LETTER • OPEN ACCESS

Are there persistent physical atmospheric responses to galactic cosmic rays?

To cite this article: Rasmus E Benestad 2013 *Environ. Res. Lett.* **8** 035049

View the article online for updates and enhancements.

Are there persistent physical atmospheric responses to galactic cosmic rays?

Rasmus E Benestad

The Norwegian Meteorological Institute, PO Box 43 Blindern, 0313 Oslo, Norway

E-mail: rasmus.benestad@met.no

Abstract

Variations in the annual mean of the galactic cosmic ray flux (GCR) are compared with annual variations in the most common meteorological variables: temperature, mean sea-level barometric pressure, and precipitation statistics. A multiple regression analysis was used to explore the potential for a GCR response on timescales longer than a year and to identify 'fingerprint' patterns in time and space associated with GCR as well as greenhouse gas (GHG) concentrations and the El Niño–Southern Oscillation (ENSO). The response pattern associated with GCR consisted of a negative temperature anomaly that was limited to parts of eastern Europe, and a weak anomaly in the sea-level pressure (SLP), but coincided with higher pressure over the Norwegian Sea. It had a similarity to the North Atlantic Oscillation (NAO) in the northern hemisphere and a wave train in the southern hemisphere. A set of Monte Carlo simulations nevertheless indicated that the weak amplitude of the global mean temperature response associated with GCR could easily be due to chance (p-value $= 0.6$), and there has been no trend in the GCR. Hence, there is little empirical evidence that links GCR to the recent global warming.

Keywords: cosmic rays, climate, regression analysis, temperature, pressure, precipitation

S Online supplementary data available from stacks.iop.org/ERL/8/035049/mmedia

1. Motivation

Speculations about a solar influence on Earth's climate have been around for centuries (e.g. table 8.1 in [1] provides some examples of scholars who discussed various hypotheses since the discovery of sunspots in 1610), and many of the hypotheses suggested in the past remain unresolved today. The oldest ones proposed that the Earth's climate was influenced by changes in the sunspots, solar output, solar magnetism, the solar wind, faculae, and eruptions. Cosmic rays were only discovered in 1912 [2], and thus constitute a more recent ingredient in the mix of explanations. Several scholars have since suggested that high-energy particles in the shape of galactic cosmic ray flux (GCR) exert an influence on Earth's climate by affecting cloud processes. Indeed, cosmic rays can be studied through cloud chambers, where they leave trails of condensation in containers with over-saturated vapour.

A recent review [3] provides a comprehensive account of the numerous studies on the link between GCR and climate, spanning the early attempts by Ney in 1959 [4] and Dickinson in 1975 [5], to recent work. Only since the turn of the millennium, there have been several papers suggesting a link between GCR and clouds [6–16]. The conclusions from some of these publications were that the GCR do affect the clouds [9, 11], whereas others suggest that GCR play a negligible role [3, 12, 10, 17]. The relationship between high-energy particles and aerosol nucleation involved in forming cloud condensation nuclei (CCN) were also examined in controlled experiments at CERN [18], however, it is not certain from those results that GCR is important for

CCN in real atmospheric conditions. A Danish lab too has run a 'low-budget' experiment 'SKY' which suggested that high-energy particles may affect ultra-small aerosols which theoretically could grow into CCNs [19], albeit on a very different level of complexity and accuracy.

Some scholars have implied a false dichotomy between GCR and greenhouse gases (GHGs), arguing that a global warming caused by GCR would be at the expense of an effect from rising concentrations of GHGs [20, 6]. Such propositions have resulted in a public controversy about the role of the sun, GCR, and GHGs [21–23]. In June 2011, a multidisciplinary European network was established between scientists from 18 countries in order to consolidate the scientific understanding of a potential influence that solar and high-energy particle may exert on Earth's climate. Another motivation was to provide the society with the best knowledge on these questions. This network is supported through the European intergovernmental COST framework, TOSCA—'towards a more complete assessment of the impact of solar variability on the Earths climate'. Some of TOSCA's objectives involve reviewing science relevant to the question about solar-terrestrial connections and pursuing new scientific questions which may shed further light on the matter.

Here the question about the link between GCR and climate is revisited, rephrasing it in terms of robust and visible effects from GCR on common meteorological elements. Furthermore, the focus here is on long-term variations and long timescales (years), rather than the timescale of weather (hours and days). The purpose is not to determine whether the GCR is important for physical processes in Earth's atmosphere, but whether there is a clear, direct, and lasting response to GCR that can be found in general meteorological parameters affecting people. The lack of a clear response does not exclude subtle and indirect effects.

