INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **14** (2002) 2275–2282 PII: S0953-8984(02)32494-9

Lack of scaling in global climate models

D Vjushin[1,3], R B Govindan[1,3], S Brenner[2], A Bunde[3], S Havlin[1] and H-J Schellnhuber[4]

[1] Minerva Centre and Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel
[2] Department of Geography, Bar-Ilan University, Ramat-Gan 52900, Israel
[3] Institut für Theoretische Physik III, Justus-Liebig-Universität Giessen, Heinrich-Buff-Ring 16, 35392 Giessen, Germany
[4] Potsdam Institute for Climate Impact Research, D-14412 Potsdam, Germany

E-mail: govindan@yafit.ph.biu.ac.il

Received 9 January 2002
Published 22 February 2002
Online at stacks.iop.org/JPhysCM/14/2275

Abstract

Detrended fluctuation analysis is used to test the performance of global climate models. We study the temperature data simulated by seven leading models for the greenhouse gas forcing only (GGFO) scenario and test their ability to reproduce the universal scaling (persistence) law found in the real records for four sites on the globe: (i) New York, (ii) Brookings, (iii) Tashkent and (iv) Saint Petersburg. We find that the models perform quite differently for the four sites and the data simulated by the models lack the universal persistence found in the observed data. We also compare the scaling behaviour of this scenario with that of the control run where the CO_2 concentration is kept constant. Surprisingly, from the scaling point of view, the simple control run performs better than the more sophisticated GGFO scenario. This comparison indicates that the variation of the greenhouse gases affects not only trends but also fluctuations.

1. Introduction

Indications of weather persistence over months and seasons are known [1]. Using DFA and wavelet analysis, it has been shown recently [2, 3] that the variations of the daily temperatures from their average values (seasonal trend) exhibit a scaling law which is independent of the location of the site. Indications of such a persistence law through spectral analysis have also been found [4, 5]. It was found that the persistence, characterized by the autocorrelation function $C(s)$ of temperature variations separated by s days, approximately decays as $s^{-\gamma}$, with roughly the same correlation exponent $\gamma \cong 0.7$ for the different sites considered [2, 3]. The range of this universal persistence law exceeds one decade, and there is no evidence that this law breaks down even on longer time scales.

Climate models are important since they are believed to predict the climate changes that occur as a result of anthropogenic interference with atmosphere. Here we study seven

state-of-the-art models: GFDL-R15-a, CSIRO-Mk2, ECHAM4/OPYC3, HADCM3, CGCM1, CCSR/NIES, and NCAR-PCM. Each model has certain unique characteristics such as the numerical methods, the type of subgrid-scale parameterizations, the spatial resolutions, and the period of integration. However, all are based on the same principles, all have common variables such as temperature, pressure and precipitation and all are coupled to ocean dynamics.

As one test of the validity of the model simulations, we would expect the universal scaling law found in the real observations to also be identifiable in the virtual world as represented by the model outputs.

This paper is structured as follows: in section 2, we discuss the methodology of fluctuation analysis (FA) and detrended fluctuation analysis (DFA). In section 3, we consider the record analysis and discuss the results, and in section 4 we present conclusions.

2. Methodology

2.1. FA and DFA

In order to explain the methods, let us consider the monthly mean temperature data T_i measured at a meteorological station. The index i counts the months in the record, $\{i = 1, 2, \ldots, N\}$. Figure 1(*a*) shows a portion of the monthly temperature data (○) for Brookings simulated by the GFDL model. For eliminating periodic, seasonal trends we obtain the deviations of T_i, $\Delta T_i = T_i - \overline{T_i}$ (see figure 1(*a*), solid curve shows the climatological annual cycle, $\overline{T_i}$), from the average of the monthly temperature $\overline{T_i}$ for each month i, say January, which has been obtained by averaging over all years in the temperature series. Figure 1(*b*) shows the temperature fluctuations ΔT_i about the annual average. Rather than calculating $C(s)$ directly from ΔT_i, which is hindered by the possible non-stationarities in the data and by the finite length of the records, we study the fluctuations in the temperature 'profile', $Y_n = \sum_{i=1}^{n} \Delta T_i$ in windows of size s.

