} noise samples for precise prewhitening, that is, properly weighting the input vectors. Therefore, we iterate this analysis 500 times by shuffling the full sample of t_{PIC} series before halving it into subsamples for calculating $\mathbf{C}_1$ and $\mathbf{C}_2$. Our final results for $\boldsymbol{\beta}$ and its confidence intervals use the median of these 500 iterations.

Inferences for detecting the signal of a fingerprint, f, from $\mathbf{X}$ in $\mathbf{y}$ operate under the null hypothesis that $\boldsymbol{\beta}_f \pm 90\%\ \text{CI} = 0$, or that f is indistinguishable from internal climate variability. Rejecting this null hypothesis is therefore achieved when $\boldsymbol{\beta}_f \pm 90\%\ \text{CI} > 0$, implying an emergent signal of f in $\mathbf{y}$ relative to background climate noise. Upon detection, the general requirements of attributing a causal factor to this observed change are that the observations are consistent with model simulations including the factor and inconsistent with other plausible simulations excluding the factor. In the context of assessing t_{LU} as f, this would require that $\boldsymbol{\beta}_{t_{LU}} \pm 90\%$ CI additionally overlaps with 1 in the two-way regression case.

d. *EOF-fingerprinting tree cover change in t_{LU}*

For each ESM, t_{LU} contains forced responses to land use—both locally and nonlocally—as well as the influences of residual internal variability from the model means of the t_{HIST} and t_{HISTNL}. The nonlocal effects of land use, however, must co-occur to some extent with internal variability as, for example, a column of perturbed air from deforestation would be advected in the direction of ongoing weather patterns. This merits further exploration of the strength of local effects of land use versus the nonlocal effects, which has the side effect of assessing the noise contamination of t_{LU}. We therefore compare trends in t_{LU} against trends in tree cover (with the CMIP6 variable treeFrac) and compute the correlations between their time series (see appendix). Finally, we apply a signal to noise ratio (S/N) assessment of the fingerprint of tree cover change in t_{LU}.

Another detection and attribution methodology compliments the regression-based studies from which ROF was developed, wherein fingerprints are generated through a principal component analysis (PCA; Marvel et al. 2019; Santer et al. 2007, 2011, 2018). Whereas regression-based detection and attribution methods assess the significance of a simulated forced response in observed changes, here, PCA can evaluate the pattern similarity between ESM treeFrac and t_{LU}. This explores how t_{LU} is composed of local or nonlocal responses to land use to further understand this simulated forced response (see below). In the conventional application of this approach,

a fingerprint is taken as the leading empirical orthogonal function (EOF) of a multimodel mean estimate of the target climate variable. For a multimodel mean estimate of dimensions longitude × latitude × time, its EOF fingerprint has dimensions longitude × latitude. The statistical meaning of a leading EOF translates suitably to the climate change context. An EOF communicates the contribution or "loading" of each grid cell to the primary mode of variability in the data, the first principal component (PC). In applying PCA to climate simulations, wherein the overall shift in the data is due to historical climate change, the leading EOF therefore provides a response pattern or fingerprint of the climate change forcings in the model simulations. The presence of this fingerprint is then evaluated in an observational series of the same dimensions by projecting the observations onto the fingerprint, yielding a 1D pseudoprincipal component (pseudoPC) time series matching the length of the observed series. The pseudoPC, also known as the projection time series, measures how the observations project onto the patterns of the fingerprint. Sections of the pseudoPC with positive slopes communicate the increasing presence of the fingerprint in the observations, and vice versa. Trends of varying lengths in the pseudoPCs are therefore used to assess the signal S of a fingerprint in observations. Taking samples of likewise computed trends, but from substituting observations with independent preindustrial control simulations in the projection step, allows one to characterize noise N through assessing the standard deviation of the trends in its pseudoPCs. This allows for an estimation of S/N.

We adapt this approach to assess the agreement between changes in treeFrac and tasmax, accounting for potentially different cooling or warming responses of tasmax to deforestation. Because the units of treeFrac (grid-scale fraction) and tasmax (Kelvin) are different, we first standardize the data to avoid the influence of the different magnitudes of their ranges (Wilks 2019). For each ESM, given its data for treeFrac spanning 1965–2014, we take its fingerprint by conducting PCA using weights for latitude to find its leading EOF. We define a signal of treeFrac in t_{LU}, S, as the trend in the 50-year pseudoPC computed by projecting the ESM's mean t_{LU} onto the fingerprint of treeFrac, revealing the agreement or disagreement of tasmax responses to land use with patterns of forest cover in the treeFrac fingerprint. Reprojecting for each sample of t_{PIC} simulated by each of our four ESMs, we define noise N for an individual ESM, as the standard deviation of the resulting S_{PIC} samples from the t_{PIC} pseudoPCs. For a given ESM, N therefore captures the dispersion in the signal of treeFrac due to internal climate variability. We use N from the t_{PIC} projections to compute both S/N for the projection of t_{LU} onto the fingerprint of treeFrac as well as for each S_{PIC}. The distribution of S_{PIC}/N derived from individual ESM samples of t_{PIC} realizations defines the spurious likeness between patterns of tree cover change and changes in tasmax due to internal climate variability, thereby acting as a null distribution or control against which to compare the S/N derived from t_{LU}.

In other studies using this method, the shared variable across the fingerprint and observations chosen for projection means that a positive loading for a grid cell in the fingerprint will agree with positive trends in the observations, thereby contributing to a positive slope in the projection series. The response of t_{LU} to ΔtreeFrac, however, does not consistently occur with the same sign of change (see section 1). Therefore, we separate the treeFrac fingerprinting for tropical to temperate-south (50°S–23°N), temperate-north (23°–50°N), and boreal (50°–90°N) regions. To additionally account for the different model responses of t_{LU} to ΔtreeFrac, we also analyze inverted treeFrac datasets by scaling them by −1. This creates oppositely loaded EOFs which will capture agreements in cases where the treeFrac EOF has locations anticorrelated with trends in t_{LU}, effectively yielding signals in absolute terms and conservatively testing for the presence of treeFrac in t_{LU} or t_{PIC}.

3. Results

a. Detection and attribution

1) The detectability of t_{LU} in the two-way regression

The response of tasmax to land use forcings, t_{LU}, is inconsistently detectable in observations (BE) across ESMs at both continental and global scales, but its detectability is internally consistent across scales for specific ESMs (Fig. 2). For CanESM5, only South America contains a detectable t_{LU} signal (Fig. 2b), and at the global scale t_{LU} is undetected and indistinguishable from internal climate variability (Fig. 2c). Likewise, for UKESM1-0-LL, Africa is the only continent with a detectable signal (Fig. 2k), yet, globally, t_{LU} is undetectable (Fig. 2l). Across ESMs, Africa is the most consistent continent with a detectable t_{LU} signal, as all but CanESM5 simulate a detectable signal there. CNRM-ESM2-1 and IPSL-CM6A-LR have the most detectable signal of t_{LU}. At the global scale, their scaling factors are consistent with unity (Figs. 2f,i), which is apparent in their continental scale results in that they both simulate a detectable t_{LU} signal in at least four of six continents (Fig. 2e). The inconsistency in the signal of t_{LU} across ESMs is confirmed in the multimodel mean case (Figs. 2m–o), where no continents contain a significant expression of t_{LU} and its signal is undetectable globally. All other forcings (t_{HISTNL}) are nearly unanimously detected in all continents, across scales and ESMs. When repeating the analysis with CRU as an alternative set of observations, some of the continents and ESMs with detectable signals are different (Fig. A1), where, for example, IPSL-CM6A-LR loses its detectability of t_{LU}. However, the same outcomes around (i) inconsistent signals across ESMs, (ii) internally consistent signals propogating through continental and global scales for individual ESMs and (iii) null results for the multimodel mean case all remain true.

2) Does t_{LU} affect the detectability of the historical climate change signal?

