

# ARIMA representation for daily solar irradiance and surface air temperature time series

Olavi Kärner

*Tartu Observatory, Observatory 1, Tõravere, 61602, Estonia*

---

## Abstract

Autoregressive integrated moving average (ARIMA) models are used to compare long-range temporal variability of the total solar irradiance (TSI) at the top of the atmosphere (TOA) and surface air temperature series. The comparison shows that one and the same type of the model is applicable to represent the TSI and air temperature series. In terms of the model type surface air temperature imitates closely that for the TSI. This may mean that currently no other forcing to the climate system is capable to change the random walk type variability established by the varying activity of the rotating sun. The result should inspire more detailed examination of the dependence of various climate series on short-range fluctuations of TSI.

*Key words:* Solar variability impact, Time series analysis, Climate variability, Solar irradiance

PACS codes: 92.70.Qr, 95.75.Wx, 92.60.Ry, 96.60.Ub

---

---

*Email address:* olavi@aai.ee (Olavi Kärner).

## 1 Introduction

2 Since November 1978 a set of total solar irradiance (TSI) measurements from  
3 space is available, yielding a time series of more than 25 years. Presently,  
4 there are three TSI composites at the top of the atmosphere (TOA) avail-  
5 able on line. The ACRIM composite was published by Willson (1997) and  
6 updated by Willson and Mordvinov (2003). The PMOD composite was pre-  
7 sented by Fröhlich and Lean (1998a,b) and is updated periodically (Fröhlich  
8 2006). Recently a third composite, called IRMB, was presented by Dewitte et  
9 al. (2004). All three are constructed from the same original data, but use dif-  
10 ferent methodology to construct the series. The problem is that no one sensor  
11 collected data over the entire period from 1978. Fröhlich and Lean (1998a,b)  
12 found that no increase in solar irradiance had occurred in the 1980s and 1990s.  
13 Willson and Mordvinov (2003) found a TSI trend of 0.04% per decade during  
14 solar cycles 21-23. Due to somewhat different construction of the composite  
15 series a mutual analysis of their temporal variability is useful.

16 The first goal here is to fit autoregressive and integrated moving average  
17 (ARIMA) models for describing long-range temporal variability in these se-  
18 ries. The results are believed to show whether the different composing schemes  
19 produced the series still varying similarly in terms of the used models. Pro-  
20 vided that one and the same model type appears to be applicable for all three  
21 versions the difference in fitted parameter values might serve as a measure of  
22 difference between the versions.

23 It is assumed that some influence of the temporal variability of TSI can be  
24 carried over to the climate system response series. Thus, the second goal here

25 is to examine whether the same type of ARIMA models are applicable to  
26 represent temporal variability of the daily TSI and station based surface air  
27 temperature time series. The station based series are selected in order to get  
28 the longest available records for fitting.

29 The algorithm for fitting ARIMA models (Box and Jenkins 1976) is used here.  
30 This is due to an earlier experience of applying random walk type models in  
31 order to represent the temporal variability of various atmospheric temperature  
32 series (e.g. Gordon 1991, Kärner 1996, 2002). In the present paper our aim is  
33 to fit a model from the ARIMA family to daily series using a time step longer  
34 than one day. The longer time step leads to better fitting conditions because  
35 the TSI and temperature series are with varying intensity of non-stationarity  
36 in terms of the self-similarity parameter  $H$ . The latter is determined by means  
37 of the structure function (Monin and Yaglom 1975).

## 38 **2 Data**

### 39 *2.1 Three versions for TSI*

40 All three TSI versions are available on line. An updated version of  $I_{PMOD}$   
41 series (Fröhlich 2006) contains also unpublished data from the VIRGO Exper-  
42 iment on the cooperative ESA/NASA Mission SoHO. The version  $d41\_61\_0710$   
43 contains also an extended part by the proxy model calibrated during cycle 21.  
44 The file, created on Oct. 19 2007, and downloaded from PMOD/WRC  
45 <ftp://ftp.pmodwrc.ch/pub/data/irradiance/composite/> is used in the current  
46 study. That is actually longer (from March 6, 1976 to Oct. 12, 2007) than two  
47 other composites. The ACRIM and IRMB composites are downloaded from

48 <http://www.acrim.com/Data>

49 and <http://remotesensing.oma.be/solarconstant/>, respectively.