To this end, we divide the profile into $K = 2[N/s]$ non-overlapping windows of size s (indexed ν) starting from the beginning as well as from the end. In the simple FA we calculate the square of the difference of the profile at both ends of each segment ν and their average over all segments which gives the square of the fluctuation function

$$F^2(s) = \frac{1}{K} \sum_{\nu=1}^{K} (Y_{\nu s} - Y_{(\nu-1)s+1})^2. \tag{1}$$

This procedure is repeated for different length scales s to obtain a scaling relation between $F(s)$ and s. For long-range correlated data, $F(s)$ increases with s according to the following power law [6]

$$F^2(s) \sim s^{2\alpha} \qquad \alpha = 1 - \frac{\gamma}{2} \tag{2}$$

where α is the fluctuation or scaling exponent. If $\alpha = 0.5$, the signal is uncorrelated; if $\alpha > 0.5$, there are positive correlations; if $\alpha < 0.5$, the signal is anticorrelated.

The drawback of FA like the Hurst analysis and the power spectrum analysis is that $F(s)$ is strongly affected by the presence of trends. To eliminate trends in a systematic way and thereby learn about them we also employ several orders of DFA [7–9]. The method has been applied to a variety of records ranging from DNA sequences to atmospheric data through heart-beat time series, see e.g. [2, 10, 11].

In DFA, we fit the data of the profile inside each segment (ν) of size s by using a kth order polynomial function $p_\nu^k(i)$, which is called the local trend. The order of the polynomial used

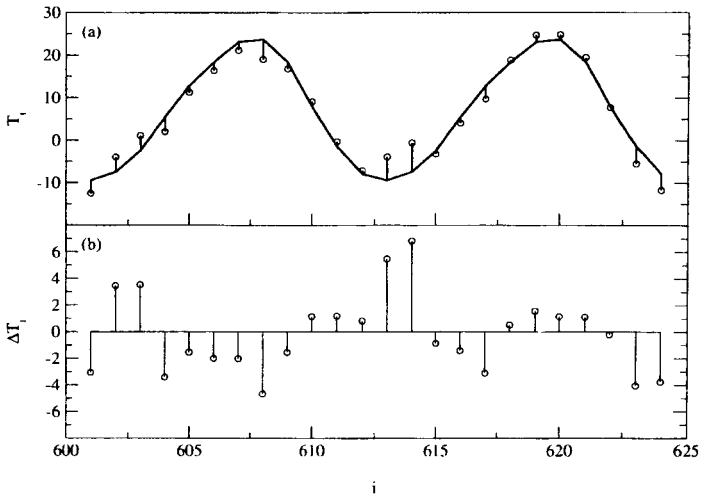

Figure 1. Interpolated monthly temperature data of Brookings simulated by the GFDL model. (*a*) Solid curve represents the annual cycle $\overline{T_i}$ for 2 years and the monthly temperatures are shown by O. (*b*) Temperature fluctuations around the mean value ΔT_i.

for the fit fixes the order of DFA. We detrend the profile $Y(i)$ by subtracting the local trend $p_\nu^k(i)$ in each box, and calculate the detrended fluctuation function

$$F_\nu(i) = Y_\nu(i) - p_\nu^k(i). \tag{3}$$

Thus, DFA order k eliminates the polynomial trends of order k from the profile and hence trends of order $k-1$ from the original signal.

For a given segment of size s, we calculate the DFA mean square fluctuation function

$$F^2(s) \equiv \frac{1}{K}\sum_{\nu=1}^{K}\frac{1}{s}\sum_{i=((\nu-1)s+1)}^{\nu s}[F_\nu(i)]^2 \sim s^{2\alpha}. \tag{4}$$

The mean square fluctuation function obtained from DFA has the same scaling behaviour as equation (2).

It is now quite clear from the above arguments that the trends will not be removed in the simple FA method, but they are systematically removed by different orders of DFA. In fact, the spectral method and the Hurst method lack this ability to distinguish between trends and correlations. In order to obtain a reliable estimate of the fluctuation exponent, one should increase the order of DFA until the convergence in the α value is achieved. Thus, by comparing the results of FA and DFA, one can identify both trends and correlations in the records.