The role of GCR is contrasted with other known factors in order to assess its significance, such as GHGs and ENSO. In other words, if the GCR really has profound effects on Earth's climate, then the most common meteorological elements should also be affected, such as temperature near the surface T_{2m}, the barometric pressure at the mean sea level p_{sl}, or the precipitation statistics.

2. Methods and data

The question whether there is a connection between GCR and T_{2m}, p_{sl}, or precipitation was examined in terms of historical data. The GCR data was from the climax neutron monitor, and is the longest record available, spanning from 1951 to 2006. The temperature was taken from two different data sets: the gridded GISTEMP [24] (1200 km smooth) and the 20th Century (20C) reanalysis [25]. Reanalyses are usually not appropriate for long-term climate studies due to the introduction of new observing systems over time, however, the 20C reanalysis only relied on surface observations (for which the observational network has been more stable than space-based remote sensing) and the analysis emphasized the year-to-year variations rather than the slow trends. Furthermore, the GISTEMP and the 20C reanalysis

represented two distinct accounts of the past temperatures. The p_{sl} data were also drawn from two different analyses: HadSLP2 [26] and 20C reanalysis. The precipitation statistics was based on GDCN [27–30]. The GDCN was used rather than the GPCP data set [31], even though the GPCP data set provides a better coverage. The analysis here was based on the wet-day mean μ and the wet-day frequency f_w, which can only be derived from 24-h amounts.

A multiple general linear regression was used to detect links between GCR and meteorological parameters, and it was assumed that a potential response to GCR could be approximated as linear. Similar assumptions are made in perturbation analyses [32] where it is assumed that the change is small compared to the mean level. The variations in T_{2m} and p_{sl} tend to be small in terms of their mean value (absolute value and the mean p_{sl}). Furthermore, the variations in the annual mean precipitation statistics can be treated as mere perturbations to the mean level as the standard deviation was less than 20% of the mean for most locations (figure S1 available at stacks.iop.org/ERL/8/035049/mmedia). The standard deviation of the daily mean GCR was 7% compared to the mean, and hence could be regarded as perturbations to the background level. Hence, if the response to GCR can be approximated as being linear, the analysis may provide a description of a first-order effect to forcings. Similar use of regression has been used in earlier work [33–37].

The analysis explored the link of these meteorological elements to GCR, GHGs, and ENSO. The latter two were included to verify the analytic set-up, as their effects have already been established. The analysis explored the link to GCR in both reanalysis data and gridded observations, and the consistency between these provided an indication of the robustness of the 'fingerprints'. The two data sets gave roughly similar values for the global mean temperature after 1950, but exhibited a systematic difference before ∼1950.

The analysis was applied to annually aggregated values, in order to emphasis moderately long timescales and lasting effects of GCR. The mathematical equation describing the regression was:

$$v_j(t) = \alpha_{0,j} + \alpha_{1,j}X_1(t) + \alpha_{2,j}X_2(t) + \alpha_{3,j}X_3(t) + \eta, \quad (1)$$

where t represented time and v was either T_{2m}, p_{sl}, or the wet-day mean 24-h precipitation (μ). The co-variates X_1–X_3 contained GCR data, forcing from well-mixed GHGs, and NINO3.4-index (representing ENSO), all of which were standardized (zero mean and unit standard deviation; figure S2 available at stacks.iop.org/ERL/8/035049/mmedia). The data for GHGs was the same as used in [35].

The motivation for including several co-variates was to account for some influences which are known to influence T_{2m} and p_{sl} in order to get an optimal fit. It is important, however, to keep in mind that such regression analyses also can give spurious results when there are independent co-variates that are co-linear [35]. However, here the connection between T_{2m} and e.g. ENSO was known, and hence served as a yard stick and an additional quality check.

The regression analysis used generalized linear models (GLM) and was applied to the 50-year interval 1951–2000,

which covered the climax GCR record. This period also coincided with a large number of rain gauge records with complete series, as many of the US stations in the GDCN data set had missing data after 2000.

In order to reduce the risk of fortuitous fit to accidental correlation, the regression was applied to spatio-temporal patterns and used to compute spatial response patterns. The use of regression to identify spatial response structure has also been adopted in other scientific studies [36, 37]. Here the combination of space and time were represented in the form of a set of empirical orthogonal functions (EOFs) [38], estimated through an singular value decomposition (SVD): $Y = UDV^T$ [39], where Y was the original data matrix with time and space dimensions, U contained the EOFs, D the eigenvalues, and V the principal components (PCs). The dependent variable $v_j(t)$ in equation (1) refers to single columns of V.