50 To get the best conditions a common time interval is chosen for comparison.

51 It coincides with the available interval for IRMB series, extending in between

52 Julian days from 2443829 to 2453264, corresponding to 9436 day interval from

53 November 16, 1978 to September 14, 2004. The series are composed on the

54 basis of the same satellite information. The latter is not available for all days.

55 Due to methodical differences in the resulting series have different number of

56 gaps during that period. For the whole period the IRMB series contains 240,

57 ACRIM 269 and PMOD 673 days with missing irradiance value. Main amount

58 of them are one day gaps between the existing irradiance data. But there are

59 also some longer periods with missing data, and the longest of them is 35 days

60 in 1992. One and the same simple scheme was applied to fill the gaps in all

61 three series. One day gaps were filled with the mean value of the neighbors.

62 The longer gaps were filled with constants obtained averaging the last and

63 first values before and after the gap, respectively. This guarantees that the

64 sample variance of the series is not increased during the operation.

65 Summary information about the analyzed TSI series is shown in Table ??.

66 First three columns show the series name, analyzed years and series length

67 in days, respectively. Fourth and fifth column show the mean and standard

68 deviation for each sample. Last three columns show some characteristics for

69 the daily and 56 day increments of TSI. The reason to use 56 day increments is

70 explained later in connection with modeling details. The sixth column shows

71 standard deviation for the daily increments. The value for  $I_{PMOD}$  appears to be

72 somewhat smaller than two others. A partial reason for that is likely the larger

73 portion of filled data. 7th and 8th columns show mean and standard deviation

74 for the 56-day increments. These are useful to characterize the modeling. It  
75 is remarkable that all three increment means are negative showing that the  
76 start date of the series is in a region of high solar activity and the end in a  
77 low activity region.

## 78 *2.2 Surface air temperature data*

79 Station based air temperature series are much longer than the TSI series. Two  
80 approaches are used to model them. The primary goal is a detailed compar-  
81 ison with the irradiance variability. For that purpose the same time interval  
82 9436 days starting from November 16, 1978 is selected from the available tem-  
83 perature record. Mean seasonal cycle is calculated for that period and the  
84 corresponding anomaly series variability is modeled.

85 Surface air temperature data from 10 meteorological stations in Europe and 4  
86 in Asia are used to fit an ARIMA family model to daily mean time series. Nine  
87 of them are downloaded from <http://eca.knmi.nl> (see Klein Tank et al. 2002  
88 for details). The Central England temperature (CET) series from 1772-2005  
89 (Parker et al. 1992) is downloaded from <http://www.badc.rl.ac.uk>. Data for  
90 4 Asian stations are downloaded from <http://www.meteo.ru>. These datasets  
91 contain daily mean air temperature values in the corresponding stations. The  
92 record length in series varies from 50 to 250 years. Four first columns in Table  
93 ?? are dedicated to the station name, latitude, longitude and record duration  
94 years, respectively. All the analysis is carried out using anomaly series in  
95 respect to local seasonal cycle. The seasonal cycle is calculated by averaging  
96 separately the observed temperature values for each calendar day (including  
97 February 29).

99 TSI series is more complicated to model than those for surface air temperature  
100 because they are evidently heteroscedastic, i.e. having unequal variances for  
101 different periods of solar activity. The reason is in variations in TSI due to  
102 variable solar output. Brightening of the Sun with increasing sunspot number  
103 is explained by existence of bright magnetic areas called faculae that are im-  
104 mediately concentrated in bright clumps around the sunspot groups. Foukal  
105 et al. (2006) give a detailed description how these dark and bright structures  
106 generate the TSI variations. The variation equals to the product of the struc-  
107 ture's projected area and its photometric contrast relative to the adjacent,  
108 undisturbed photo-spheric disc. Empirical models reconstruct a TSI variation  
109 by adding up these contributions using records of the changing areas of the  
110 spots and magnetic brightenings. These models can account for over 80% of  
111 the variance in the irradiance time series. This shows that additional contribu-  
112 tions to the TSI variation from photospheric temperature variations outside  
113 spots and magnetic brightenings must be small (Foukal et al. 2006).

114 To get an idea about the range of the sample variance change during various  
115 stages of the solar activity cycle an example is presented in the upper panel  
116 of Figure ???. It shows the sample standard deviation for  $I_{PMOD}$  and its daily  
117 increment calculated for the consecutive non-overlapping 56 day long time  
118 intervals. over the available period. The solar activity cycles 21 to 23 are clearly  
119 observable by both, the TSI and its increment series standard deviations.