The response patterns of tasmax to the forcings in HIST and HISTNL, t_{HIST} and t_{HISTNL}, represent alternative fingerprints of the historical climate with and without the effects of land use, and their consistency with observations (BE) is sometimes affected by the presence of t_{LU} in t_{HIST} (Fig. 3). All ESMs produce t_{HIST} and t_{HISTNL} signals that are detectable in the observations at the global scale, and Australia is

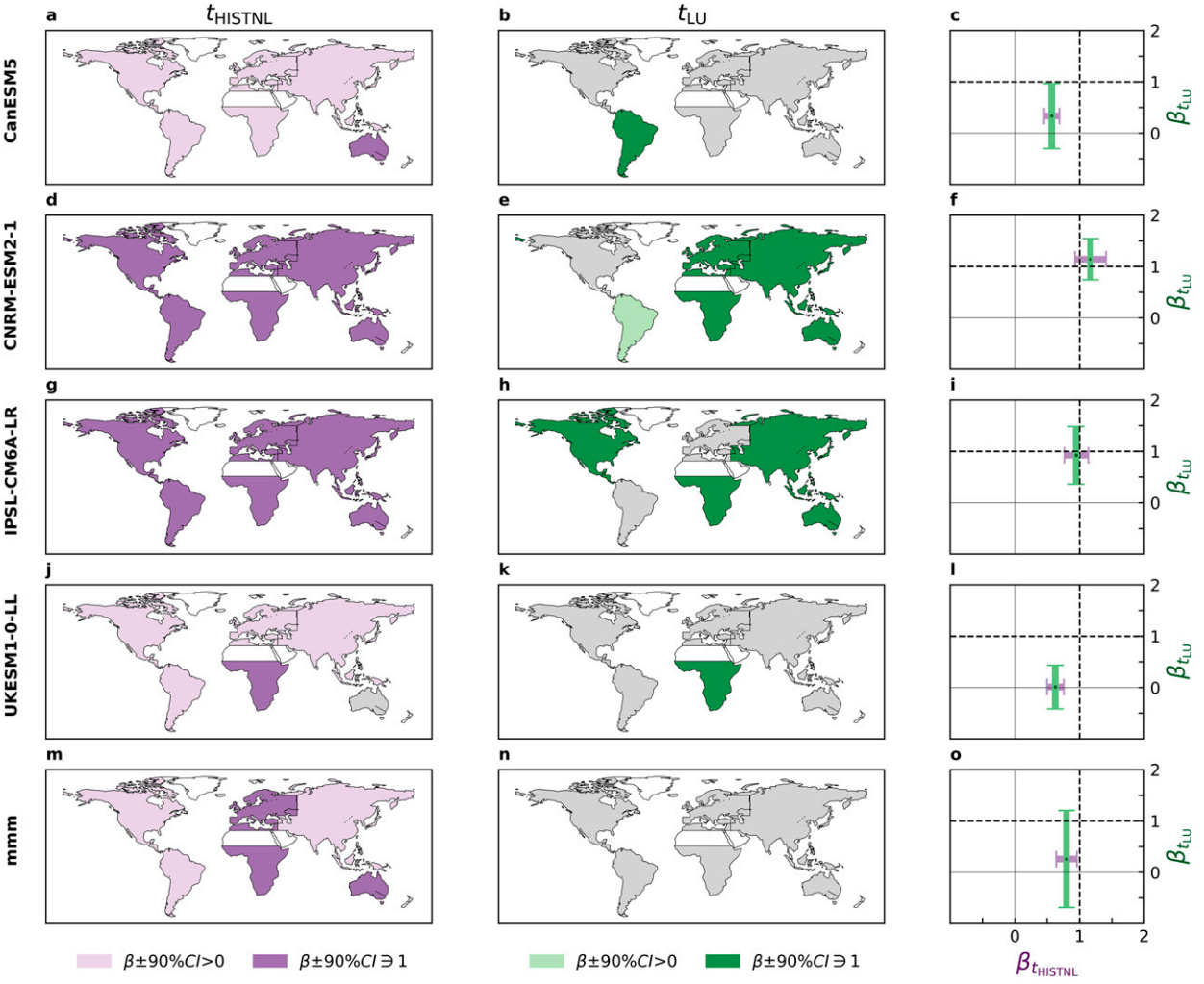

FIG. 2. Continental and global detection and attribution results for two-way analyses with observations from Berkeley Earth. The continental scale detectability of (a),(d),(g),(j),(m) t_{HISTNL} (purple) and (b),(e),(h),(k),(n) t_{LU} (green) fingerprints. Light colors represent scaling factor results implying detection ($\boldsymbol{\beta}_f \pm 90\%$ CI > 0), dark colors represent scaling factor results consistent with unity ($\boldsymbol{\beta}_f \pm 90\%$ CI $\ni 1$), and gray implies an undetected signal. (c),(f),(i),(l),(o) The global detectability of t_{HISTNL} (purple; x axes) and t_{LU} fingerprints (green; y axes) through their scaling factors and confidence intervals. In these panels, dark, hashed lines are drawn to mark unity and light gray lines mark 0. Multimodel mean results are labeled as "mmm".

the only continent where, in some instances, these signals are not detected. For the ESMs that produce a detectable signal of t_{LU} in the two-way regression, the presence of t_{LU} in t_{HIST} improves the detectability of t_{HIST} relative to t_{HISTNL} for some continents in CNRM-ESM2-1 (Figs. 3d,e), IPSL-CM6A-LR (Figs. 3g,h), and UKESM1-0-LL (Figs. 3j,k). However, for the multimodel mean results at both continental and global scales, there are no differences in the detectability of t_{HIST} and t_{HISTNL}. This feature is consistent for analyses with both observational datasets (Figs. 3m–o, A2 m–o).

b. *S/N of treeFrac in t_{LU}*

The *S*/*N* of treeFrac in t_{LU} is generally weak (Fig. 4). The *S*/*N* for t_{LU} exists mostly around the center of S_{PIC}/N distributions (the control), and, with the exception of the tropics in CanESM5, they only reach so far as the 66th percentile of the control distribution. CNRM-ESM2-1 and IPSL-CM6A-LR have the best pattern similarity between treeFrac and t_{LU}, as the *S*/*N* for t_{LU} always exceeds the 66% percentile of the control distribution (Figs. 4b,c), which is consistent with these ESMs simulating t_{LU} with the best detectability using ROF for detection and attribution. In line with the two-way regression results, CanESM5, IPSL-CM6A-LR and UKESM1-0-LL have the best pattern matching between treeFrac and t_{LU} in the tropics to temperate-south latitude band (Figs. 4a,d), which contains the continents for which they report detection using both BE and CRU as observations (with exception to detecting UKESM1-0-LL in CRU). In all ESMs except for CNRM-ESM2-1, the tropical to temperate-south latitude band shows the likeliest signal of treeFrac in t_{LU}, implying that

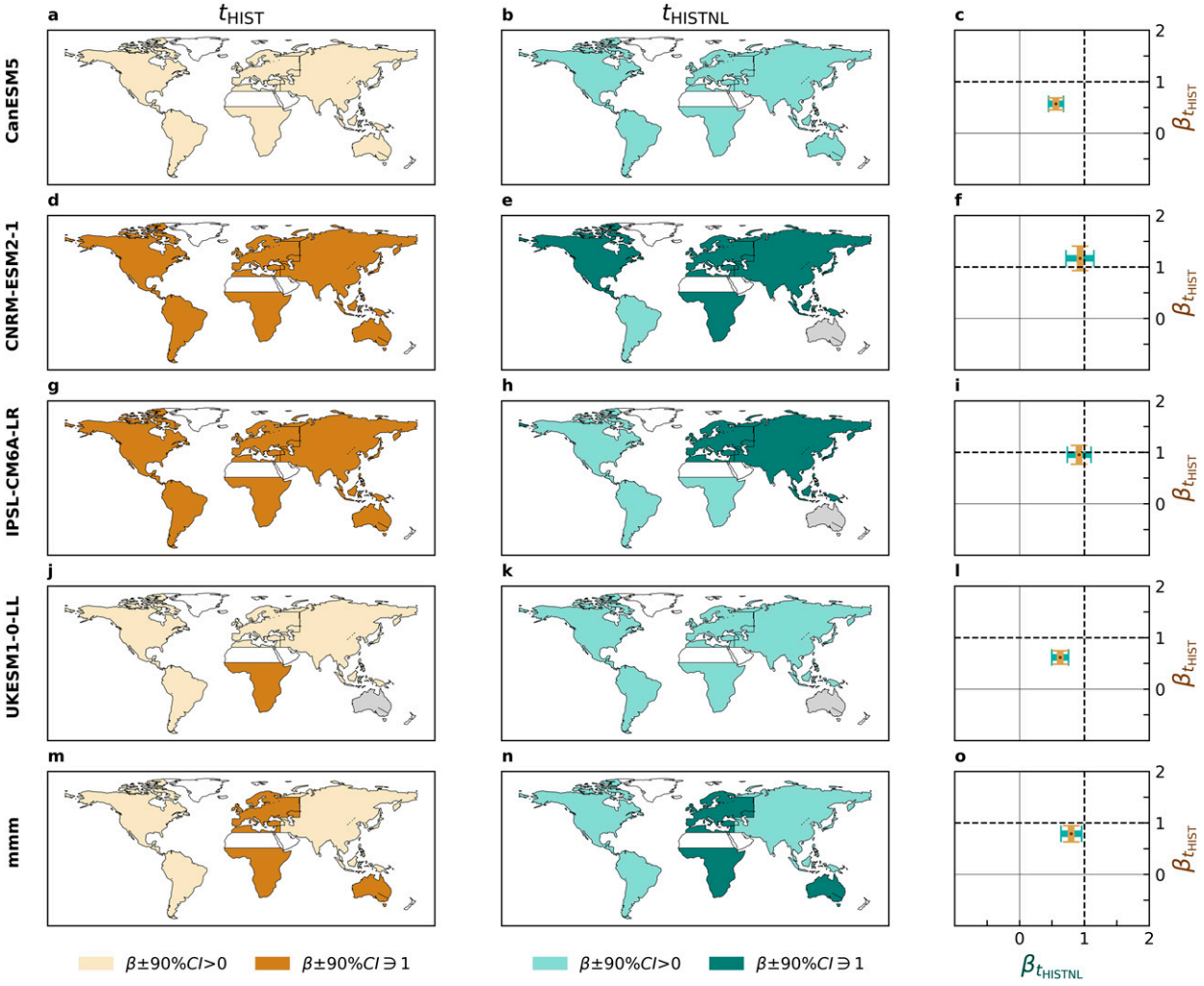

FIG. 3. As in Fig. 2, but for one-way analysis results reported separately for t_{HIST} (browns) and t_{HISTNL} (blue-greens).

the local effects of land use are most strongly expressed in these regions. For the tropical to temperate-south latitude band, it is contradictory that CNRM-ESM2-1 produces the lowest *S*/*N* of treeFrac in t_{LU} while t_{LU} for this ESM is also consistently detectable in southern continents across observational datasets using ROF (see section 4). Since CNRM-ESM2-1 has the smallest sample of t_{PIC} realizations from which to compute *N*, it is possible that the sampled *N* is considerably larger than its true value, suppressing *S*/*N* for this ESM. In the other latitude bands, the lesser pattern similarity between treeFrac and t_{LU} may imply that noise or nonlocal effects of land use are dominant. Overall, these results emphasize the weak signal case of t_{LU} because the treeFrac *S*/*N* is consistent with preindustrial control runs, namely that t_{LU} probably contains spurious trends from internal climate variability.

