Warming Trends and Long-Range Dependent Climate Variability Since Year 1900: A Bayesian Approach

Eirik Myrvoll-Nilsen, Hege-Beate Fredriksen, Sigrunn H. Sørbye and Martin Rypdal*

Department of Mathematics and Statistics, UiT The Arctic University of Norway, Tromsø, Norway

Temporal persistence in unforced climate variability makes detection of trends in surface temperature difficult. Part of the challenge is methodological since standard techniques assume a separation of time scales between trend and noise. In this work we present a novel Bayesian approach to trend detection under the assumption of long-range dependent natural variability, and we use estimates of historical forcing to test if the method correctly discriminates trends from low-frequency natural variability. As an application we analyze $2° × 2°$ gridded data from the GISS Surface Temperature Analysis. In the time period from 1900 to 2015 we find positive trends for 99% of the grid points. For 84% of the grid points we are confident that the trend is positive, meaning that the 95% credibility interval for the temperature trend contained only positive values. This number increased to 89% when we used estimates of historical forcing to specify the noise model. For the time period from 1900 to 1985 the corresponding ratios were 42 and 52%. Our findings demonstrate that positive trends since 1900 are now detectable locally over most of Earth's surface.

Keywords: trend detection, climate change, long-range dependence, fractional Gaussian noise, bayesian methods

1. INTRODUCTION

Since the year 1900, the global mean surface temperature (GMST) has increased by almost 1 degree K due to increasing concentrations of greenhouse gases in the atmosphere. While we are far from a full understanding of the complex dynamics of Earth's climate, the cause of the industrial-era warming is well understood, and the question of detection of global warming is today of little relevance. Recent detection studies for global temperature have instead focused on identifying the onset time of the anthropogenic warming, as well as the time when the warming became statistically detectable (Abram et al., 2016). For local and regional temperatures the situation is different. Not all of Earth's surface has warmed since 1900, and in some locations the warming is small compared to the natural variability. Sutton et al. (2015) point out that the question of local detectability is highly relevant since it provides insight into the strength of the warming signal relative to the natural fluctuations for which ecosystems are adapted.

From a statistical point of view, temperature trends provide a unique challenge since the climate naturally fluctuates on an extended range of time scales. A standard set-up is to assume that a temperature anomaly time series $\Delta T(t)$ can be approximated by a model on the form:

$$\Delta T(t) = m(t) + \varepsilon(t), \tag{1}$$

where $m(t) = a + bt$ is a linear trend, and $\varepsilon(t)$ is a stochastic process (Bloomfield, 1992). In the zero-dimensional global energy-balance model (EBM),

$$C\Delta T = -\lambda\Delta T dt + F(t)dt + \sigma_{OU}dB(t), \quad (2)$$

the unforced temperature fluctuations $\varepsilon(t)$ are defined by the equation $Cd\varepsilon = -\lambda\varepsilon dt + \sigma_{OU}dB(t)$, which describes an Ornstein-Uhlenbeck (OU) process (the continuous interpretation of a first-order auto-regressive (AR1) process) with a characteristic correlation scale $\tau = C/\lambda$:

$$\varepsilon(t) = \frac{\sigma_{OU}}{C}\int_{-\infty}^{t}e^{-(t-s)/\tau}dB(s).$$

In the equations above, C is the average heat capacity of Earth's surface, λ is the feedback parameter, and $F(t)$ is a forcing record. For this model, parameter estimates in temperature time series yield values of τ that are much shorter than those relevant for the warming trend. Hence, there is a separation of time scales that makes it easy to estimate the characteristics of the noise process without influence from the long-term trend. However, the zero-dimensional EBM in Equation (2) does not model the slow thermal response of the deep oceans, and the generalization to the so-called 2-box EBM (Geoffroy et al., 2013) gives a noise model on the form:

$$\varepsilon(t) = \int_{-\infty}^{t}G(t-s)dB(s),$$

where the response function $G(t) = c_1e^{-t/\tau_1} + c_2e^{-t/\tau_2}$ is a sum of two exponential functions. The generalization to N-box models gives response functions that are sums of N exponential functions. Models with multiple characteristic time scales are consistent with observations. Estimated power spectral densities (PSDs) of temperature reconstructions show approximate scale invariance, i.e., $S(f) \sim f^{-\beta}$, for frequencies corresponding to time scales from months to several hundred thousand years (Rypdal and Rypdal, 2016). Analyses of the relation between reconstructed forcing and reconstructed temperatures, as well as experiments in Earth System Models (ESMs), show approximate scale invariance in the climate response and in unforced climate variability on time scales from months to several hundred years (Rypdal and Rypdal, 2014; Rypdal et al., 2018). The implication for trend detection is that the noise processes that represent natural variability should be allowed to exhibit long-range dependence (LRD). A parsimonious model with LRD is obtained by using a power-law response function $G(t) = (t/\mu)^{\beta/2-1}$. With this choice, the noise model $\varepsilon(t)$ is a fractional Gaussian noise (fGn). The parameter β is identical to the exponent in the PSD and related to the so-called Hurst exponent via the relation $\beta = 2H - 1$.