120 The heteroscedasticity in time series brings about some problems in trend  
121 analysis e.g. McKittrick (2002) and generally compels us to use non-linear

122 models for describing their variability (Tong 1993). In the current study we are  
123 attempting to overcome the difficulties analyzing sub-series over a sparse grid  
124 separated from the original daily time series. The approach helps assimilation  
125 of the temporal variability in produced sub-series. so that their variance tends  
126 to be more homogeneous.

127 Surface air temperature series behave differently. As an example standard  
128 deviation time series for CET and its daily increment over the consecutive  
129 non-overlapping 56 day intervals is shown in the lower panel of Figure ???. The  
130 shown time interval approximately coincides with that for  $I_{PMOD}$  i.e. 1976-  
131 2006. Much more homogeneous short-range variance illustrates the difference.  
132 No solar activity change can be detected using that record.

133 The TSI distribution at the TOA does not precisely determine the amount of  
134 solar energy absorbed in the climate system. Due to a high albedo of clouds  
135 an essential part of it is reflected back to space (e.g. Webster and Stephens  
136 1984). This means that any existence of direct dependence between the TSI  
137 and temperature variations is unlikely. But time series analysis can help to  
138 examine the relationship in terms of statistical models.

### 139 **3 Methodical background**

140 The goal here is to fit a simple model to represent long-term variability in TSI  
141 and air temperature time series. Thus, one and the same modeling scheme  
142 is applied to both variables, TSI and temperature. Initially, it is useful to  
143 quantify variability in the series in order to determine possible changes in its  
144 intensity, and thus, get information for choosing the time interval necessary

145 to get rid of the short-range variability. The quantification is carried out by  
 146 means of the structure function that calculates the second moment of the  
 147 series  $X(t)$  increment  $X(t + \tau) - X(t)$  as a function of the increment interval  
 148  $\tau$  (Monin and Yaglom 1975):

$$149 \quad D(\tau) = \frac{1}{n - \tau} \sum_{i=1}^{n-\tau} [X(i + \tau) - X(i)]^2. \quad (1)$$

150 Using some simple transformation, it can be written as (e.g Kärner 2005);

$$151 \quad D(\tau) = \tau [C(0) + 2 \sum_{i=1}^{\tau-1} (1 - i/\tau) C(i)], \quad (2)$$

152 where  $C(i)$  stands for auto-covariance of the increment series  $x(t) = X(t +$   
 153  $1) - X(t)$  at the lag  $i$ . Equation ( ??) shows that the growth rate of  $D(\tau)$  is  
 154 determined by the correlations between the increments.

155 Estimating the growth rate as a function of  $\tau$ , it has been found that a con-  
 156 siderable number of geophysical variables present self-similar time series, i.e  
 157 such that produce the growth rate proportional to  $\tau^{2H}$  where  $H$  is a constant  
 158  $0 < H < 1$  in some interval from some  $\tau_1$  to  $\tau_2$  (e.g. Mandelbrot and van Ness  
 159 1968):

$$160 \quad D(\tau) \propto \tau^{2H}. \quad (3)$$

161 Examining the surface air temperature series has shown that they have two  
 162 regimes with the scale break approximately at two weeks (e.g. Lovejoy and  
 163 Schertzer 1986). The shorter scale variability is characterized by rapid growth  
 164 of  $D(\tau)$  and by the value  $H \approx 0.35$ . The longer scale i.e that starting approx-  
 165 imately at  $\tau = 30$  (days) is characterized by nearly saturated behavior (i.e  
 166  $H \approx 0$ ).

167 A short demonstration of the quantification is shown in Figure ?? using 3  
168 TSI and 3 air temperature series from the same time interval. The upper panel  
169 shows results for TSI and the lower one those for temperature. A comparison  
170 shows that the growth rate decreases rapidly at  $\tau$  values in between 4 and 16  
171 days. There appears to be somewhat different growth rate if  $\tau$  grows beyond  
172 two weeks. The curve for TSI samples keeps growing at moderate rate (corre-  
173 sponding to  $H \approx 0.1$ ) up to a hump at  $\log_2\tau = 11$ . The latter is an indication  
174 of cyclic behavior and the hump is situated at the scale showing half of the  
175 solar activity cycle length.