4. Discussion and conclusions

We find (i) a null detection and attribution result for t_{LU}—the response of the local warmest monthly mean of daily maximum temperature to land-use and land-cover changes (land use)—using regularized optimal fingerprinting (ROF) in a two-way regression and (ii) no evidence that land use improves the detection of the historical forced response pattern. Even though results vary across ESMs, for the multimodel mean analysis, t_{LU} is consistently undetected and indistinguishable from internal climate variability in both two-way and one-way regressions and across observational datasets. CNRM-ESM2-1 is the exceptional model in the ensemble. If taken alone as a means of communicating the significance of the t_{LU} signal, its detectability across scales and observational datasets, reinforced to a minor degree by its pattern matching in the EOF-fingerprinting exercise, might be convincing evidence of detection. Alongside its detection results, opposite to CanESM5, the t_{LU} response pattern of CNRM-ESM2-1 is the most aligned with treeFrac changes north of the tropical to temperate-south latitude band. Altogether, this may imply a greater expression of local effects of land use in this ESM, thus agreeing with earlier assessments of the scale dependence of the land use signal (Chen and Dirmeyer 2020).

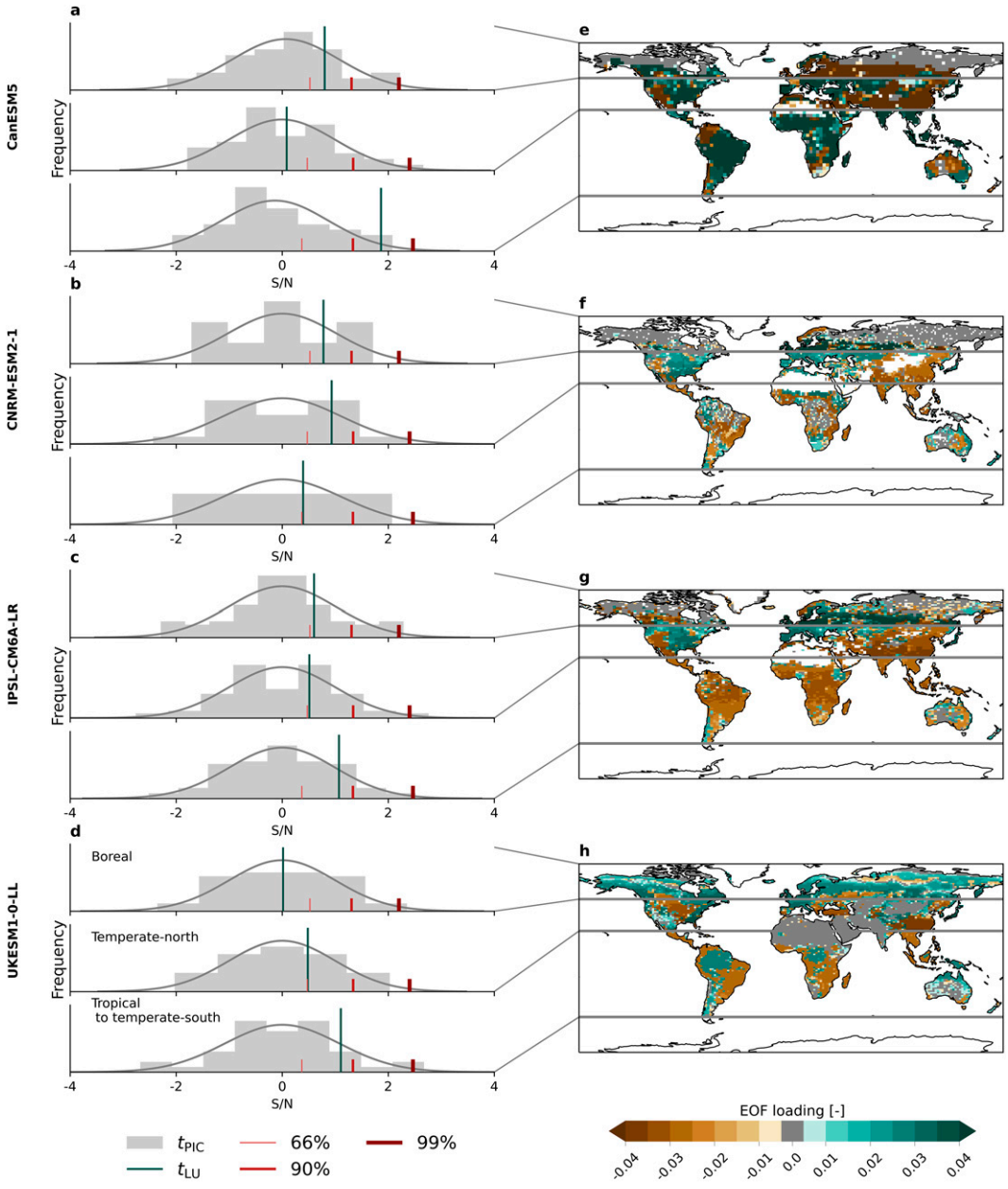

FIG. 4. (a)–(d) Signal-to-noise ratios (S/N) of (e)–(h) EOF fingerprints of treeFrac in t_{LU} vs t_{PIC} realizations for each model. Histograms showing the distribution of S_{PIC}/N across boreal, temperate-north, and tropical to temperate-south latitudinal bands are plotted behind S/N for t_{LU}. The S_{PIC}/N ratios derived from t_{PIC} are sampled for each latitude band across all four models in the detection and attribution analysis for computing the following quantiles for signal detectability: 66th, 90th, and 99th.

Notably, excluding IPSL-CM6A-LR, ESMs in our ensemble simulate above average mean cooling responses to an idealized deforestation experiment (Table 1; Boysen et al. 2020). However, as discussed below, the t_{LU} response pattern for CNRM-ESM2-1, like the rest of the ensemble, should be viewed with caution given the difficulties in assessing why its parameterization scheme should be trusted over others for computing surface water and energy fluxes due to land use.

We also caution against extending comparisons between the results of our ROF detection and attribution and the EOF-fingerprinting analyses, as these methods address different questions. ROF for detection and attribution asks whether or not t_{LU} is significantly detectable in observed changes in tasmax. EOF-fingerprinting, however, seeks to test how patterns of change in in treeFrac agree with t_{LU}. Although we can infer from a strong S/N of treeFrac in t_{LU} that this could