3. Record analysis

In this section, we apply FA and DFA methods to the observed records at four locations in order to corroborate the earlier finding [2]. We then apply the same methods to the data simulated by seven leading GCMs for the same four sites.

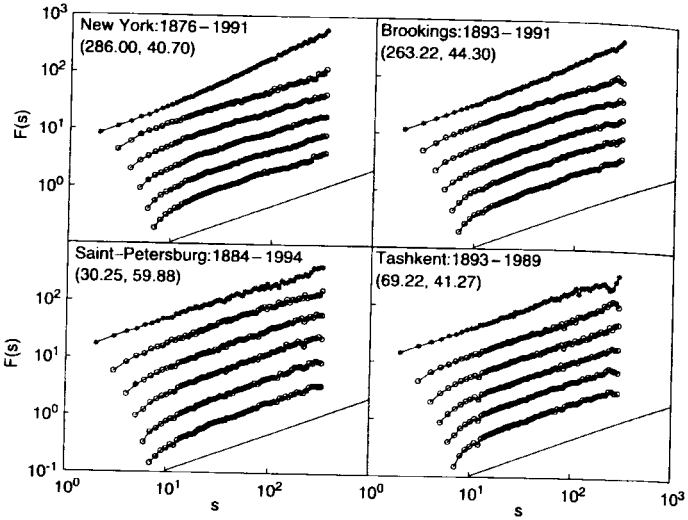

Figure 2. Results of FA (*) and DFA (O) for observed data. In each panel curves from top to bottom represent FA and DFA1–DFA5. The scale of fluctuation function is arbitrary. Unit of s is month. Solid line shown (for comparison) at the bottom of each panel is the theoretical line with a slope 0.65. Numbers within parentheses represent the geophysical co-ordinates of the considered site.

3.1. Application to the observed temperature data

Figure 2 shows the results of FA (∗) and DFA (○) for the observed data for the four sites New York (USA), Brookings (USA), Saint Petersburg (Russia) and Tashkent (Uzbekistan). In each panel, curves from top to bottom represent the results of FA and DFA1–5. After about 1 year the DFA curves are parallel to each other with a slope of 0.65 ± 0.02 (compare with the line with slope 0.65 shown at bottom). This finding is in agreement with earlier results [2].

In contrast the FA curves for New York (after 1 year), and Brookings (after 2 years) bend upwards yielding a higher exponent. This indicates the presence of a weak trend, probably due to urban warming, which is eliminated by the second order and subsequently by higher orders of DFA. Thus, it should be adequate to consider the results of DFA3 for comparing the scaling performance of the global climate models (GCMs).

3.2. Application to the data simulated by GCMs

Before considering the results of the FA and DFA, we will present some of the essential details of GCMs. During the past two to three decades, general circulation models have proven to be an indispensable tool for the study of geophysical fluid dynamics systems. These models provide numerical solutions of filtered forms of the Navier–Stokes equations devised for simulating meso-scale to large-scale atmospheric and oceanic dynamics. In addition to the explicitly resolved scales of motion, these models also contain parameterization schemes representating the so-called subgrid-scale processes, such as radiative transfer, turbulent mixing, boundary layer processes, cumulus convection, precipitation, and gravity wave drag. A radiative scheme,

for example, is necessary for simulating the role of various greenhouse gases such as CO_2. When run for the global domain, the only boundary conditions that need to be specified are those describing the effects of the surface of the Earth. A typical coupled atmosphere ocean (AOGCM) climate model will have a grid spacing of 300–500 km and 10–20 vertical layers as compared to weather forecasting models which have a grid spacing of 100 km or less and 30–40 layers. This difference in spatial resolution is due to the much longer integration times needed for climate models. The differences among the models developed by the leading research centres usually lie in the selection of the numerical methods employed, the choice of the spatial resolution, and the detailed formulation of the subgrid-scale parameterization schemes.