176 The growth rate for temperature series shows a tendency to saturation. A de-  
177 crease in the growth rate for  $D(\tau)$  while  $\tau$  grows larger than two weeks shows  
178 that the statistical properties of longer increments are changing in compari-  
179 son with those for the daily ones (Kärner 2005). The saturation means that  
180 longer range increment series have approximately one and the same variance,  
181 independent upon the increment interval.

182 Using linear models to represent short-range temporal variability in the daily  
183 surface air temperature series is complicated due the various problems includ-  
184 ing heteroscedasticity (Tong 1993, McKittrick 2002). The TSI series are new  
185 and no modeling attempts is published until now. Change in the growth rate  
186 is accompanied by changes in certain statistical properties of the time series  
187 increments. In the present study we are interested in longer range variability.  
188 Thus, it is justified to select the time step in the series long enough to be  
189 outside the scale of the rapid growth for  $D(\tau)$ .

190 According to the theory of fitting the ARIMA family models, primary informa-  
191 tion can be obtained from sample correlations between the initial time series

192 terms and their increments (Box and Jenkins 1976). To use all information  
 193 and simplify the correlation calculations between series of various increment  
 194 ranges it is useful to divide the initial series into groups of sub-series. Time  
 195 series  $X(t)$  where  $t = 1, 2, 3, \dots, n$  is unfolded into  $\tau$  sub-series  $j = 1, 2, \dots, \tau$   
 196 of increments as follows:

$$197 \quad x_j(t) = X((t + 1)\tau + j) - X(t\tau + j). \quad (4)$$

198 Here,  $\tau$  is the increment range and  $t = 1, 2, \dots, n/\tau - 1 = n_1$ . The variable  
 199  $x_j(t)$  is actually also a function of  $\tau$  but the corresponding index is omitted  
 200 because its value can be seen from the amount of sub-series. A useful value of  
 201  $\tau$  will be selected empirically from the region where  $D(\tau)$  is nearly saturated.

202 Calculation of covariance for the newly determined sub-series over the available  
 203 sample is straight forward.

$$204 \quad C_j(k) = \frac{1}{n_1 - k} \sum_{i=1}^{n_1-k} (x_j(i) - \bar{x}_j)(x_j(i+k) - \bar{x}_j) \quad (5)$$

205 Sample mean values  $\bar{x}_j$  for the sub-series can generally differ from zero (even  
 206 for anomaly series due to varying sub-series length in ( ??)). Thus, it is im-  
 207 portant to take them into account. Autocorrelations for the sub-series  $x_j(t)$   
 208 over a large  $\tau$  interval are calculated for every sub-series. In the current task,  
 209 correlations between the consecutive increments (i.e. for the lag  $k = 1$ ) appear  
 210 to be important because the sample values for larger lags ( $k > 1$ ) are practi-  
 211 cally zero (not shown). As every  $j$  corresponds to one determined sub-series  
 212 of non-overlapping increments, the averaging gives a representative value for  
 213 the whole sample  $r_\tau(1) = \tau^{-1} \sum_{j=1}^{\tau} C_j(1)/C_j(0)$ .

214 Comparison of the sample correlations  $r_\tau(1)$  calculated for the equal (9436 day  
 215 long) daily TSI and temperature series is shown in Figure ?? for the increment  
 216 ranges from 1 to 100 days. The upper panel shows the growth rate for three  
 217 TSI series and the lower panel for surface air temperature anomaly series,  
 218 corresponding to the observations in Central England, Prague and Zagreb.

219 Figure ?? shows that the value of  $r_\tau(1)$  for the TSI versions appears to  
 220 undulate as  $\tau$  increases with the same frequency for all versions. No actual  
 221 saturation is observed over the presented  $\tau$  range although the undulation has  
 222 low amplitude. One can observe a slow tendency of increasing correlation as  
 223  $\tau$  grows longer than 15 days.

224 The corresponding correlations between the temperature increments behave  
 225 differently. The lower panel in Figure ?? shows that the value of  $r_\tau(1)$  sat-  
 226 urates rapidly as  $\tau$  grows and stays near constant  $-0.5$  over a wide range of  
 227  $\tau$ . This is an indication that for all these increment intervals, the applicable  
 228 model should be of one and the same type. The situation where only the first  
 229 sample autocorrelation differs remarkably from zero indicates that the first  
 230 order moving average MA(1) model should be appropriate to model the vari-  
 231 ability of increments  $x(t)$ . This leads to the type ARIMA (0,1,1) for the series  
 232 of  $X(t)$  (Box and Jenkins 1976).