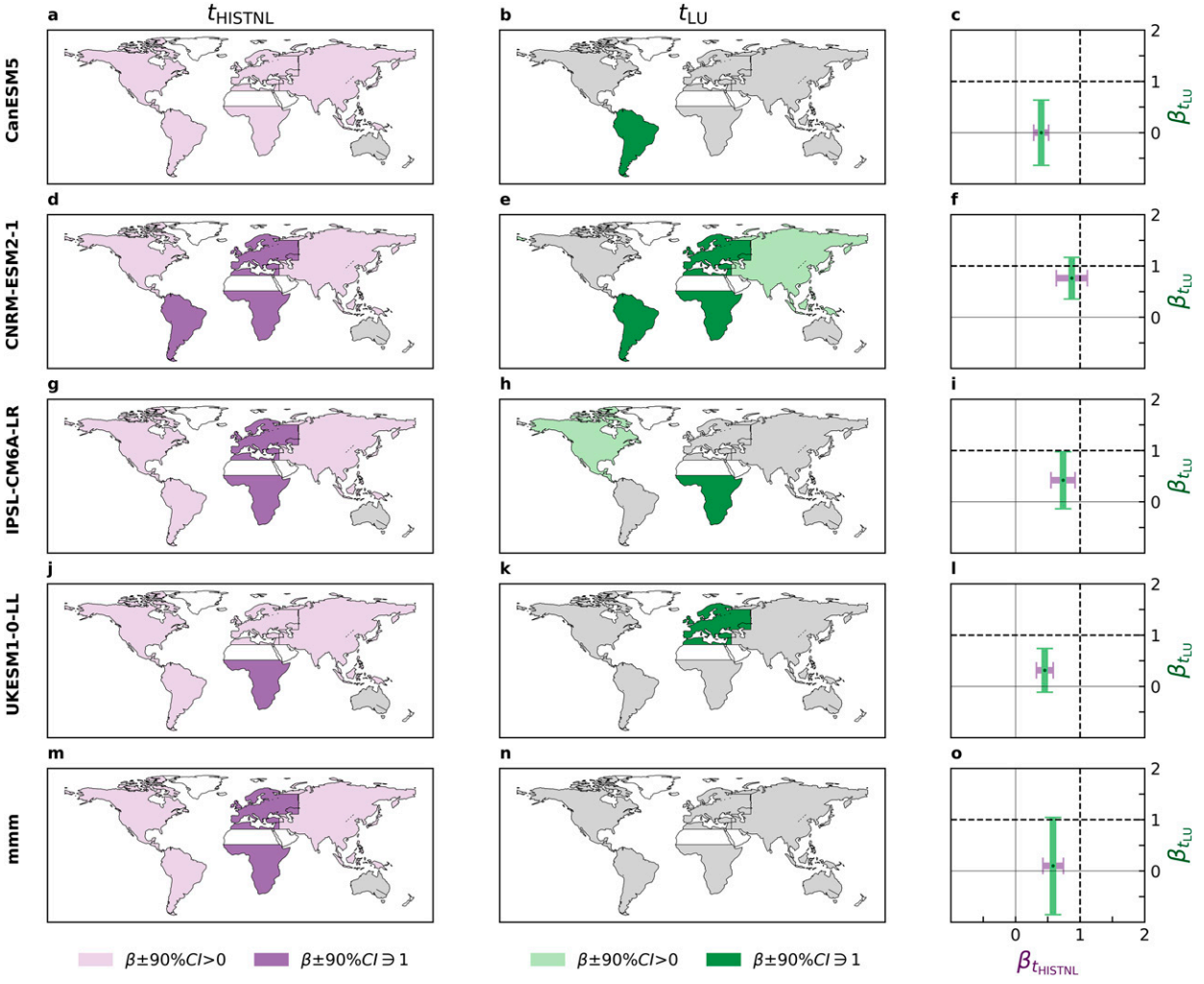

FIG. A1. As in Fig. 2, but for observations from CRU.

imply better signal detectability of t_{LU} in ROF, this only comments on the potential signal expression in t_{LU}. It says nothing about t_{HISTNL}, the observations, and how they are expressed after space–time truncation and optimization with t_{PIC} variance–covariance matrices. As well, this inference assumes that land use will have a stronger signal relative to all of historical climate change if it is mostly made up of local rather than nonlocal effects. This assumption may not be true in ROF if regions with strong nonlocal effects are not aligned with those of high variance in t_{PIC} simulations from the entire CMIP6 ensemble.

The null outcome of this analysis is acceptable for a number of reasons. For some ESMs, the inclusion of t_{LU} in t_{HIST} in the one-way regression analysis improves the consistency of modeled response patterns with observations. The opposite does not occur, wherein the inclusion of t_{LU} reduces the likeness between observations and models (Fig. 3, Fig. A2). Yet, the cancellation of t_{LU} signals in the multimodel mean underscores model uncertainties that are too high for this set of forcings relative to internal climate variability to allow for formal detection and attribution, yielding null results across scales and observational datasets. This questions the assumption that these models are all valid in representing land use. The weak *S*/*N* assessment of treeFrac in t_{LU} and the fact that the regions with the strongest trends in t_{LU} do not consistently align with regions of high land use (Figs. 4, B1) supports that factorial estimates of t_{LU} may be dominated by nonlocal processes and/or internal climate variability. The t_{LU} response pattern in our study period is a weak signal in comparison to the earlier historical period due to the stronger effects of land use relative to other forcings in the early twentieth century. This is apparent in more consistent correlations and slightly greater *S*/*N* when repeating these analyses for 1915–64 (Figs. B2, B3). Finally, we tested multiple other approaches to our detection and attribution analysis, which all provided similar results. These tested alternatives included (i) separately assessing summer and winter seasonal tasmax; (ii) making selections of regions of high land use with respect to LUH2-prescribed tree cover change before aggregating with AR6-weighted means for continental analyses; (iii) limiting the inclusion of AR6 regions to those of

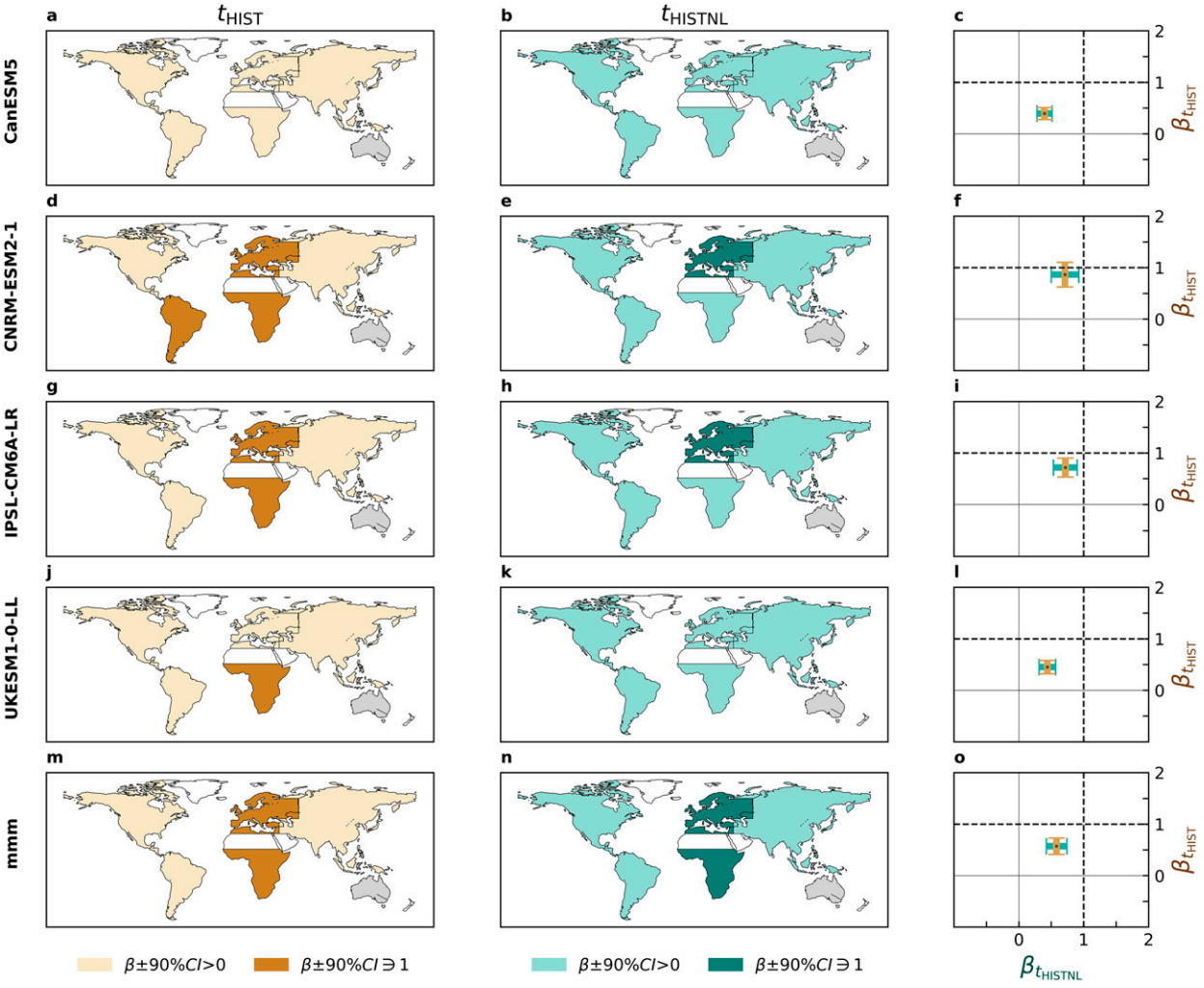

FIG. A2. As in Fig. 3, but for observations from CRU.

considerable land use; (iv) testing other temporal aggregation methods, such as either computing 10-yr means or eliminating the temporal dimension of the analysis altogether by representing each region by its difference between the means of 1965–89 and 1990–2014; and (v) restricting our t_{PIC} samples to those simulated by the four ESMs involved in this study.