EOFs were estimated for the respective data sets of gridded annual values for T_{2m} and p_{sl}. For the precipitation, the EOFs were replaced by a principal component analysis (PCA) for annual wet-day mean precipitation (μ) from rain gauges. EOFs and PCA only differed by the fact that the EOFs were area-weighted according to the relative area of the grid boxes. The PCs for both were used in the multiple regression as v, and the analysis was repeated for each component. The regression coefficients $\alpha_{i,j}$ was estimated for each PC (distinguished by the index j) and each co-variate i, and a spatial pattern was then estimated by adding the contribution from each mode j: $\beta_i = \sum_j \alpha_{i,j} d_j u_j$, where d_j are the diagonal elements of D and u_j the columns of U. The global mean response was then estimated from the spatial patterns derived from the EOFs and the regression coefficients by taking the grid box area a_k into account:

$$\langle f(X_j(t)) \rangle = \frac{\sum_k (\beta_{i,k} a_k)}{\sum_k a_k} \times X_j(t). \tag{2}$$

For the wet-day frequency (f_w), a GLM with a logit link function ($\text{logit}(p) = \ln |p/(1-p)|$) was used, assuming a binomial distribution for the error. This choice was also appropriate for modelling a variable $y \in [0, 1]$. Due to the more complicated link function for f_w and its implication for the superposition of modes, the regression analysis was applied to each rain gauge record respectively, rather than to PCA.

The results were evaluated by extending the predictions to years outside the 1951–2000 interval, and the global mean temperatures derived from the spatial patterns were then compared with the observations. The GISTEMP and 20C reanalysis global mean temperatures diverged in the period before 1950, suggesting higher degree of uncertainty in their estimate in the early part of the period, due to more sparse data coverage.

In order to make prediction beyond the 1951–2006 interval, the GCR was extended based on a regression analysis between annual mean GCR from Climax and from GCR-reconstructions [40] (to ensure corresponding levels), for which the R^2 was 0.68 (a standard statistical metric corresponding to the correlation squared and 0.68 means that

the reconstruction explained 68% of the variance seen in GCR when the records overlap). Although the reconstruction by all means was imperfect, the purpose of this extension of the GCR was to provide a means to evaluate the analysis for ENSO and GHG.

The evaluation of the GCR results was limited to the 1951–2000 period, and two sets of Monte Carlo simulations were carried out where the standardized annual mean GCR was replaced by random numbers with similar mean and standard deviation. Here 3000 simulations were carried out for each set. In the first set of simulations, all the co-variates were replaced by random numbers mimicking the detrended annual mean values. The second set was run to test whether the results for GCR was influenced by the two other co-variates, where both GHG and ENSO were the kept same as in the original analysis while only the GCR was replaced by random numbers. The annual mean temperature response was then estimated based on the spatial patterns derived from the multiple regression against the EOFs (equations (1) and (2)), and these results were then used as a null-distribution for the global mean temperature response to GCR, GHG, and ENSO. The Monte Carlo simulations were not applied to the SLP, as the global mean SLP is expected to be invariant with respect to the co-variates and a function of the atmospheric mass. Likewise, a global mean value for the wet-day mean precipitation amount are difficult to interpret due to incomplete geographical sampling and because the wet-day frequency too plays a role for the hydrological cycle. Here, the autocorrelation was not taken into account when generating random numbers.

3. Results

Figure 1 shows spatial patterns revealing geographical dependencies of the direct linear response to the forcings. These patterns were composites of the regression results for the different annual mean temperature EOFs, and the co-variance can account for 45% of the total variance. The statistical significance of these patterns will be examined later in terms of the global mean response, and the patterns for GHG an ENSO were used to check for well-known teleconnection patterns and finger prints. For ENSO, the analysis identified the anomaly along the equator in the eastern part of the Pacific [41], with a magnitude of ∼0.6 °C for the annual mean values. The spatial response pattern associated with the GHGs revealed a high sensitivity over the Arctic (∼1.5 °C) for the 20C reanalysis, consistent with the 'Arctic amplification' [42–45]. The corresponding pattern for GISTEMP revealed maximum response over the northern continent (figure S3 available at stacks.iop.org/ERL/8/035049/mmedia). These differences were due to lack of data in the Arctic and different strategies for filling the voids.