Each model is usually subjected to a rigorous evaluation procedure by its developers including consistency and sensitive tests. Models are also updated by comparing them with the historical data. These efforts have been, however, restricted to traditional time series analysis of overall characteristics of the entire simulation period. Recently, model inter-comparisons have become quite popular. Even in these inter-comparison projects the evaluations are mostly based on the same diagnostic tools as used for the individual model assessments and relate, therefore, only to mean states and simple measures of variability. In general, the traditional methods assume that the statistical properties of a signal remain the same throughout its unfolding period which is not true for the climate records due to the superimposed anthropogenic effect.

Today, selected results of state-of-the-art climate models are readily available through the Internet. Here we consider the monthly temperature records with the greenhouse gas forcing only scenario[5] from seven AOGCMs: GFDL-R15-a [12], CSIRO-Mk2 [13], ECHAM4/OPYC3 [14], HADCM3 [15], CGCM1 [16], CCSR/NIES [17], NCAR-PCM [18] that are available from the IPCC Data Distribution Centre [19]. We extracted the data for four representative sites around the globe (New York, Brookings, Tashkent and Saint Petersburg). For each model, we selected the four closest grid points to each site and bilinearly interpolated the data to the geographical location of the site.

Some of the primary characteristics of the models are listed in table 1. All models are global, include a full suite of subgrid-scale processes, and explicitly consider ocean dynamics. Note that HADCM3 is the only grid-point AOGCM while the other six models use the spectral method. All of the ocean models are grid-point models and do not necessarily have the same spatial resolution and grid as the atmospheric models. The periods for which the models were run are: GFDL-R15-a (1958–2057), CSIRO-Mk2 (1881–2100), ECHAM4/OPYC3 (1860–2099), HADCM3 (1860–2099), CGCM1 (1900–2099), CCSR/NIES (1890–2099) and NCAR/PCM (1871–2098).

All the models used a historical level of CO_2 until 1990 and after that they increased the CO_2 level at a rate of 1% per year. The latter effect should mainly cause an increase in the trend when we consider the entire record and should not affect the scaling. Therefore, if the models perfectly mimic the real data then they should yield a scaling exponent as found in the observed record. Results obtained from the real data (see figure 2) show that in double logarithmic representation, the higher order DFA (DFA2–5) curves are parallel to each other and hence for comparison it is appropriate to use the results of DFA3.

Results of DFA3 for $F(s)/s^{1/2}$ for the data simulated by the seven GCMs for the four sites New York, Brookings, Saint Petersburg and Tashkent are shown in figure 3. Also, the line with slope 0.15 (corresponding to the real data) is drawn at the bottom of each panel for comparison. It can be readily seen from figure 3 that, for a given site, all the models show

[5] In addition to this scenario there is one more scenario available for all models which is the greenhouse gas plus sulfate aerosol forcings.

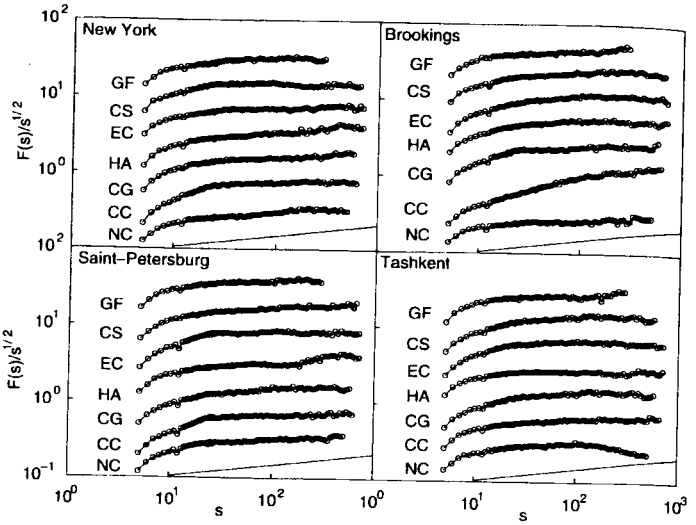

Figure 3. Results of DFA3 for four sites (New York, Brookings, Saint Petersburg, Tashkent) simulated by the seven GCMs, GF: GFDL, CS: CSIRO, EC: ECHAM4/OPYC3, HA: HADCM3, CG: CGCM1, CC: CCSR/NIES, NC: NCAR PCM. The curves shown in each panel are the DFA3 fluctuation function divided by $s^{1/2}$ obtained for different models for the site considered in that panel. Theoretical line with slope 0.15 is drawn at the bottom of each panel for comparison. The scale of fluctuation function is arbitrary. Unit of s is month.