We write the discrete-time version of Equation (1) on the following vector form:

$$\mathbf{y} = (y_1, \ldots, y_n)^\top = (m(t_1) + \varepsilon(t_1), \ldots, m(t_n) + \varepsilon(t_n))^\top, \quad (3)$$

where $\mathbf{y}$ is the temperature time series. Using the short-hand notation $\varepsilon_i = \varepsilon(t_i)$, the vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_n)^\top$ is a

zero-mean stationary Gaussian process with covariance function (Mandelbrot and Ness, 1968):

$$\text{Cov}(\varepsilon_i, \varepsilon_j) = \frac{\sigma_\varepsilon^2}{2}\left(|i-j+1|^{2H} + |i-j-1|^{2H} + |i-j|^{2H}\right). \quad (4)$$

Several previous studies have modeled climatic data using a linear trend model $m(t_i) = a + bt_i$, $i = 1, \ldots n$, and an LRD noise term (Cohn and Lins, 2005; Koutsoyiannis and Montanari, 2007; Franzke, 2012; Løvsletten and Rypdal, 2016). Some of these studies conclude that trends that have been identified as statistically significant based on AR1 noise models, are not found to be significant when LRD noise models are used. Ideally, the parameters in the noise model (the Hurst exponent H and the scale parameter σ_ε) and the parameters in the trend model (the intercept a and the slope b) should be estimated simultaneously, but most previous studies have used non-parametric measures of the second-order statistics to determine the Hurst exponent. A standard approach is to estimate a trend using a least-squares method, and to subsequently estimate a Hurst exponent and a scale parameter of the de-trended signal using a fluctuation function over a range of time scales. The fluctuation function could be a wavelet-based fluctuation function, a variogram, the de-trended fluctuation (DFA) function, or the PSD. The significance of a positive trend can then be tested using Monte-Carlo simulations, or by theoretical estimates based on the specified noise model. The disadvantage of the two-step approach is that it does not fully take into account the dependence between the estimates of the trend and the estimates of the noise parameters. Part of the reason why fluctuation-based estimators of H and σ_ε are popular is that likelihood-based methods are computationally costly for processes with LRD. Computing the likelihood function involves inversions of the dense covariance matrix $\boldsymbol{\Sigma}$ defined by the elements $\Sigma_{ij} = \text{Cov}(\varepsilon_i, \varepsilon_j)$ in Equation (4).

In this paper we take advantage of recent work by Sørbye et al. (2019) who incorporate fGn models within a Bayesian hierarchical formulation using the computational framework of latent Gaussian models. These models can be analyzed efficiently using the methodology of integrated nested Laplace approximation (INLA) developed in Rue et al. (2009). The INLA methodology provides accurate estimates of the posterior marginal distributions for all of the model parameters which can then be used to calculate summary statistics like posterior means, variances, credible intervals, and posterior probabilities. Of particular interest is the posterior marginal distribution $p(b \mid \mathbf{y})$, which is used to calculate the probability $\text{Prob}[b > 0 \mid \mathbf{y}]$ of a positive trend given the observed temperature anomalies $\mathbf{y}$. The Bayesian modeling approach is described further in section 2. In section 3 we discuss an alternative approach to trend detection where data of historical forcing is used to discriminate between forced response and natural variability. The results of the latter is used to test and validate the methods described in section 2. Results of analyses of gridded temperature data from the GISS Surface Temperature Analysis are presented in section 4, and discussed in section 5.

2. BAYESIAN INFERENCE

To analyse the regression model defined by Equation (3) for a large number of gridded time series, computational efficiency is crucial. The presented Bayesian approach makes use of the computational framework of latent Gaussian models. These models can be seen as a flexible class of three-stage Bayesian hierarchical models where the different stages specify the distribution of an observational vector $\mathbf{y}$, a Gaussian prior for a latent random field $\mathbf{x}$ and priors for random hyperparameters $\boldsymbol{\theta}$. The first stage of the model assumes that the observations are conditionally independent given the latent field and the hyperparameters. The resulting joint conditional distribution of the observations is then expressed as the product of the marginals:

$$p(\mathbf{y} \mid \mathbf{x}, \boldsymbol{\theta}) = p(y_1 \mid x_1, \boldsymbol{\theta}) \cdots p(y_n \mid x_n, \boldsymbol{\theta}).$$

In our case, the observations represent the temperature time series at some grid point and are assumed to have a Gaussian distribution and the marginals are just univariate Gaussian distributions. We consider each local time series independently and thus do not include spatial correlation.