Relative to the analysis of daily temperature extremes in Christidis et al. (2013), it is understandable that we find a weaker signal case using the monthly time scale. As they discuss regarding their attribution of warm day extremes relative to the insignificant effect of land use on cold and night extremes, ample solar insolation is required to drive the expression of the radiative effects of land use that control temperature changes. This includes albedo controlling the available energy and the partitioning of this energy into turbulent heat fluxes. Therefore, on the one hand, in dulling the temporal scale of this analysis from daily to monthly resolution, the peak assessment of these radiative effects on near-surface temperatures is lost, effectively weakening our version of the land use signal. On the other hand, the signal to noise ratio of t_{LU} in our study is benefited by analyzing the monthly resolution, which is less influenced by internal variability than daily temperatures. However, comparisons between this study and Christidis et al. (2013) can never be perfect. We isolated spatial aggregations to land regions and avoided ocean domains, while they used a T4 truncation that included ocean domains. Repeating our analysis to include the oceans, however, invites potentially greater inconsistencies in ESM-specific advection processes and ocean–atmospheric coupling. Although we choose a likewise 50-yr-long study period, ours begins 15 years later, which means we are including different land use trends in some regions. This study is therefore not a correction to their detection and attribution results, but rather an addition to the outlook of the detectability of land use on warm temperature indices.

Our detection and attribution analysis has a number of limitations. Our observational datasets, CRU and BE, likely do not fully ingest historical tasmax responses to land use. By convention, weather stations do not measure areas before and after land use transitions such as deforestation. Moreover, in CRU, tasmax is estimated using daily mean temperature and

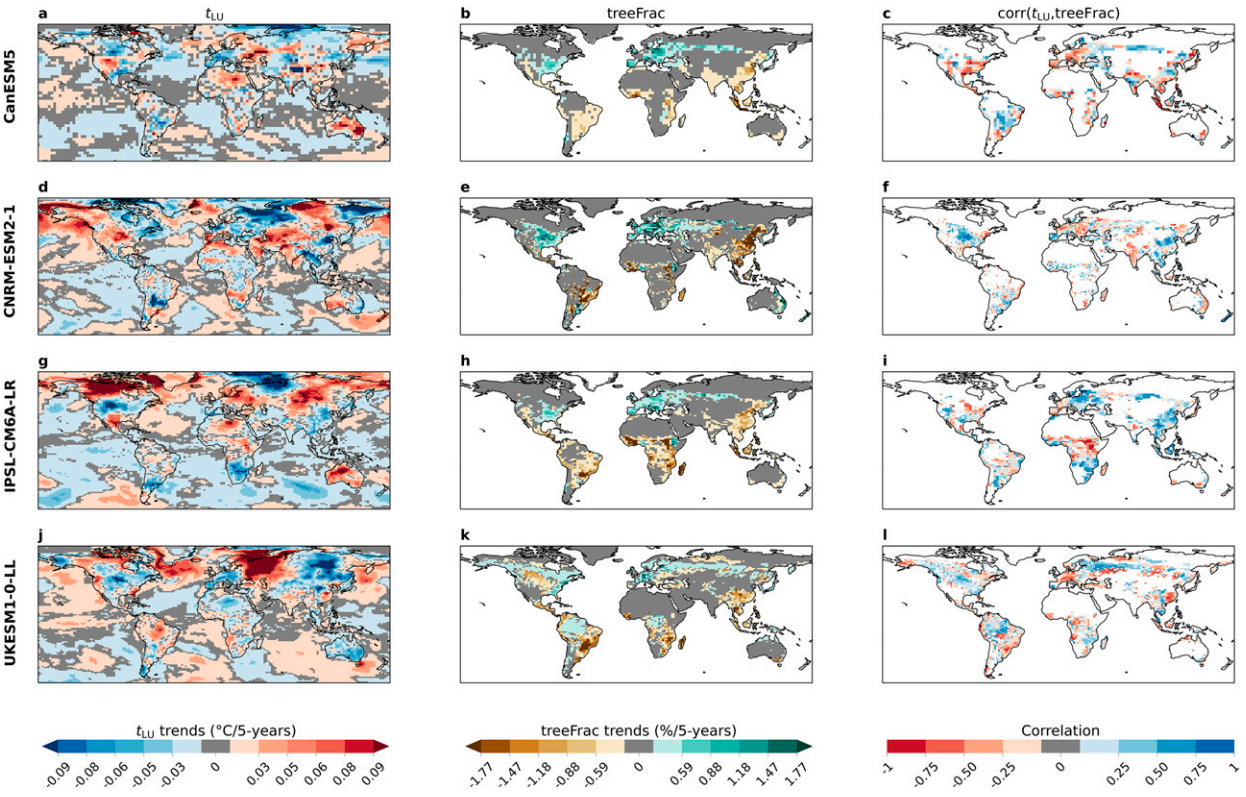

FIG. B1. Trends in (a),(d),(g),(j) 5-yr mean t_{LU} and (b),(e),(h),(k) treeFrac and (c),(f),(i),(l) the correlations between t_{LU} and treeFrac for 1965–2014.

the diurnal temperature range (Thiery et al. 2020). The same occurs on the model side, wherein daily tasmax is computed at the model time step. As these time steps do not necessarily align with the timing of the peak temperature, the daily maximum temperature might be underestimated by the ESMs. However, as ESM time steps typically are between 30 min and 1 h, we assume that the potential effect is limited. Model data availability constraints also limit the quality of this analysis. There are too few ESMs involved in this study to robustly assess t_{LU} detectability across the full CMIP6 ensemble. As our four-ESM ensemble displays weak and spatially varying trends in t_{LU} and different applications of land use (Figs. B1, B2), encompassing all CMIP6 model outcomes could provide a more reliable understanding of the significance of land use as a historical forcing. A limited availability of ensemble members for the HISTNL experiment leaves our estimate of t_{LU} contaminated by internal variability (Table 1). Similarly, restricting preindustrial control runs to only those simulated by the ESMs involved in this study would cover too few years for adequately sampling the space-time variance possible via internal climate variability. Therefore, we increase this sample size as a compromise for potentially including internal variability estimates that are ill-fitted to our four ESMs. In this way, the assumed internal variability is representative of a CMIP6 multimodel mean variability, although that internal variability varies largely among CMIP6 models in terms of amplitude and characteristics (Ribes et al. 2021). Accounting for these between-model differences may inflate the variance of internal climate variability for some regions, thereby reducing the likelihood of detection and making our detection test more conservative. Finally, we apply an earlier subset of the ROF regression toolbox by using an ordinary least squares (OLS) fitting of model predictands to observed predictors. In the detection and attribution context, this ignores uncertainty in the model fingerprints by minimizing the vertical distances of the residuals in the regression, which is best imagined in the one-way case (Fig. 2; Allen and Stott 2003). Methodological advances from Allen and Stott (2003) came by the use of total least squares (TLS) fitting, which assumes the fingerprints to be a noised version of their true pattern and involves fitting the regression by minimizing distances perpendicular to the best fit line. In our case, the OLS regression acts as a conservative approach to searching for a weak signal. Our inability to reject the null hypothesis that t_{LU} is indistinguishable from noise in the multimodel mean case using OLS implies that further work with the more robust TLS fitting is not yet necessary.

Our EOF-fingerprinting analysis demonstrates the presence of internal climate variability and/or nonlocal effects of land use in t_{LU} and benefits from its basis in the treeFrac fingerprint. This approach lends a pattern assessment framework from earlier studies assessing the presence of modeled fingerprints in observations for the same variable (Marvel et al. 2019; Santer et al. 2018). We deviate from this by evaluating

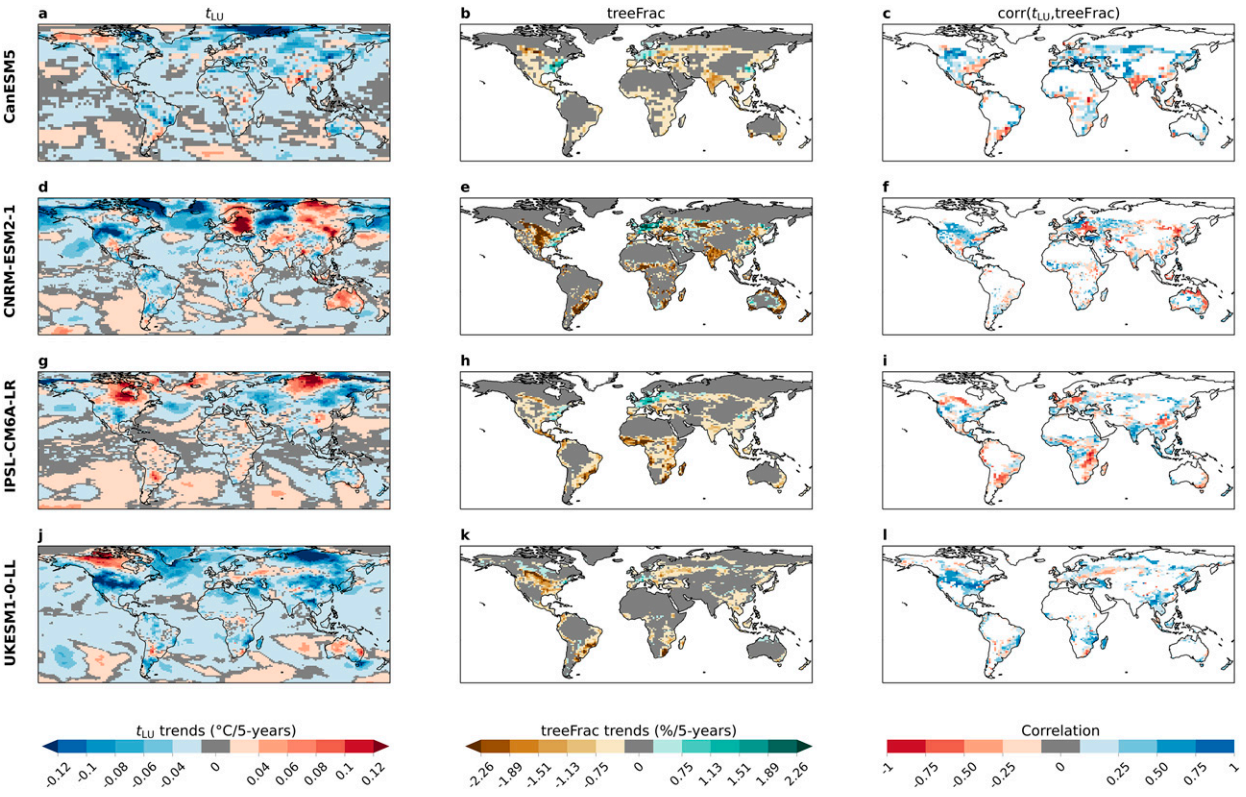

FIG. B2. As in Fig. B1, but for 1915–64.