Table 1. Primary characteristics of GCMs.

Model	Origin	Type	Atmospheric resolution	Coupling
GFDL-R15a	GFDL (Princeton)	Spectral	R15L9	Flux correction
CSIRO-Mk2	CSIRO (Australia)	Spectral	R21L9	Flux correction
ECHAM4/OPYC3	DKRZ (Hamburg)	Spectral	T4L19	Flux correction
HADCM3	Hadley Centre (UK)	Grid point	M96N73L19	Flux conservation
CGCM1	CCC (Canada)	Spectral	T32L10	Flux correction
CCSR/NIES	CCSR (Japan)	Spectral	T21L20	Flux correction
NCAR-PCM	NCAR (Colorado)	Spectral	T42L18	Flux conservation

wide differences in scaling behaviour. Some of the models show uncorrelated behaviour after a few years (see figure 3 ECHAM4/OPYC3 model and table 2) which implies loss of memory. For New York, the HADLEY model yields an exponent of 0.64 which is close to the observed data. None of the models perform well for Tashkent, Saint Petersburg and Brookings.

Table 2 summarizes the exponents obtained for the four sites simulated by different models. It can seen from table 2 that most of the exponent values are in the range of 0.50–0.56, although for some of the sites the exponents are closer to the observed data. Also, the mean and standard deviation values of these exponents are given in table 2 (in an ideal situation the mean and standard deviations of the scaling exponents obtained from the models should be 0.65 and zero, respectively), which indicates the inconsistency of the models.

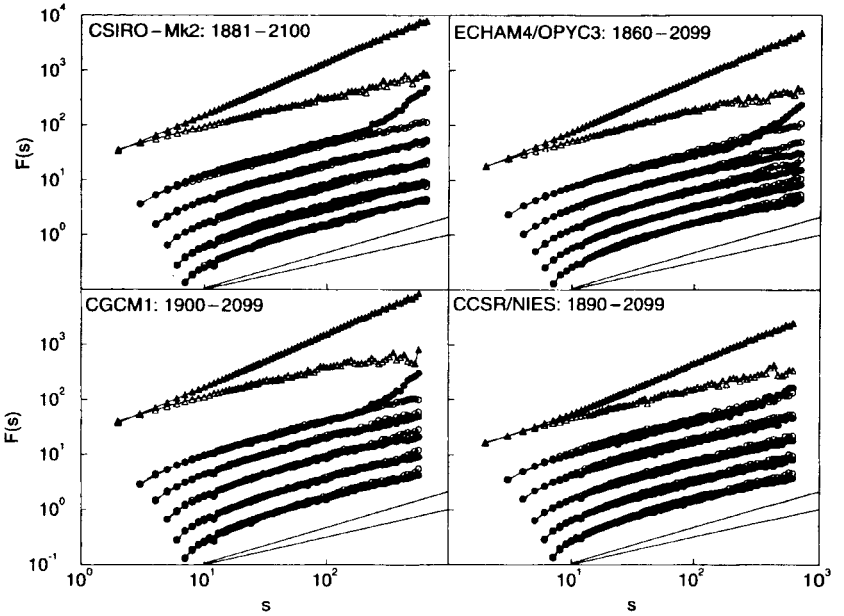

Figure 4. Comparison of fluctuation functions of GGFO scenario and CR for four GCMs. Curves from top to bottom in each panel show fluctuation functions of FA (△) and DFA1-DFA5 (○). Fluctuation functions of GGFO scenario are shown by solid symbols while those of CR are shown by open symbols. Theoretical lines with slopes 0.5 (bottom) and 0.65 (top) are drawn for comparison. The scale of fluctuation function is arbitrary. Unit of s is month.

Table 2. The fluctuation exponent α obtained for different models for different cities.