The second stage of the latent Gaussian model formulation specifies that the conditional distribution of $\mathbf{x}$ given $\boldsymbol{\theta}$ is a Gaussian random field. Based on the regression formulation in Equation (3), the expectation of $\mathbf{y}$ is modeled in terms of a linear predictor, $\eta = E(\mathbf{y})$. The latent field $\mathbf{x}$ includes all the random variables of the predictor $\mathbf{x} = (a, b, \boldsymbol{\varepsilon})^{\top}$. By assigning Gaussian priors to all of these variables, $\mathbf{x}$ will also be Gaussian and this is what separates a latent Gaussian model from general three-stage Bayesian hierarchical models. The conditional distribution of $\mathbf{x}$ given hyperparameters is then defined by:

$$\mathbf{x} \mid \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{Q}^{-1}) \tag{5}$$

where $\boldsymbol{\mu}$ denotes the mean vector and $\mathbf{Q}$ is the precision (inverse covariance) matrix of all the random variables in $\mathbf{x}$. The matrix $\mathbf{Q}$ reflects conditional independence properties of the elements in $\mathbf{x}$, giving zeros for all combinations of elements x_i and x_j that are independent conditioned on the other elements of $\mathbf{x}$. Usually, $\mathbf{x}$ is assumed to be a Gaussian Markov random field (GMRF) implying that the precision matrix $\mathbf{Q}$ will be sparse.

The final stage of the model formulation specifies priors for the hyperparameters which here include $\boldsymbol{\theta} = (H, \sigma_\varepsilon)$. Assuming independent priors, the probability density function is $p(\boldsymbol{\theta}) = p(H)p(\sigma_\varepsilon)$ where both parameters are assigned penalized complexity priors (Simpson et al., 2017). This is a recently developed class of priors which introduces a framework to compute priors for hyperparameters based on specific principles. For scaling parameters such as σ_ε, the PC prior can be computed to equal the exponential distribution. The PC prior for the Hurst exponent is computed numerically as explained in Sørbye and Rue (2018). Using Bayes theorem, the posterior joint distribution of the latent field and the hyperparameters is expressed as:

$$p(\boldsymbol{\theta}, \mathbf{x} \mid \mathbf{y}) \propto \prod_{i=1}^{n} p(y_i \mid x_i, \boldsymbol{\theta}) p(\mathbf{x} \mid \boldsymbol{\theta}) p(H) p(\sigma_\epsilon). \tag{6}$$

The main aim of the current analysis is to find the posterior marginal distributions $p(b \mid \mathbf{y})$, $p(H \mid \mathbf{y})$ and $p(\sigma_\epsilon \mid \mathbf{y})$. These distributions are then used to find summary statistics for the parameters. Especially, significance of warming trends are assessed by the probability of $b > 0$ according to the density $p(b \mid \mathbf{y})$. More generally, posterior marginal distributions for all components of $\mathbf{x}$ and the hyperparameters in $\boldsymbol{\theta}$ might be of interest. Theoretically, such marginals are expressed by integrating out all other variables in Equation (6), but this is not a computationally feasible approach. Posterior samples from the posterior marginals can be obtained using Markov chain Monte Carlo approaches (Gamerman and Lopes, 2006), but this is computationally slow as such approaches are simulation-based. The INLA methodology (Rue et al., 2009) represents an accurate and computationally superior alternative as it estimates the posterior marginals without any simulations, combining numerical approximations with numerical interpolation and integration (see Rue et al., 2017) for a recent review. In order for INLA to be computationally fast, the latent field $\mathbf{x}$ needs to be a GMRF having a sparse precision matrix $\mathbf{Q}$ in Equation (5). This is not the case when the noise term $\boldsymbol{\epsilon}$ is fGn having an LRD structure, but the precision matrix of an AR1 process is tridiagonal. In Sørbye et al. (2019), the fGn process is approximated using a weighted sum of AR1 processes where the weights and the first-lag autocorrelation coefficients of the approximation are optimized such that the covariance function of the approximation matches the exact covariance function of fGn defined in Equation (4). The latent field x is extended to include the AR(1) components that make up the approximation. This implies that full Bayesian inference is obtained for these as well. For the time scales of interest, the approximation is very accurate using a sum of only four AR1 processes. This speeds up the model fitting of Equation (3) considerably, see section 4 for results.

3. USING HISTORICAL FORCING TO SPECIFY NOISE MODELS

The alternative approach to trend detection that we present in this paper makes use of the historical global data of radiative forcing $F(t)$ (an updated version of the forcing in Hansen et al., 2011). This is done to validate the results of the approach in section 2 which does not account for information about radiative forcing. For the EBM in Equation (2), the forced temperature response is:

$$\Delta T(t) = \frac{1}{C} \int_{-\infty}^{t} e^{-(t-s)/\tau} F(s) ds + \frac{\sigma_{\mathrm{OU}}}{C} \int_{-\infty}^{t} e^{-(t-s)/\tau} dB(s).$$

Expressed as in (1) we get that $\varepsilon(t)$ is an OU model with characteristic correlation length τ and,

$$m(t) = \frac{1}{C} \int_{-\infty}^{t} e^{-(t-s)/\tau} F(s) ds$$

is a convolution of the exponential kernel $e^{-t/\tau}$ with the historical forcing.