233 In the current situation, it is useful to write the model in general form (i.e  
 234 including the term  $\Theta_0$  to estimate whether the series show any permanent  
 235 tendency over the sample).

236

$$237 \quad x_j(t) = \Theta_{0,j} + a_j(t) - \Theta_{1,j}a_j(t-1), \quad (6)$$

238 where  $j = 1, 2, \dots, \tau$ ,  $\Theta_1$  is fitted coefficient, and  $a_j(t)$  is white noise (WN).

239 Averaging eq ( ??) gives  $\Theta_{0,j} = \overline{x_j}$  over each sub-series. The increment anomalies  
240 lies in respect to the mean values  $\overline{x_j}$  are used in order to estimate  $\Theta_1$ .

241 In order to build a model for daily values over a sparse grid we can select  
242 a  $\tau$  value over a wide interval in between 15 and 100 (or more) days. There  
243 is no basis to prefer any particular time step for the temperature anomaly  
244 series. The situation is different for TSI. Due to the solar rotation a logical  
245 time interval between the successive data points in the sub-series should be  
246 connected to the rotation period. At the equator the solar rotation (sidereal)  
247 period is 25.4 days. The synodic period of rotation is 27.3 days. It is the  
248 time for a fixed feature on the sun to rotate to the same apparent position as  
249 viewed from earth. The synodic period is longer because the sun must rotate  
250 for a sidereal period plus an extra amount due to the orbital motion of the  
251 earth around the sun. Thus, an astronomically justified time step should be  
252 an  $m$ -fold synodic period ( $m = 1, 2, \dots$ ). An attempt has been made to use 28  
253 day interval, but the residuals remained too correlated. Thus the interval of  
254 56 days was selected and it appears to be satisfactory.

255 The 56 day increment mean values  $\overline{x_j}$  are shown in seventh column of Table  
256 ??. Because the series length is not an  $m$ -fold of the solar activity cycle in  
257 any case, the trend estimates are meaningless. The disputable value of it,  
258 0.04% per decade (Willson and Mordvinov (2003)) is comparatively low. The  
259 MA(1) model is capable to describe stochastic level changes (Box and Jenkins  
260 1976) without special trend term. This means, that we can put  $\Theta_0 = 0$  fitting  
261 the model ( ??) to the TSI series. Numerous analyzes of temperature series  
262 have revealed the necessity to examine the existence of trends carefully (e.g.

263 Harvey and Mills 2003). Thus, the question about including the term  $\Theta_0$  must  
264 be solved separately for every series and sample length.

## 265 4 Results

266 Estimation of the coefficient and diagnostic testing are well known operations  
267 (see Box and Jenkins (1976) for details.) The coefficient  $\Theta_{1,j}$  is calculated  
268 by means of the maximum likelihood method, separately for each sub-series.  
269 The values of  $\Theta_{1,j}$  are then used with the same equation to compute the  
270 corresponding residuals, which are in turn used to test whether they can be  
271 treated as white noise (WN). In the current study, the portmanteau test (e.g.  
272 Box and Jenkins 1976) is used at the 99% significance level.

273 In the fitted models two sub-series for PMOD and one series for ACRIM and  
274 IRMB failed to pass the portmanteau test at 99% level. Thus, there appears  
275 to be a satisfactory accordance between the versions in terms of applicability  
276 of the model.

277 The fitted model depends on one coefficient  $\Theta_1$ . Distribution of its values over  
278 the sub-series can help a comparison of the TSI versions. The corresponding  
279 histograms for three versions are shown in the consecutive panels of Figures  
280 ?? and ??. The coefficients vary remarkably due to the fact that the sub-series  
281 contain different portions of values from high and low activity regimes. The  
282 variations inside of versions appear to have larger extent than those between  
283 the mean coefficients of the versions. The mean values are shown in the last  
284 column in Table ??. Internal variations for the coefficient are from 0.636  
285 to 0.804 (PMOD), from 0.645 to 0.766 (ACRIM) and from 0.620 to 0.756

286 (IRMB).