For the GCR, there were indications of a cool anomaly over eastern Europe with high GCR flux. The amplitudes have magnitudes of ∼0.2 °C in both data sets. Others too have noted coincidental regional variations with solar activity, such as cold winters in the UK when there has been low solar activity [46]. The results presented here are for the entire year,
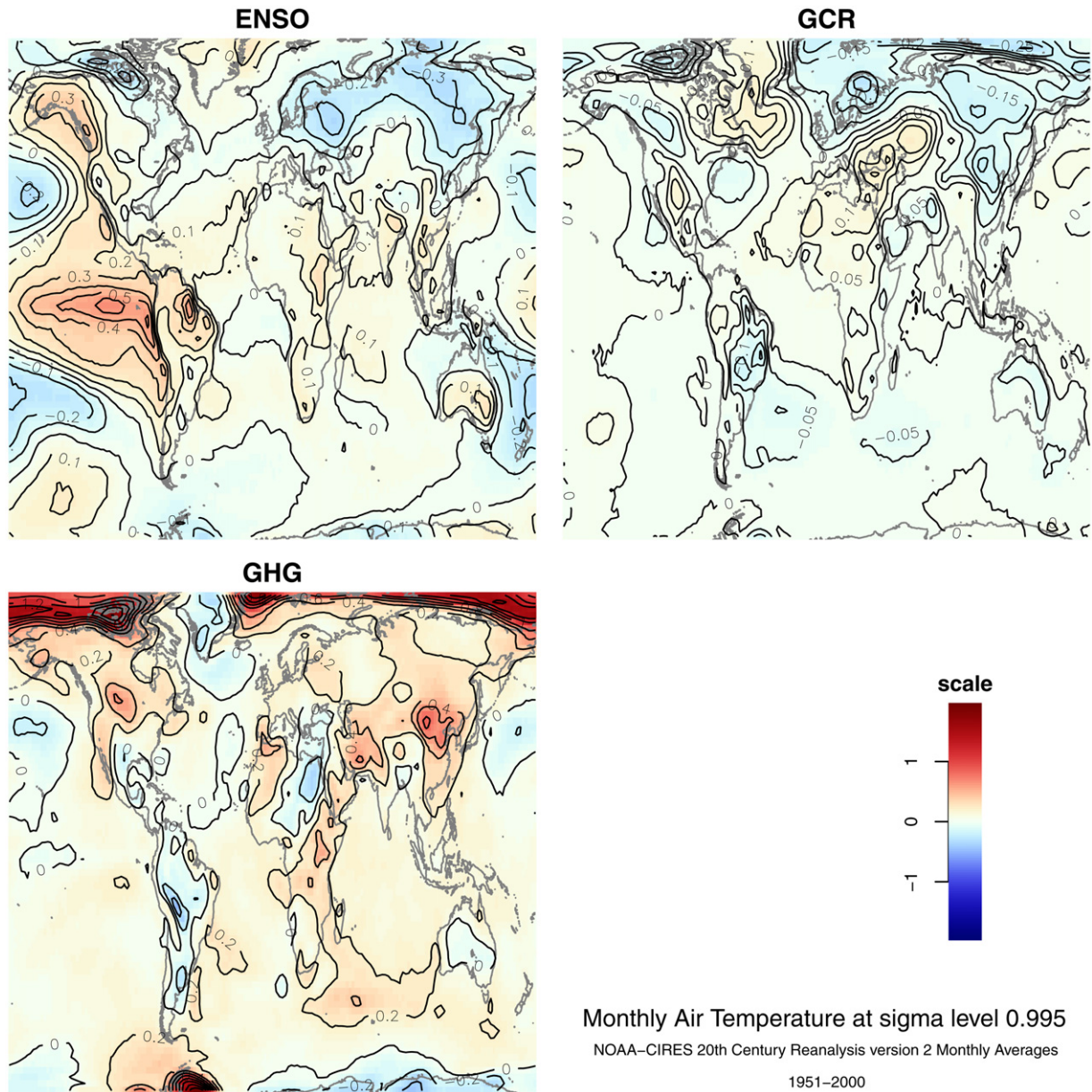
Figure 1. Geographical distribution of the regression coefficients (similar to spatial correlation) for ENSO, GHGs, and GCR. The signature of ENSO in the eastern equatorial region of the Pacific and the Arctic amplification are both prominent features. The main feature associated with GCR was a cold anomaly over eastern Europe. The R^2-value for the total patterns was 0.45, and estimated as the sum of the products between the multiple regressions for each EOF: $\sum_i R_i^2 d_i^2 / D$ where $D = \sum_j d_j^2$. The physical unit of colour scale is K per standard deviation (annual mean).

and indicated a similar direction as the previous findings (low sunspot number $\rightarrow$ high GCR $\rightarrow$ low T_{2m}), as GCR and the sunspots are anti-correlated. The regional amplitude of the maximum GCR response was about half of that of the GHG and ENSO response patterns.