Rypdal and Rypdal (2014) proposed an LRD modification of this model in discrete time where $\boldsymbol{\varepsilon}$ in Equation (3) is an fGn process with Hurst exponent H, variance σ_ε^2, and mean function:

$$m(t_i) = \sigma_f \sum_{s=-\infty}^{t_i} (t_i - s)^{H-3/2} \big(F_0 + F(s) \big), \quad i = 1, \ldots, n. \quad (7)$$

The parameter F_0 is introduced to make sure that the forcing records $F(t_i)$ have the correct mean as this a relative measure, while σ_f is an additional scaling parameter. Myrvoll-Nilsen et al. (submitted) extend the methodology described in section 2 to analyse models where the fGn process has mean defined by Equation (7). This approach is more computationally demanding as it introduces the additional hyperparameters σ_f and F_0, such that $\boldsymbol{\theta} = (H, \sigma_\varepsilon, \sigma_f, F_0)$. Here, F_0 is assigned a vague Gaussian prior, while the other hyperparameters are assigned penalized complexity priors.

Figure 1 shows surface temperature anomalies from two grid points, one located around the city of Moscow (55N, 37E), and one location in the tropical Pacific ocean (33S, 120W). The dotted blue lines show the estimated linear trend $m(t_i) = a + bt_i$ using the methods described in section 2, and the solid black curves are the estimated forced responses in Equation (7). Using historical forcing implies that we are imposing a global warming signal, and hence estimates of increasing functions $m(t_i)$ can not be considered as detection of warming trends. However, our purpose is not to obtain estimates of $m(t_i)$, but rather to estimate the parameters in the noise term $\boldsymbol{\varepsilon}$ when the forced response is modeled more realistically than a linear function. Hence, our second method for trend detection is to use a linear model $m(t_i) = a + bt_i$ together with a noise model $\boldsymbol{\varepsilon}$, where the parameters σ_ε and H are fixed and equal to the posterior marginal mean values obtained from the model where $m(t_i)$ is given by Equation (7).

4. RESULTS

In this section we present results for $2° \times 2°$ gridded data from the GISS Surface Temperature Analysis. Annual data is used and parameter estimates are given for those time series that have no more than 5% missing values. We have used the two different methods described in sections 2 and 3, respectively. **Figure 2** shows maps of estimates for the trend parameter b, and the noise parameters σ_ε and H for the time period 1900–2015. The presented estimates are the posterior means obtained from the estimated posterior marginal distributions. The method described in section 2 is used to obtain the estimates in **Figures 2A,C,E**, and the method described in section 3 is used to obtain the estimates in **Figures 2B,D,F**. It is well-known that the Hurst exponents are higher for sea-surface temperatures than for land temperatures (Fraedrich and Blender, 2003; Monetti et al., 2003; Fredriksen and Rypdal, 2016), and this is confirmed in this study. We also observe stronger warming trends in the Arctic compared with the rest of Earth's surface, consistent with polar amplification. Of the 11,997 grid points that are analyzed, 11,883 and 11,906 had positive estimates for the trend parameter b for the two methods, respectively.