the likeness between the tree cover fingerprint, a prominent variable in the land use forcing (Lejeune et al. 2017), and t_{LU}. In the context of the effects of land use—originating local to the site of land perturbation yet affecting other regions due to advection and nonlocal feedbacks (Winckler et al. 2017)—the pseudoPC series from projecting t_{LU} onto the treeFrac fingerprint indicates the degree to which patterns of change in t_{LU} match those of tree cover. Therefore, this also indicates if our tasmax metric for 1965–2014 is dominated by local or nonlocal processes. Since the *S*/*N* of treeFrac in t_{LU} is mostly consistent with the control distribution, it cannot be claimed that t_{LU} is composed mostly of local responses to land use. Instead, t_{LU} is likely influenced by internal climate variability or the nonlocal effects of land use. Importantly, t_{LU} also contains forced responses to other types of land use, such as agricultural expansion and urbanization. Therefore, features in t_{LU} that might appear to be nonlocal feedbacks to forest cover change will sometimes be local responses to other land perturbations. However, tree cover changes are known to produce a stronger climatic signal than crop cover expansion (Lejeune et al. 2017), and our results did not change much when testing this approach for cropFrac. The presentation style of these results is typical in detection and attribution studies. A metric derived from different scenarios, be it trends in or correlations between response patterns and a historical mean series, is compared against a control distribution derived from preindustrial control simulations (Qian and Zhang 2015; Padrón et al. 2020; Grant et al. 2021). However, applying the same approaches in this study would not achieve much given the weak and inconsistent signal of t_{LU}. In our approach, despite that there is also implicit uncertainty behind the land cover change data that the ESMs apply, on an individual ESM level, the patterns of the treeFrac fingerprints act as concrete basis against which to assess t_{LU}.

There are many sources of uncertainty in t_{LU} which control its inconsistency across ESMs. ESMs respond differently at the local scale to land use, owing to their individual systems for representing vegetation (Sellar et al. 2020; Séférian et al. 2019; Swart et al. 2019; Boucher et al. 2020) and for computing surface flux exchanges for water and energy cycles (de Noblet-Ducoudré et al. 2012; Lejeune et al. 2020). For example, it is interesting to note that the models without dynamic vegetation schemes have the best expressed signals of land use. These local responses to land use are then propagated nonlocally by model-specific advection and atmospheric feedback processes, which are inconsistent across ESMs in this study and in earlier research (Pitman et al. 2009; Winckler et al. 2019c; Hirsch et al. 2015). One such process includes cloud cover simulation, a known limitation to modern ESMs which has a demonstrated cooling bias from overestimating cloud shortwave radiative effects (Miao et al. 2021). Future work could distinguish the drivers of model uncertainties between their surface flux responses or atmospheric feedbacks by comparing the coupled simulations applied here to the

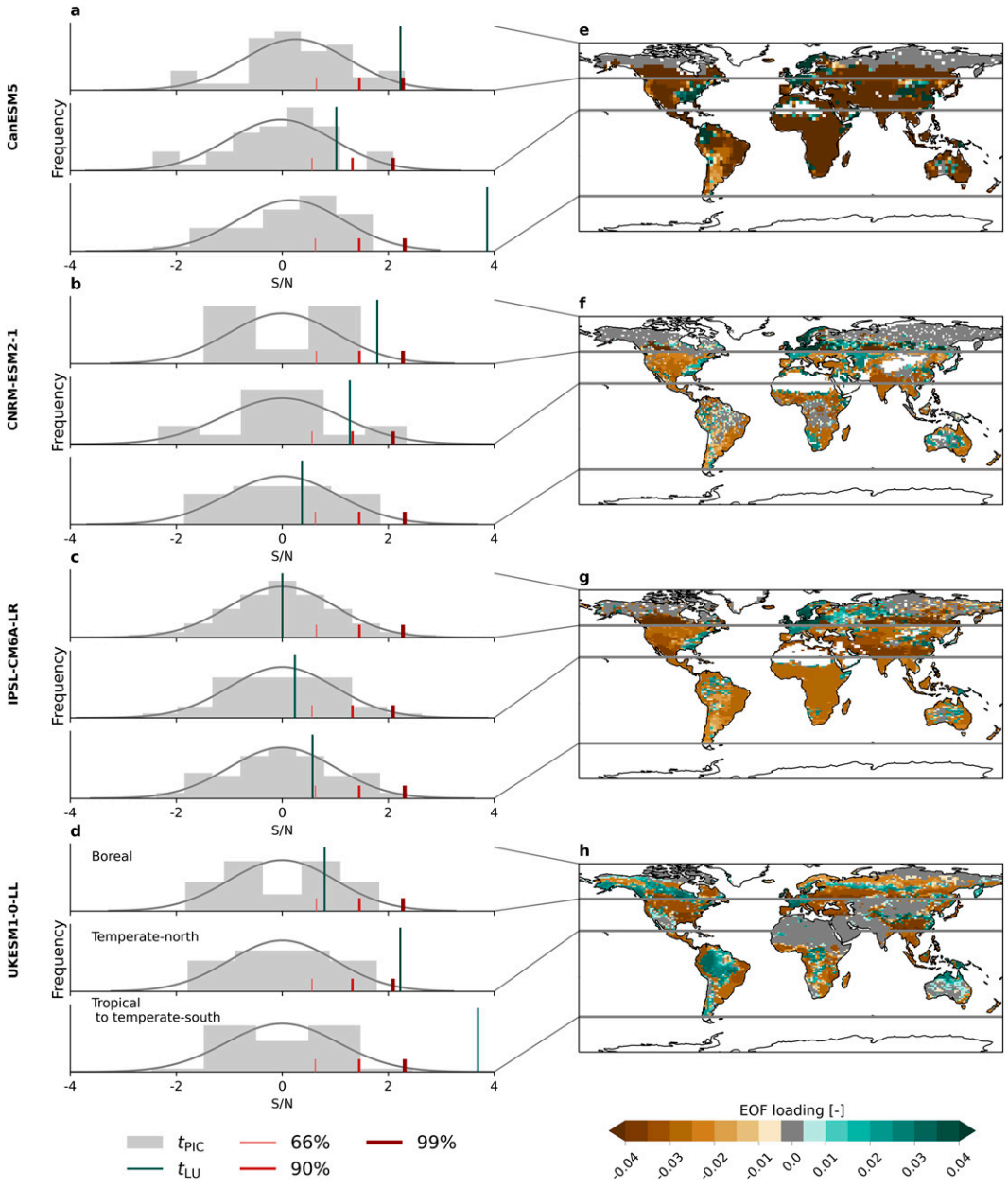

FIG. B3. As in Fig. 4, but for 1915–64.

historical Land Use Model Intercomparison Project land-only simulations with and without land use (Lawrence et al. 2016). Meanwhile, it is beyond the scope of this study to assess what drives these differences in temperature outcomes in detail. This would require factorial experiments with incremental changes to understand at a process level what may drive differences between ESMs, a challenging task due to cascading interactions. Finally, the basis for these t_{LU} response patterns, the land use patterns themselves, are differently applied in the models owing to model-dependent land parameterization and tiling schemes. This includes how the number of PFTs vary between the land surface schemes and whether or not they include land management. The background information from LUH2, which prescribes the locations and timing of historical land use, also has demonstrated uncertainties and limitations (Hartung et al. 2021; Prestele et al. 2017).

To conclude, our detection and attribution analysis found an insignificant influence of biogeophysical land use on the daily maximum temperatures averaged over the warmest month, t_{LU}. This temperature metric is the best available climate variable across CMIP6 ESMs for expressing the influence of this set of forcings on warm temperatures given the data requirements of a detection and attribution exercise. A null result was found confidently by assessing t_{LU} across different observational datasets,

ESMs, spatial scales and analytical approaches to detection and attribution. This result, when viewed through the lens of the extensive research showing the difficulty of representing land use impacts in ESMs, is not necessarily surprising. This study clarifies that, given existing model imperfections, it is not yet possible to find a significant fingerprint of the effects of land use on warm temperatures via traditional detection and attribution methods that rely on accurately modeled response patterns.