Figure 2 shows the global mean temperature estimated from the product between the spatial response patterns shown in figure 1 and the co-variates. The black and dark grey curves show the GISTEMP and 20C reanalyses respectively, whereas the grey curve shows the sum of all co-variates. All curves show anomalies with a 1961–1990 base line. The regression analysis suggested that the GCR had a weak influence on the global mean temperature if any, as the variations associated

with the GCR were well within the range indicated by the Monte Carlo simulations (figure 3). Lagged cross-correlation between the GCR and GISTEMP gave no correlation above the level of statistical significance at the 5% level, and furthermore suggested highest correlation for when changes in $\langle T \rangle$ lead the GCR by one year (figure S4 available at stacks.iop.org/ERL/8/035049/mmedia). The results gave no evidence for any long-term trend over the 1951–2000 period associated with the GCR.
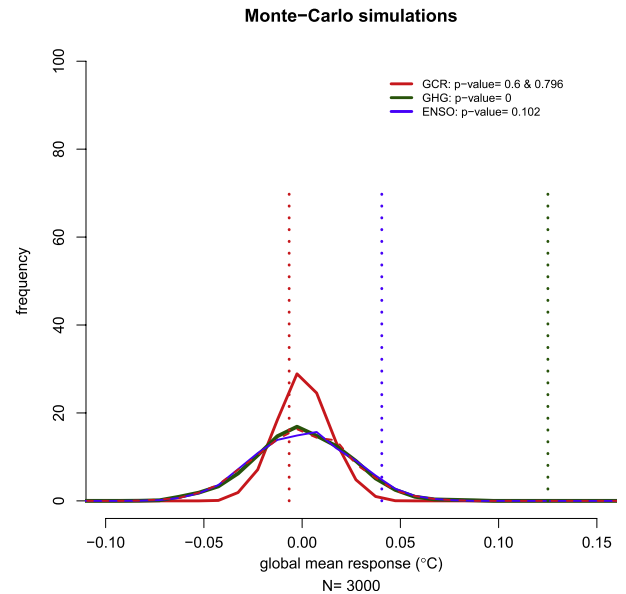
The residual from the fitted global mean temperature to the GHGs, GCR and ENSO can be used to reconstruct the spatial structures. The projection of this residual onto the EOFs gave a close match in the representation of the spatial
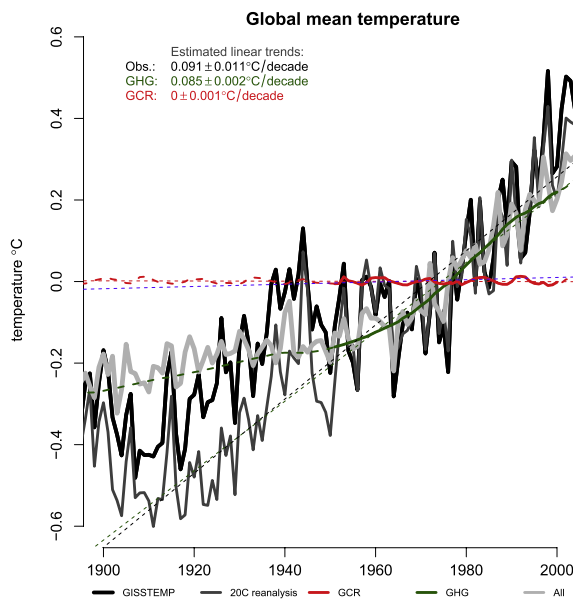
Figure 2. Estimated global contributions to the global mean temperature based on the product between the spatial patterns shown in figure 1 and the time series for ENSO, GHGs, and GCR respectively. The grey curve shows the sum of the three variables. GISSTEMP is in black and 20C reanalysis in dark grey.