The two methods presented do give quite similar results for the parameters b, σ_ε, and H, indicating that the method described in section 2 produces reasonable estimates of the noise parameters. However, **Figure 3** shows that the root mean square error, defined by:
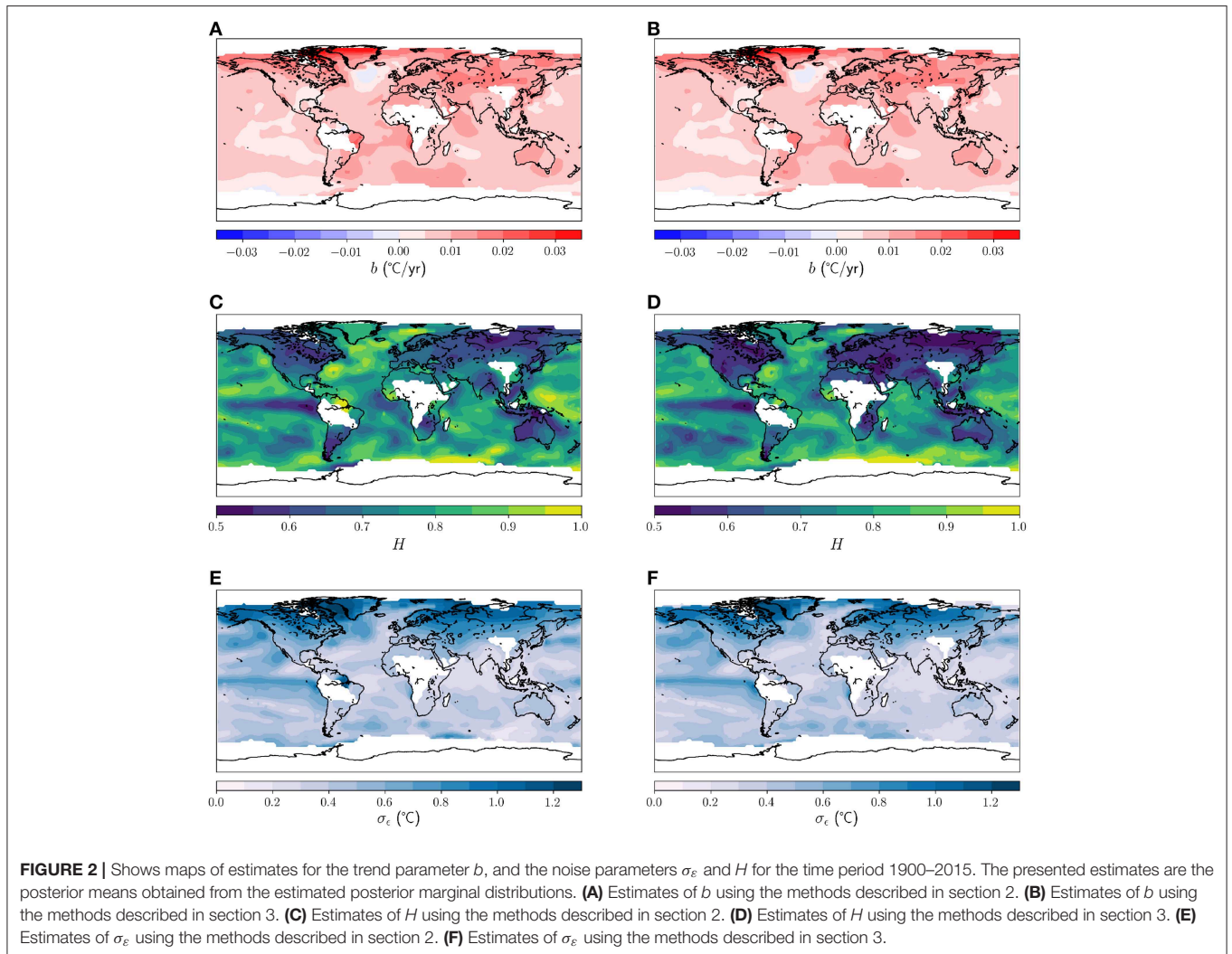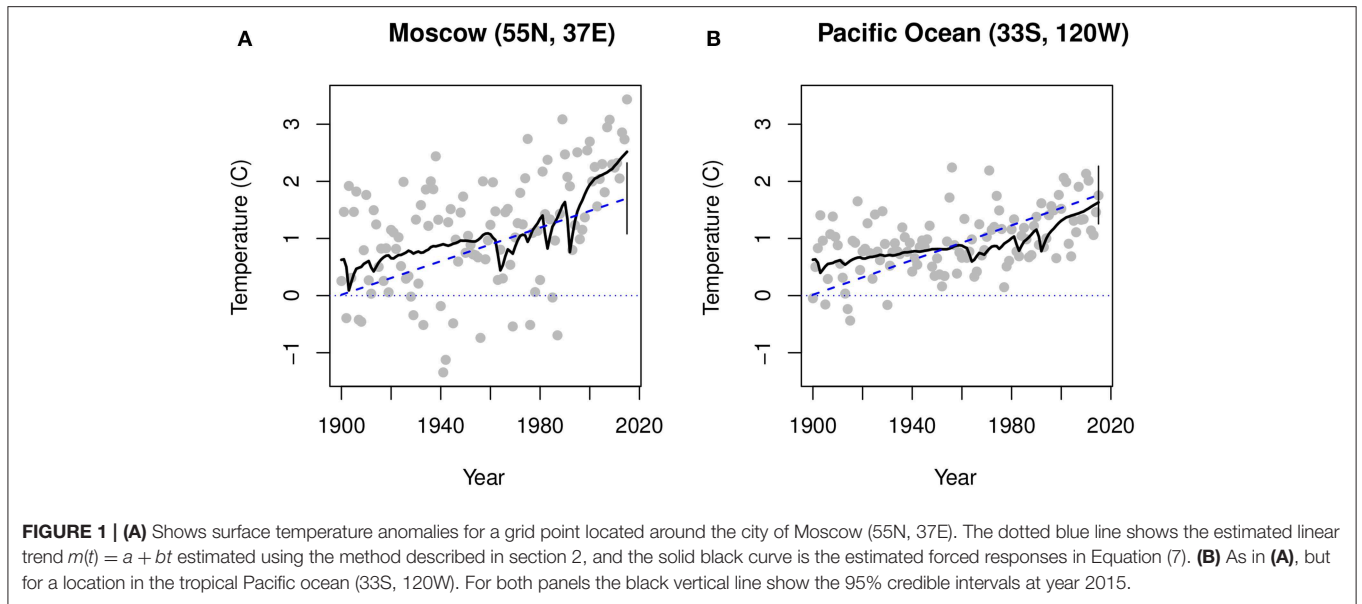
$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{m}(t_i))^2}{n}}, \quad (8)$$