	Brookings	New York	St Petersburg	Tashkent	Mean	Std.
GFDL-R15a	0.6	0.55	0.57	0.6	0.58	0.02
CSIRO-Mk2	0.5	0.5	0.56	0.55	0.53	0.03
ECHAM4/OPYC3	0.5	0.56	0.51	0.5	0.52	0.02
HADCM3	0.52	0.64	0.57	0.51	0.56	0.05
CGCM1	0.55	0.57	0.57	0.55	0.56	0.01
CCSR/NIES	0.5	0.52	0.54	0.58	0.53	0.03
NCAR PCM	0.52	0.61	0.56	0.52	0.55	0.04

Next, we compare the scaling behaviour of the temperature data obtained from the GGFO scenario with that of the control run (CR). Figure 4 shows representative results for Tashkent for the four models: (i) CSIRO-Mk2, (ii) ECHAM4/OPYC3, (iii) CGCM1 and (iv) CCSR/NIES. The other three GCMs do not have data for the periods discussed in figure 4 for CR and hence we have not considered them for the comparison. Figure 4 shows the fluctuation functions of FA (△) and DFA1–5 (○) for the GGFO scenario (solid symbols) and CR (open symbols). Since trends are not eliminated in FA, fluctuation functions of FA for CR yield lower exponents when compared to those obtained for the GGFO scenario, as expected. For the case of CR, the

fluctuation functions of DFA1–5 show a scaling exponent close to 0.59 for CSIRO-Mk2, 0.65 for ECHAM4/OPYC3 and CGCM1, and about 0.6 for CCSR/NIES, which are less far from the universal exponent found in the observed data ($\alpha = 0.65$). This shows that the addition of increased CO_2 in the models causes the simulated results to lose the observed scaling (for comparison see the fluctuation exponents for Tashkent for the same models in table 2).

4. Conclusions

None of the GCMs for the GGFO scenario show scaling performance close to the observed data. The lower exponent values obtained for the GGFO scenario show that the models lack the persistence found in the observed data. In addition, the models display a wide range of scaling performance. Also, the comparison of the scaling exponents of CR versus GGFO scenario shows that the gradual addition of CO_2 makes the models lose their memory during the course of simulation. Even though the additions of CO_2 in the models are based on fair assumptions, they are not being done in a proper way. The introduction of a gradual increase of CO_2 causes some other side effects apart from its usual role of increasing the temperature, which causes the models to lack the universal scaling behaviour. A careful analysis of the methodology of the addition of CO_2 may certainly improve the performance of the models.

References

[1] Shukla J 1998 *Science* **282** 728
[2] Koscielny-Bunde E *et al* 1998 *Phys. Rev. Lett.* **81** 729
[3] Koscielny-Bunde E, Bunde A, Havlin S and Goldreich Y 1996 *Physica* A **231** 393
[4] Talkner P and Webber R O 2000 *Phys. Rev.* E **62** 150
[5] Pelletier J D and Turcotte D L 1997 *J. Hydrol.* **203** 198
[6] Bunde A and Havlin S 1995 *Fractals in Science* (New York: Springer)
[7] Peng C K *et al* 1994 *Phys. Rev.* E **49** 1685
[8] Kantelhardt J W *et al* 2001 *Physica* A **295** 441
[9] Hu K *et al* 2001 *Phys. Rev.* E **64** 1114
[10] Bunde A *et al* 2000 *Phys. Rev. Lett.* **85** 3736
[11] Buldyrev S V *et al* 1995 *Phys. Rev.* E **51** 5084
[12] Dixon K W and Lanzante J 1999 *Geophys. Res. Lett.* **26** 2749
[13] Gordon H B and OFarrell S P 1997 *Mon. Weather. Rev.* **125** 875
[14] Christoph M *et al* 2000 *J. Clim.* **13** 2536
[15] Pope V D *et al* 2000 *Clim. Dyn.* 123
[16] Flato G M *et al* 2000 *Clim. Dyn.* **16** 451
[17] Emori S *et al* 1999 *J. Meteorol. Soc. Jpn.* **77** 1299
[18] Washington W M *et al* 2000 *Clim. Dyn.* **16** 755
[19] http://ipcc-ddc.cru.uea.ac.uk/dkrz/dkrz_index.html