Figure 3. The null-distribution for the global mean temperature response derived through Monte Carlo simulations, where the annual mean GCR was replaced by random values and repeated 3000 times. The solid red line shows the results for which only the GCR was replaced by random numbers whereas the dashed red line shows the results for GCR when all co-variates were set to random numbers. The global response was calculated from the global spatial temperature patterns.

modes (figure S5 available at stacks.iop.org/ERL/8/035049/mmedia), and the action of the residual was most pronounced over the polar regions and over high-latitude land areas in the northern hemisphere (figure S6 available at stacks.iop.org/ERL/8/035049/mmedia). However, some of the residual is expected to be connected with the error in the data, such sampling issues and density of measuring stations. Most of the trend in the global mean temperature could be attributed to GHGs as a slow and gradual increase with weak inter-annual variability. The predictions based on GHGs followed the observations until the 1990s, when there was a jump in GISTEMP and the 20C reanalysis $\langle T \rangle$. A scatter plot between the residual from all the predictions for $\langle T \rangle$ based on the co-variates in equation (1) and the respective co-variates gave no indication of further dependences that was not captured a linear model (figure S7 available at stacks.iop.org/ERL/8/035049/mmedia).

The ENSO pattern in p_{sl} included the well-known east–west southern oscillation structure as well as the Pacific-North-America pattern (PNA) and corresponding southern teleconnections (figures S8 and S9 available at stacks.iop.org/ERL/8/035049/mmedia). There was a stronger response in the extremely data sparse region just west of the Drake Strait, both in the 20C reanalysis and the HadSLP. A similar response was found for GHGs, albeit slightly further to the west and with the opposite sign. These were probably spurious and caused by insufficient data sampling. The spatial p_{sl} anomalies associated with GCR indicated high pressure over the British isles in both data sets. In the southern hemisphere, where the data coverage is sparse, the patterns in the different data sets were inconsistent, as HadSLP2 indicated a pronounced anomaly west off the Drake Strait that was absent in the 20C reanalysis. The amplitude of the GCR

response patterns was 1/2–1/5 of those for GHG and ENSO. The response pattern near northern Europe had a similarity to the north Atlantic oscillation (NAO), and is an interesting feature. It is possible that this is pure coincidence or that there is a real physical connection.

The wet-day mean 24-h precipitation μ was lower along the west-coast and eastern parts of the USA during ENSO years, and more intense in the interior parts (figure S10 available at stacks.iop.org/ERL/8/035049/mmedia). A somewhat similar spatial pattern was associated with the GHGs, except for the reduction in the east. The pattern associated with GCR indicated more intense μ in the west and less in the east. A histogram suggests that increases in the GHGs tend to favour higher values for the wet-day mean μ while greater GCR flux coincides with a slight reduction (figure 4). These results may suggest a weak tendency for a response in the cloudiness, as both μ and the wet-day frequency are consequences of cloud processes, although the coefficients are scattered around zero, and the bias in the distribution may be due to uneven geographical sampling. Hence, these results do not show a pronounced response, but may hint at the possibility of a weak effect in precipitation, mostly due to GHGs.

The frequency of rainy days f_w exhibited a clear large-scale coherent structure connected to ENSO, with more wet days in the west and central parts of the USA during El Niño years (figure S11 available at stacks.iop.org/ERL/8/035049/mmedia). Also the GHGs suggested a coherent spatial response, with increase almost everywhere. The results for GCR, on the other hand, revealed a heterogeneous pattern
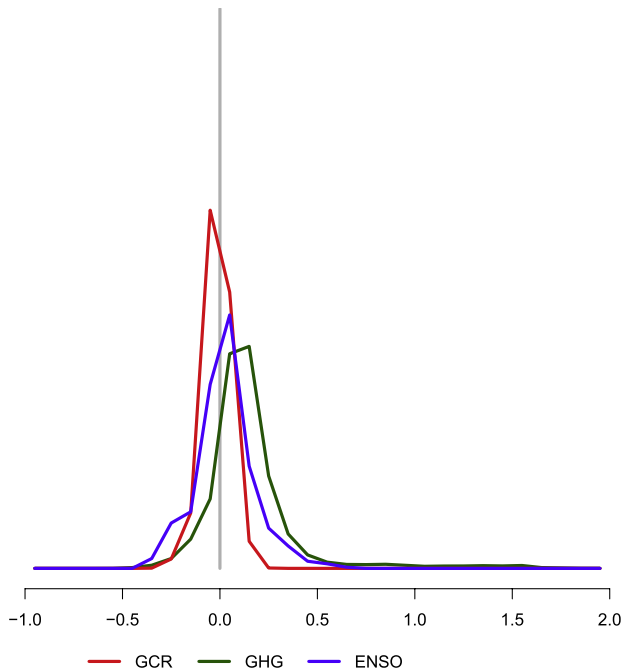
Figure 4. Histograms of the regression coefficients for the wet-day mean values μ associated with ENSO, GHGs, and GCR respectively. Increased levels of GCR coincided with slightly lower than usual μ whereas increased GHGs appeared to be associated with higher values for μ.

with both increases and decreases. High GCR-levels have coincided with fewer rainy days in south central parts of the USA, and more rainy days to the north of this region.