where $\hat{m}(t_i)$ is the posterior marginal mean of $m(t_i)$, is generally higher when a linear trend is used. Hence, the noise processes need to account for more variability in the models described in section 2 than in the models described in section 3. This effect can also be seen in **Figure 4**, which summarizes the main findings of this work, but the effect is very subtle. **Figures 4A,C,E** show the posterior probability of a positive linear trend given the observed temperature time series for each grid point using the method described in section 2, and **Figures 4B,D,F** show the same numbers obtained using the method described in section 3. **Figures 4A,B** show results for the years 1900–1950, **Figures 4C,D** show results for the years 1900–1985, and **Figures 4E,F** show results for the time period 1900–2015. Using data up to the year 1950 we find $\text{Prob}[b > 0 \mid \mathbf{y}] > 0.95$ for 3571 and 4,606 out of the 11,997 analyzed grid points, for the two methods, respectively. These numbers increase to 5,140 and 6,223 is the time period extends to year 1985. And to 10,121 and 10,683 when the analysis includes all years from 1900 to 2015. **Figure 4** shows that there are large areas in the oceans where the sea-surface temperature warming signal has become detectable over the last 30 years.

Fitting the model in Equation (3) where the linear trend and the parameters of the fGn model are estimated simultaneously, required on average 2.7 s per time series. This gives a total elapsed computation time of $\sim$9 h for all 11,997 grid points. Using the approach described in section 3, we first fit an fGn process where the mean is specified by Equation (7). The average run time to fit this model to a single time series was almost 8 s, giving a total run time of $\sim$25.9 h. The second step of the method in section 3 fits the linear trend combined with the fGn noise term using fixed parameters. Fitting of this model required on average 1.5 s for individual time series, giving a total elapsed computation time of $\sim$5 h for all grid points. The main reason for the increased computation time of the approach described in section 3 compared with the method in section 2 is the increased number of hyperparameters. Also, fitting of the model including radiative forcing required extensions to existing software, see Myrvoll-Nilsen et al. (submitted) for further details.

5. DISCUSSION AND CONCLUSIONS

The main contribution of this paper is to present a computationally efficient Bayesian method for trend-detection under the assumption of LRD noise, and to apply the method to detection of global warming in gridded temperature data. By considering two different methods, where the second
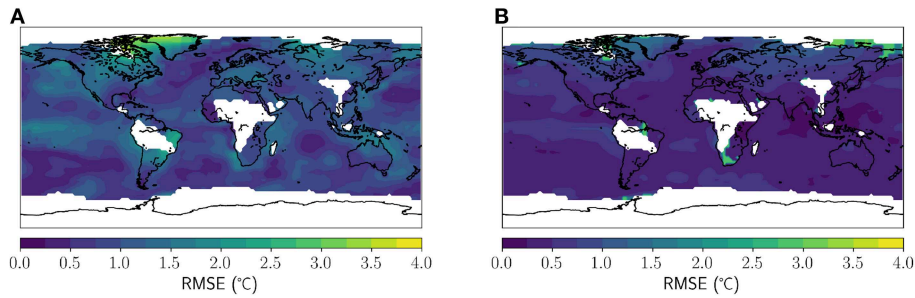
FIGURE 1 | (A) Shows surface temperature anomalies for a grid point located around the city of Moscow (55N, 37E). The dotted blue line shows the estimated linear trend $m(t) = a + bt$ estimated using the method described in section 2, and the solid black curve is the estimated forced responses in Equation (7). **(B)** As in **(A)**, but for a location in the tropical Pacific ocean (33S, 120W). For both panels the black vertical line show the 95% credible intervals at year 2015.

FIGURE 2 | Shows maps of estimates for the trend parameter b, and the noise parameters σ_ε and H for the time period 1900–2015. The presented estimates are the posterior means obtained from the estimated posterior marginal distributions. **(A)** Estimates of b using the methods described in section 2. **(B)** Estimates of b using the methods described in section 3. **(C)** Estimates of H using the methods described in section 2. **(D)** Estimates of H using the methods described in section 3. **(E)** Estimates of σ_ε using the methods described in section 2. **(F)** Estimates of σ_ε using the methods described in section 3.