Acknowledgments. L. Grant and W. Thiery designed the analysis. L. Grant performed the analysis and wrote the manuscript. L. Gudmundsson and A. Ribes guided the detection and attribution methodology. D. M. Lawrence is a primary coordinator of the Land Use Model Intercomparison Project. All authors provided guidance on the analysis and contributed to writing the manuscript. D. M. Lawrence is supported by the National Center for Atmospheric Research, which is a major facility sponsored by the NSF under Cooperative Agreement 1852977, and by the U.S. Department of Energy, Office of Biological and Environmental Research Grant DE-FC03-97ER62402/A0101. R. Séférian acknowledges the European Union's Horizon 2020 research and innovation program under grant agreement No. 101003536 (ESM2025—Earth System Models for the Future). This study was supported by the LAMACLIMA project, part of AXIS, an ERA-NET initiated by JPI Climate, and funded by BELSPO (BE, Grant No. B2/181/P1) with co-funding by the European Union (Grant No. 776608). E. Robertson is supported by the Joint U.K. BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101). IPSL-CM6A-LR experiments were run on the HPC resources of TGCC under the allocations 2016-A0030107732, 2017-R0040110492 and 2018-R0040110492 (project gencmip6) provided by GENCI (Grand Équipement National de Calcul Intensif) to conduct CMIP6 projects at IPSL. The computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation—Flanders (FWO) and the Flemish Government—department EWI.

Data availability statement. Data from models is available under the CMIP6 archive on the Earth System Grid Federation search engine here: https://esgf-data.dkrz.de/search/esgf-dkrz/. Observations from Berkeley Earth can be found here: http://berkeleyearth.org/data/. Observations from the Climatic Research Unit can be found here: https://catalogue.ceda.ac.uk/uuid/c26a65020a5e4b80b20018f148556681. Code for this analysis is here: https://github.com/VUB-HYDR/2022_Grant_etal_JOC.

APPENDIX A

Further Detection and Attribution Results

For detection and attribution with ROF, the sensitivity of our results to different observational datasets was tested. While individual models differed in their signal detectability with CRU as an observational reference product, the multimodel mean produced consistent results as with BE (Figs. A1, A2).

APPENDIX B

Trends and Correlations

There is little consistency in the trends in t_{LU} across ESMs, and in some regions of model agreement there has been no land use (Fig. B1). CNRM-ESM2-1 and IPSL-CM6A-LR have similar patterns of strong nonlocal cooling and warming trends in northern Eurasia along the border of the Arctic Ocean (Fig. B1d,g), yet these regions are considerably north of grid cells with treeFrac changes (Fig. B1e,h). In the same region, however, UKESM1-0-LL has opposite trends in t_{LU} (Fig. B1j). Alternatively, the same models disagree in northern Canada, a region likewise void of land use, where IPSL-CM6A-LR shows the strongest warming trends among all models and CNRM-ESM2-1 simulates cooling in the east of the country. Generally, CanESM5 has the smallest trends in t_{LU} (Fig. B1a).

Modeled treeFrac change acts as the best indicator of land use in the ESMs because of the historical dominance of converting forested land to agriculture, therefore reflecting a rough composite of cropland trends (Fig. B1). Generally, the ESMs simulate a reciprocal regrowth of forests in the Northern Hemisphere and deforestation in the Southern Hemisphere, with exceptions in East Asia (Figs. B1b,e,h,k) and, for UKESM1-0-LL, the Great Plains in North America (Fig. B1k). However, modeled patterns of t_{LU} do not share this reciprocity, which is evident in the mixed, regional correlation and anticorrelation across ESMs (Figs. B1c,f,i,l). The regions of best consistency in the correlation between t_{LU} and treeFrac are eastern North America and Europe. In Africa, the ESM application of treeFrac change is as inconsistent as the t_{LU} responses; ISPL-CM6A has the most widespread deforestation in central Africa, producing moderate warming centrally and cooling in southern Africa (Figs. B1g,h), yet here the remainder of the models simulate lesser deforestation and variable t_{LU} responses.

As mentioned in section 2b, the relative strength of the LULCC forcing compared to other historical forcings is likely enhanced in an earlier period covered by our observational datasets (1915–64). This is evident in some models' higher regional consistency in correlations between t_{LU} and treeFrac, such as UKESM1-0-LL (Fig. B2l). Similarly, when repeating the *S/N* analysis of EOF fingerprints of treeFrac in t_{LU} for 1915–64, CanESM5 and UKESM1-0-LL show improved *S/N*, implying a stronger pattern similarity between treeFrac and t_{LU} for this period (Figs. B3a,d).

REFERENCES

Allen, M. R., and P. A. Stott, 2003: Estimating signal amplitudes in optimal fingerprinting, Part I: Theory. *Climate Dyn.*, **21**, 477–491, https://doi.org/10.1007/s00382-003-0313-9.

Boucher, O., and Coauthors, 2020: Presentation and evaluation of the IPSL-CM6A-LR climate model. *J. Adv. Model. Earth Syst.*, **12**, e2019MS002010, https://doi.org/10.1029/2019MS002010.

Boysen, L. R., and Coauthors, 2020: Global climate response to idealized deforestation in CMIP6 models. *Biogeosciences*, **17**, 5615–5638, https://doi.org/10.5194/bg-17-5615-2020.

Breil, M., and Coauthors, 2020: The opposing effects of reforestation and afforestation on the diurnal temperature cycle at the surface and in the lowest atmospheric model level in the European summer. *J. Climate*, **33**, 9159–9179, https://doi.org/10.1175/JCLI-D-19-0624.1.

Bright, R. M., E. Davin, T. O'Halloran, J. Pongratz, K. Zhao, and A. Cescatti, 2017: Local temperature response to land cover and management change driven by non-radiative processes. *Nat. Climate Change*, **7**, 296–302, https://doi.org/10.1038/nclimate3250.

Chen, L., and P. A. Dirmeyer, 2020: Reconciling the disagreement between observed and simulated temperature responses to deforestation. *Nat. Commun.*, **11**, 202, https://doi.org/10.1038/s41467-019-14017-0.

Christidis, N., P. A. Stott, G. C. Hegerl, and R. A. Betts, 2013: The role of land use change in the recent warming of daily extreme temperatures. *Geophys. Res. Lett.*, **40**, 589–594, https://doi.org/10.1002/grl.50159.

Davin, E. L., and N. de Noblet-Ducoudre, 2010: Climatic impact of global-scale deforestation: Radiative versus nonradiative processes. *J. Climate*, **23**, 97–112, https://doi.org/10.1175/2009JCLI3102.1.

De Hertog, S. J., and Coauthors, 2022: The biogeophysical effects of idealized land cover and land management changes in Earth system models. *Earth Syst. Dyn.*, **13**, 1305–1350, https://doi.org/10.5194/esd-13-1305-2022.

de Noblet-Ducoudré, N., and Coauthors, 2012: Determining robust impacts of land-use-induced land cover changes on surface climate over North America and Eurasia: Results from the first set of LUCID experiments. *J. Climate*, **25**, 3261–3281, https://doi.org/10.1175/JCLI-D-11-00338.1.

Duveiller, G., J. Hooker, and A. Cescatti, 2018: The mark of vegetation change on Earth's surface energy balance. *Nat. Commun.*, **9**, 679, https://doi.org/10.1038/s41467-017-02810-8.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, **9**, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

Grant, L., and Coauthors, 2021: Attribution of global lake systems change to anthropogenic forcing. *Nat. Geosci.*, **14**, 849–854, https://doi.org/10.1038/s41561-021-00833-x.

Harris, I., T. J. Osborn, P. Jones, and D. Lister, 2020: Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Sci. Data*, **7**, 109, https://doi.org/10.1038/s41597-020-0453-3.

Hartung, K., and Coauthors, 2021: Bookkeeping estimates of the net land-use change flux—A sensitivity study with the CMIP6 land-use dataset. *Earth Syst. Dyn.*, **12**, 763–782, https://doi.org/10.5194/esd-12-763-2021.

Hegerl, G. C., T. R. Karl, M. Allen, N. L. Bindoff, N. Gillett, D. Karoly, X. Zhang, and Z. Francis, 2006: Climate change detection and attribution: Beyond mean temperature signals. *J. Climate*, **19**, 5058–5077, https://doi.org/10.1175/JCLI3900.1.

——, and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 663–745.

Hirsch, A. L., A. J. Pitman, J. Kala, R. Lorenz, and M. G. Donat, 2015: Modulation of land-use change impacts on temperature extremes via land–atmosphere coupling over Australia. *Earth Interact.*, **19**, https://doi.org/10.1175/EI-D-15-0011.1.

Hurtt, G. C., and Coauthors, 2011: Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic Change*, **109**, 117, https://doi.org/10.1007/s10584-011-0153-2.