4. Discussion

There are several caveats associated with this analysis and these results: (1) the use of reanalysis; (2) assuming a linear response; and (3) because of strong co-linearity between GCR and other solar activity indicators, this study examines the potential influence of solar variability on Earth's climate, and not a specific theoretical solar-forcing mechanism.

Although reanalyses must be used with caution, the reanalysis variables examined were surface-variables that are strongly constrained by meteorological observations. However, the quality of these, as well as other gridded data, are expected to be lower over the oceans and remote regions of the world due to sparse sampling network.

The weak response patterns reported here may justify the assumption of linearity and that a linear approximation is sufficient to characterize the GCR/solar connection. Furthermore, there is no known reason why the GCR should have a more complicated effect with e.g. threshold values, and the hypothesized mechanism is a rapid effect on the low-level cloudiness and increased albedo [3].

Concerning co-linearity between solar activity and GCR, there is an asymmetry with respect to the results, and positive findings suggest that any of the solar mechanisms may be involved, but do not provide information as to which. It is possible that the effects shown here are due to

aspects connected to solar variations other than GCR such as variations in the total solar irradiance [47] or EUV. Negative results are reported here, on the other hand, suggest that none play an important role, as the present result may be interpreted as an upper limit of the effects from GCR on Earth's climate for timescales longer than a year. The results from the regression analysis suggested that GCR do not have a detectable effect on common meteorological variables, and from the p-value derived from a set of Monte Carlo simulations, the global mean temperature response associated with GCR was 0.6. The Monte Carlo used to generate the null-distribution did not take into account persistence (autocorrelation, long-term persistence), and is conservative. A null hypothesis taking the persistence into account is expected to give wider confidence interval, however, this will not change the results for the GCR-connection which was not statistically significant anyway.

Although it has been postulated that the GCR affect the nucleation of aerosols that may grow to CCNs, there are still missing links in the process leading up to cloud drops [14, 15, 48]. Furthermore, there is no known physical reason for why such processes should be limited to the Norwegian Sea region. A likely reason may by that the observed pattern may be just a coincidental result, and indeed, simple Monte Carlo simulations suggest just that. An interesting finding was a feature which could be associated with the NAO, however, it was not clear from the analysis whether this was due to coincidence or if there is a real physical link between the NAO and the GCR (the NAO characteristics may be due to it being a pronounced feature in the EOFs). If the NAO is chaotic and sensitive to external conditions near a set of bifurcation points, it is plausible that small perturbation could lead to a more pronounced response, however, a non-linear system may not necessarily respond the same way for similar conditions.

Here the GHGs and ENSO were included to provide a better fit for the GCR by reducing 'noise' as in [34]. These furthermore provided a check of the regression analysis in terms of spatial pattern and temporal evolution, as their 'fingerprint' and teleconnection are known *a priori*.

The analysis suggest that most of the trends can be attributed to GHGs, especially for T_{2m} and precipitation. There is no evidence suggesting that the recent global warming has any connection with the GCR. The results suggested that higher GHG concentrations were associated with spatially homogeneous increases in μ and f_w, and during ENSO years, the response in f_w has been spatially coherent of about 30 days over the central and south-western USA.

5. Conclusions

An analysis of spatio-temporal variability of near surface temperature, the mean sea-level barometric pressure, and precipitation data suggested that there is no evidence for the GCRs having contributed to the recent global warming. There was some interesting regional feature, however, which implies that a link between GCR and the NAO cannot be ruled out. There is no indication that GCR has a persistent effect on the most common meteorological elements measured by

the network of instruments, with the NAO being a possible exception.