FIGURE 3 | The root mean square error (RMSE) as defined in Equation (8), measuring a standardized difference between the observed temperature signal and the trend model. **(A)** The RMSE for the model defined in section 2. **(B)** The RSME for the model defined in section 3.
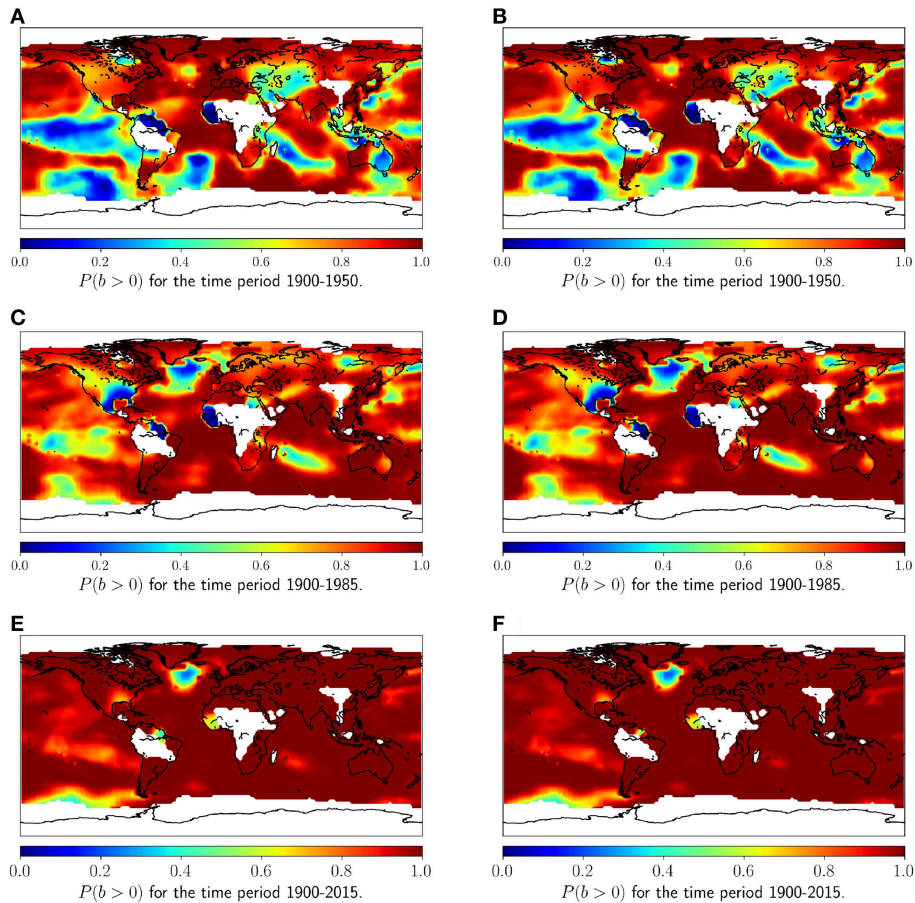

FIGURE 4 | The posterior probability of a positive linear trend given the observed temperature time series for each grid point. **(A)** For the time period 1900–1950 using the method described in section 2. **(B)** For the time period 1900–1950 using the method described in section 3. **(C)** For the time period 1900–1985 using the method described in section 2. **(D)** For the time period 1900–1985 using the method described in section 3. **(E)** For the time period 1900–2015 using the method described in section 2. **(F)** For the time period 1900–2015 using the method described in section 3.

uses historical data for global radiative forcing, we validate that the inaccuracy of a linear trend model does not affect the specification of the trend model to such a degree that it significantly affects the trend-detection results. Hence, we avoid a situation where the presence of non-linear temperature trends

produce biased estimates of σ_ε and H, which again affect the trend detection.

The results presented in this paper imply that even by the most conservative estimates, more than 84% of the analyzed grid points show significant warming at the 0.05-level. This ratio is

higher than the 80% ratio that Løvsletten and Rypdal (2016) obtained after making a model selection between AR1 and fGn for each grid point, and much higher than what they obtain for a pure fGn model. Our results also show a striking difference for the time period 1900–2015 compared to 1900–1985, and clearly demonstrate that even locally, the warming signal has emerged from the noise over the last 30 years.

The limitations of the approach presented here is that noise models are restricted to the class of fGns. Whereas LRD models are highly accurate on decadal time scales in GMST, local temperatures exhibit modes of internal variability that are not well-approximated by scale invariance. The most prominent example being the El Niño Souther Oscillation (ENSO). We therefore believe that future work on trend detection in surface temperature data should be based on more flexible models, with more free parameters. The introduction of more flexible models must be weighed against the risk of statistical overfitting for the relatively short instrumental temperature record. An alternative approach is to characterize internal variability as the fluctuations around ensemble means in historical runs in ESMs. Using randomization of phases one can then run Monte-Carlo simulations without specifying parametric models. However, since spatial patterns of internal variability vary between models it is difficult to ensure that the unforced fluctuations in a particular grid point is a good representation of the unforced fluctuations of this grid point in the real climate system.

DATA AVAILABILITY

Software for the INLA methodology is available at `R-INLA.org` for the programming environment `R`. INLA allows for fast Bayesian inference for latent Gaussian models such as the one described in section 2. However, some modifications are required in order to ensure efficiency. First, the latent field is of non-standard form since the mean vector and covariance matrix of the latent Gaussian field share the same parameter H, and thus requires separate specification. Second, due to the LRD assumption, the precision matrix, defined as the inverse covariance matrix, is dense and therefore unsuited for the computationally efficient algorithms that INLA applies. In order to retain sparsity we have to introduce an approximation such as that introduced by Sørbye et al. (2019). The model with these modifications are available in the `R`-package `INLA.climate` which can be downloaded from the GitHub repository `eirikmn/INLA.climate`. The temperature data analyzed is downloaded from https://data.giss.nasa.gov/gistemp/, and the forcing data from http://www.columbia.edu/~mhs119/Forcings/.