——, and Coauthors, 2020: Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. *Geosci. Model Dev.*, **13**, 5425–5464, https://doi.org/10.5194/gmd-13-5425-2020.

Iturbide, M., and Coauthors, 2020: An update of IPCC climate reference regions for subcontinental analysis of climate model data: Definition and aggregated datasets. *Earth Syst. Sci. Data*, **12**, 2959–2970, https://doi.org/10.5194/essd-12-2959-2020.

Jia, G., and Coauthors, 2019: Land-climate interactions. *Climate Change and Land*, P. R. Shukla et al., Eds., Cambridge University Press, 131–248.

Kumar, S., P. A. Dirmeyer, V. Merwade, T. DelSole, J. M. Adams, and D. Niyogi, 2013: Land use/cover change impacts in CMIP5 climate simulations: A new methodology and 21st century challenges. *J. Geophys. Res. Atmos.*, **118**, 6337–6353, https://doi.org/10.1002/jgrd.50463.

Lawrence, D. M., and Coauthors, 2016: The land use model intercomparison project (LUMIP) contribution to CMIP6: Rationale and experimental design. *Geosci. Model Dev.*, **9**, 2973–2998, https://doi.org/10.5194/gmd-9-2973-2016.

Ledoit, O., and M. Wolf, 2004: A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.*, **88**, 365–411, https://doi.org/10.1016/S0047-259X(03)00096-4.

Lee, X., and Coauthors, 2011: Observed increase in local cooling effect of deforestation at higher latitudes. *Nature*, **479**, 384–387, https://doi.org/10.1038/nature10588.

Lejeune, Q., S. I. Seneviratne, and E. L. Davin, 2017: Historical land-cover change impacts on climate: Comparative assessment of LUCID and CMIP5 multimodel experiments. *J. Climate*, **30**, 1439–1459, https://doi.org/10.1175/JCLI-D-16-0213.1.

——, E. L. Davin, L. Gudmundsson, J. Winckler, and S. I. Seneviratne, 2018: Historical deforestation locally increased the intensity of hot days in northern mid-latitudes. *Nat. Climate Change*, **8**, 386–390, https://doi.org/10.1038/s41558-018-0131-z.

——, ——, G. Duveiller, B. Crezee, R. Meier, A. Cescatti, and S. I. Seneviratne, 2020: Biases in the albedo sensitivity to deforestation in CMIP5 models and their impacts on the associated historical radiative forcing. *Earth Syst. Dyn.*, **11**, 1209–1232, https://doi.org/10.5194/esd-11-1209-2020.

Marvel, K., B. I. Cook, C. J. Bonfils, P. J. Durack, J. E. Smerdon, and A. P. Williams, 2019: Twentieth-century hydroclimate changes consistent with human influence. *Nature*, **569**, 59–65, https://doi.org/10.1038/s41586-019-1149-8.

Meier, R., E. L. Davin, S. C. Swenson, D. M. Lawrence, and J. Schwaab, 2019: Biomass heat storage dampens diurnal temperature variations in forests. *Environ. Res. Lett.*, **14**, 084026, https://doi.org/10.1088/1748-9326/ab4a42.

——, ——, G. B. Bonan, D. M. Lawrence, X. Hu, G. Duveiller, C. Prigent, and S. I. Seneviratne, 2022: Impacts of a revised surface roughness parameterization in the Community Land Model 5.1. *Geosci. Model Dev.*, **15**, 2365–2393, https://doi.org/10.5194/gmd-15-2365-2022.

Miao, H., X. Wang, Y. Liu, and G. Wu, 2021: A regime-based investigation into the errors of CMIP6 simulated cloud radiative effects using satellite observations. *Geophys. Res. Lett.*, **48**, e2021GL095399, https://doi.org/10.1029/2021GL095399.

Padrón, R. S., and Coauthors, 2020: Observed changes in dry-season water availability attributed to human-induced climate change. *Nat. Geosci.*, **13**, 477–481, https://doi.org/10.1038/s41561-020-0594-1.

Pitman, A. J., and Coauthors, 2009: Uncertainties in climate responses to past land cover change: First results from the LUCID intercomparison study. *Geophys. Res. Lett.*, **36**, L14814, https://doi.org/10.1029/2009GL039076.

Pongratz, J., C. Schwingshackl, S. Bultan, W. Obermeier, F. Havermann, and S. Guo, 2021: Land use effects on climate: Current state, recent progress, and emerging topics. *Curr. Climate Change Rep.*, **7**, 99–120, https://doi.org/10.1007/s40641-021-00178-y.

Prestele, R., A. Arneth, A. Bondeau, N. de Noblet-Ducoudré, T. A. Pugh, S. Sitch, E. Stehfest, and P. H. Verburg, 2017: Current challenges of implementing anthropogenic land-use and land-cover change in models contributing to climate change assessments. *Earth Syst. Dyn.*, **8**, 369–386, https://doi.org/10.5194/esd-8-369-2017.

Qian, C., and X. Zhang, 2015: Human influences on changes in the temperature seasonality in mid- to high-latitude land areas. *J. Climate*, **28**, 5908–5921, https://doi.org/10.1175/JCLI-D-14-00821.1.

Ribes, A., S. Planton, and L. Terray, 2013: Application of regularised optimal fingerprinting to attribution. Part I: Method, properties and idealised analysis. *Climate Dyn.*, **41**, 2817–2836, https://doi.org/10.1007/s00382-013-1735-7.

——, S. Qasmi, and N. P. Gillett, 2021: Making climate projections conditional on historical observations. *Sci. Adv.*, **7**, eabc0671, https://doi.org/10.1126/sciadv.abc0671.

Rohde, R. A., and Z. Hausfather, 2020: The Berkeley Earth land/ocean temperature record. *Earth Syst. Sci. Data*, **12**, 3469–3479, https://doi.org/10.5194/essd-12-3469-2020.

Santer, B. D., and Coauthors, 2007: Identification of human-induced changes in atmospheric moisture content. *Proc. Natl. Acad. Sci. USA*, **104**, 15 248–15 253, https://doi.org/10.1073/pnas.0702872104.

——, and Coauthors, 2011: Separating signal and noise in atmospheric temperature changes: The importance of timescale. *J. Geophys. Res.*, **116**, D22105, https://doi.org/10.1029/2011JD016263.

——, and Coauthors, 2018: Human influence on the seasonal cycle of tropospheric temperature. *Science*, **361**, eaas8806, https://doi.org/10.1126/science.aas8806.

Séférian, R., and Coauthors, 2019: Evaluation of CNRM Earth system model, CNRM-ESM2-1: Role of earth system processes in present-day and future climate. *J. Adv. Model. Earth Syst.*, **11**, 4182–4227, https://doi.org/10.1029/2019MS001791.

Sellar, A. A., and Coauthors, 2020: Implementation of U.K. Earth system models for CMIP6. *J. Adv. Model. Earth Syst.*, **12**, e2019MS001946, https://doi.org/10.1029/2019MS001946.

Swart, N. C., and Coauthors, 2019: The Canadian Earth system model version 5 (CanESM5.0.3). *Geosci. Model Dev.*, **12**, 4823–4873, https://doi.org/10.5194/gmd-12-4823-2019.

Thiery, W., and Coauthors, 2020: Warming of hot extremes alleviated by expanding irrigation. *Nat. Commun.*, **11**, 290, https://doi.org/10.1038/s41467-019-14075-4.

Wan, H., X. Zhang, F. Zwiers, and S. K. Min, 2015: Attributing northern high-latitude precipitation change over the period 1966–2005 to human influence. *Climate Dyn.*, **45**, 1713–1726, https://doi.org/10.1007/s00382-014-2423-y.

Wilks, D. S., 2019: *Statistical Methods in Atmospheric Sciences*. 4th ed. Elsevier, 818 pp., https://doi.org/10.1016/C2017-0-03921-6.

Winckler, J., C. H. Reick, and J. Pongratz, 2017: Robust identification of local biogeophysical effects of land-cover change in a global climate model. *J. Climate*, **30**, 1159–1176, https://doi.org/10.1175/JCLI-D-16-0067.1.

——, Q. Lejeune, C. H. Reick, and J. Pongratz, 2019a: Nonlocal effects dominate the global mean surface temperature response to the biogeophysical effects of deforestation. *Geophys. Res. Lett.*, **46**, 745–755, https://doi.org/10.1029/2018GL080211.

——, C. H. Reick, R. M. Bright, and J. Pongratz, 2019b: Importance of surface roughness for the local biogeophysical effects of deforestation. *J. Geophys. Res. Atmos.*, **124**, 8605–8618, https://doi.org/10.1029/2018JD030127.

——, and Coauthors, 2019c: Different response of surface temperature and air temperature to deforestation in climate models. *Earth Syst. Dyn.*, **10**, 473–484, https://doi.org/10.5194/esd-10-473-2019.

Zwiers, F. W., X. Zhang, and Y. Feng, 2011: Anthropogenic influence on long return period daily temperature extremes at regional scales. *J. Climate*, **24**, 881–892, https://doi.org/10.1175/2010JCLI3908.1.