Acknowledgments

References

[1] Benestad R 2002 *Solar Activity and Earth's Climate* (Berlin: Praxis-Springer)
[2] Israel M 2012 *EOS Trans. AGU* **93** 373
[3] Laken B A, Pallé E, Čalogović J and Dunne E M 2012 *J. Space Weather Space Clim.* **2** A18
[4] Ney E 1959 *Nature* **183** 451
[5] Dickinson R 1975 *Bull. Am. Meteorol. Soc.* 1240–8
[6] Marsh N and Svensmark H 2000 *Phys. Rev. Lett.* **85** 5004
[7] Harrison R 2008 *Proc. R. Soc.* A **464** 2575
[8] Harrison R and Ambaum M 2008 *Proc. R. Soc.* A **464** 2561
[9] Carslaw K, Harrison R and Kirkby J 2002 *Science* **298** 1732
[10] Sloan T and Wolfendale A 2008 *Environ. Res. Lett.* **3** 024001
[11] Harrison R G and Stephenson D B 2006 *Proc. R. Soc.* A **462** 1221
[12] Laken B, Wolfendale A and Kniveton D 2009 *Geophys. Res. Lett.* **36** L23803
[13] Čalogović J *et al* 2010 *Geophys. Res. Lett.* **37** L03802
[14] Pierce J R and Adams P J 2009 *Geophys. Res. Lett.* **36** L09820
[15] Kristjánsson J E *et al* 2008 *Atmos. Chem. Phys.* **8** 7373–87
[16] Pallé E and Butler C 2001 *Int. J. Climatol.* **21** 709
[17] Jørgensen T and Hansen A 2000 *J. Atmos. Sol.-Terr. Phys.* **62** 73–7
[18] Kirkby J *et al* 2011 *Nature* **476** 429
[19] Enghoff M, Pedersen J, Uggerhøj U, Paling S and Svensmark H 2011 *Geophys. Res. Lett.* **38** L09805
[20] Svensmark H 1998 *Phys. Rev. Lett.* **81** 5027
[21] Damon P and Laut P 2004 *EOS Trans. AGU* **85** 370
[22] Laut P 2003 *J. Atmos. Sol.-Terr. Phys.* **65** 801
[23] Benestad R E, Hygen H O, van Dorland R, Cook J and Nuccitelli D 2013 *Earth Syst. Dyn. Discuss.* **4** 451
[24] Hansen J 2009 *NASA/GISS, Giss Surface Temperature Analysis; Global Temperature Trends: 2008 Annual Summation* (http://data.giss.nasa.gov/gistemp/2008/)
[25] Compo G *et al* 2011 *Q. J. R. Meteorol. Soc.* **137** 1–28
[26] Allan R and Ansell T 2006 *J. Clim.* **19** 5816–42
[27] Peterson T, Daan H and Jones P 1997 *Bull. Am. Meteorol. Soc.* **78** 2145–52
[28] Legates D and Willmott C 1990 *Theor. Appl. Climatol.* **41** 11
[29] Legates D and Willmott C 1990 *Int. J. Climatol.* **10** 111
[30] Lanzante J 1996 *Int. J. Climatol.* **16** 1197
[31] Adler R *et al* 2003 *J. Hydrometeorol.* **4** 1147
[32] Bonnans J F and Shapiro A 2000 *Perturbation Analysis of Optimization Problems* (Berlin: Springer)
[33] Lean J L and Rind D H 2008 *Geophys. Res. Lett.* **35** L18701
[34] Foster G and Rahmstorf S 2011 *Environ. Res. Lett.* **6** 044022
[35] Benestad R and Schmidt G 2009 *J. Geophys. Res.* **114** D14101
[36] van Oldenborgh G J, Doblas Reyes F J, Drijfhout S S and Hawkins E 2013 *Environ. Res. Lett.* **8** 014055
[37] van Oldenborgh G J *et al* 2009 *Clim. Past* **5** 1
[38] Lorenz E 1956 Empirical orthogonal functions and statistical weather prediction *Scientific Report No. 1 Statistical Forecasting Project*
[39] Press W, Flannery B, Teukolsky S and Vetterling W 1989 *Numerical Recipes in Pascal* (Cambridge: Cambridge University Press)
[40] Usoskin I G, Mursula K, Solanki S K, Schüssler M and Kovaltsov G A 2002 *J. Geophys. Res.: Space Phys.* **107** 1374
[41] Philander S 1989 *El Niño, La Niña, and the Southern Oscillation* (New York: Academic)
[42] Francis J A and Vavrus S J 2012 *Geophys. Res. Lett.* **39** L06801
[43] Kumar A *et al* 2010 *Geophys. Res. Lett.* **37** L21701
[44] Serreze M, Barrett A, Stroeve J, Kindig D and Holland M 2009 *Cryosphere* **3** 11
[45] Stott P and Jones G 2009 *Geophys. Res. Lett.* **36** L10701
[46] Lockwood M, Harrison R G, Woollings T and Solanki S K 2010 *Environ. Res. Lett.* **5** 024001
[47] Kristjánsson J, Staple A, Kristiansen J and Kaas E 2002 *Geophys. Res. Lett.* **29** 2107
[48] Götz G, Mészáros E and Vali G 1991 *Atmospheric Particles and Nuclei* (Budapest: Akadémiai Kiadó)