AUTHOR CONTRIBUTIONS

EM-N, SS, H-BF, and MR designed the study. EM-N carried out the analyses. H-BF made the figures. MR, SS, and EM-N wrote the paper with input from all authors.

REFERENCES

Abram, N. J., McGregor, H. V., Tierney, J. E., Evans, M. N., McKay, N. P., Kaufman, D. S., et al. (2016). Early onset of industrial-era warming across the oceans and continents. *Nature* 536, 411–413. doi: 10.1038/nature19082

Bloomfield, P. (1992). Trends in global temperature. *Clim. Change* 21, 1–16.

Cohn, T. A., and Lins, H. F. (2005). Nature's style: naturally trendy. *Geophys. Res. Lett.* 32:L23402. doi: 10.1029/2005GL024476

Fraedrich, K., and Blender, R. (2003). Scaling of atmosphere and ocean temperature correlations in observations and climate models. *Phys. Rev. Lett.* 90:108501. doi: 10.1103/PhysRevLett.90.108501

Franzke, C. (2012). On the statistical significance of surface air temperature trends in the Eurasian Arctic region. *Geophys. Res. Lett.* 39:L23705. doi: 10.1029/2012GL054244

Fredriksen, H.-B., and Rypdal, K. (2016). Spectral characteristics of instrumental and climate model surface temperatures. *J. Climate* 29, 1253–1268. doi: 10.1175/JCLI-D-15-0457.1

Gamerman, D., and Lopes, H. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd Edn. New York, NY: Chapman & Hall/CRC.

Geoffroy, O., Saint-Martin, D., Olivié, D. J. L., Voldoire, A., Bellon, G., and Tytéca, S. (2013). Transient climate response in a two-Layer energy-balance model. Part I: analytical solution and parameter calibration using CMIP5 AOGCM experiments. *J. Climate* 26, 1841–1857. doi: 10.1175/JCLI-D-12-00195.1

Hansen, J., Sato, M., Kharecha, P., and von Schuckmann, K. (2011). Earth's energy imbalance and implications. *Atmos. Chem. Phys.* 11, 13421–13449. doi: 10.5194/acp-11-13421-2011

Koutsoyiannis, D., and Montanari, A. (2007). Statistical analysis of hydroclimatic time series: uncertainty and insights. *Water Resour. Res.* 43:W05429. doi: 10.1029/2006WR005592

Løvsletten, O., and Rypdal, M. (2016). Statistics of regional surface temperatures after 1900: long-range versus short-range dependence and significance of warming trends. *J. Climate* 29, 4057–4068. doi: 10.1175/JCLI-D-15-0437.1

Mandelbrot, B., and Ness, J. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* 18, 1088–1107.

Monetti, R. A., Havlin, S., and Bunde, A. (2003). Long-term persistence in the sea surface temperature fluctuations. *Phys. A Stat. Mech. Appl.* 320, 581–589. doi: 10.1016/S0378-4371(02)01662-X

Rue, H., Martino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *J. R. Stat. Soc. Ser. B* 71, 319–392. doi: 10.1111/j.1467-9868.2008.00700.x

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., and Lindgren, F. K. (2017). Bayesian computing with INLA: a review. *Ann. Rev. Stat. Appl.* 4, 395–421. doi: 10.1146/annurev-statistics-060116-054045

Rypdal, M., Fredriksen, H.-B., Myrvoll-Nilsen, E., Rypdal, K., and Sørbye, S. H. (2018). Emergent scale invariance and climate sensitivity. *Climate* 6:93. doi: 10.3390/cli6040093

Rypdal, M., and Rypdal, K. (2014). Long-memory effects in linear response models of earth's temperature and implications for future global warming. *J. Climate* 27, 5240–5258. doi: 10.1175/JCLI-D-13-00296.1

Rypdal, M., and Rypdal, K. (2016). Late quaternary temperature variability described as abrupt transitions on a 1/f noise background. *Earth Syst. Dynam.* 7, 281–293. doi: 10.5194/esd-7-281-2016

Simpson, D., Rue, H., Riebler, A., Martins, T. G., and Sørbye, S. H. (2017). Penalising model component complexity: a principled, practical

approach to constructing priors. *Stat. Sci.* 232, 1–28. doi: 10.1214/16-STS576

Sørbye, S. H., Myrvoll-Nilsen, E., and Rue, H. (2019). An approximate fractional Gaussian noise model with $\mathcal{O}(n)$ computational cost. *Stat. Comput.* 29, 821–833. doi: 10.1007/s11222-018-9843-1

Sørbye, S. H., and Rue, H. (2018). Fractional Gaussian noise: prior specification and model comparison. *Environmetrics* 29:e2457. doi: 10.1002/env.2457

Sutton, R., Suckling, E., and Hawkins, E. (2015). What does global mean temperature tell us about local climate? *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 373:20140426. doi: 10.1098/rsta.2014.0426