287 The comparison shows that the versions are equal in terms of the ARIMA  
288 models. The coefficient histograms are similar and consequently the explained  
289 variance undulates in the same interval, showing from 25% to 35% explained  
290 variance for the ACRIM and IRMB versions. The models for PMOD show  
291 a bit higher coefficient values and also from 23% to 40% for the explained  
292 variance. This may be caused due to larger fraction of filled irradiance values  
293 which evidently have some smoothing effect to the series.

294 The same model type is fitted to three (CET, Prague, Zagreb) station based  
295 temperature series of the same time interval as determined by the TSI record.  
296 The anomalies are computed in respect to the corresponding 26 year mean  
297 seasonal cycle. Number of used sub-series is 56 and all the fitted models passed  
298 the portmanteau test at 99% significance level (not shown). Fitted coefficient  
299 values for temperature sub-series are all remarkably higher than those for the  
300 irradiance series. This difference is caused by the different heteroscedasticity in  
301 the initial series. The difference can be followed in the charts showing growth  
302 rate for  $D(\tau)$  and saturation for  $r_\tau(1)$ , respectively. For the TSI series,  $r_\tau(1)$   
303 undulates around the value  $-0.45$  during the approximate self-similarity region  
304 over  $\tau$ . For the temperature series the saturation of it takes place near value  
305  $-0.5$ .

306 For fractional Brownian motion (FBM) series with the self-similarity parame-  
307 ter  $H$  (where  $0 < H < 1$ ), the correlation between the consecutive increments  
308 of the length  $\tau$  is written as (Mandelbrot and Van Ness 1968)

$$309 \quad r_\tau(1) = 2^{2H-1} - 1, \quad (7)$$

310 independent on the increment interval  $\tau$ . The situation enables us to compare  
 311 the behavior of  $D(\tau)$  and  $r_\tau(1)$  estimates obtained for TSI and temperature  
 312 series over the self-similarity region. These estimates of  $H$  on the basis of  $D(\tau)$   
 313 are approximately 0.1 and 0.0 for TSI and temperature, respectively. Using  
 314 eq. ( ??) and mean  $r_\tau(1)$  levels from Figure ?? gives the corresponding values  
 315 0.07 and 0.0.

316 Model ( ??) for the increments in case of  $\Theta_0 = 0$  can be written for the  
 317 anomalies as

$$318 \quad X(t) = \lambda \sum_{i=1}^{\infty} a(t-i) + a(t), \quad (8)$$

319 where  $\lambda = 1 - \Theta_1$ . The obtained version ( ??) shows that in case of the fitting  
 320 result appears to be  $\Theta_1 = 1$ , the model reduces to a trivial form where the  
 321 temperature anomaly series represents white noise if collected over the time  
 322 step of  $\tau$  days. The current results show that the coefficient appears to be very  
 323 close to unity for the temperature series and comparatively far from it for TSI.  
 324 In this case the model can be interpreted as random walk in noisy environ-  
 325 ment (Box and Jenkins 1976). This means that for the temperature series the  
 326 variability shows weak non-stationarity because the random walk generating  
 327 noise has low variance. The same is declared by nearly saturated  $D(\tau)$  esti-  
 328 mate. All the calculated characteristics ( $D(\tau), r_\tau(1), \Theta_1$ ) indicate that the TSI  
 329 series contain stronger non-stationarity than the temperature series. Due to  
 330 smaller  $\Theta_1$  variance of the random walk generating noise part is considerably  
 331 larger than in the temperature models.

332 The lower panel of Figure ?? shows two histograms of the fitted values of  
 333  $\Theta_{1,j}$  for CET series. The first one corresponds to the series from the same

334 interval with the TSI versions. Another histogram belongs to models fitted  
335 to the whole available 234 year time interval. The comparison is carried out  
336 in order to get an idea about the effect of long Gleissberg cycle in the solar  
337 activity to local temperature variability.

338 Available long temperature series from different stations are used to test  
339 whether the obtained model type is applicable for longer periods. An overview  
340 of parameter values obtained in the fitting of model ( ??) to the series from 14  
341 European and Northern Asian stations is shown in Table ???. The first four  
342 columns are dedicated to the station information. The proportion of sub-series  
343 that passed the test at 99% significance level in case of  $\tau = 64$  days is shown  
344 in the fifth column. The fraction of passed sub-series is remarkably high. Six  
345 series from 14 occasions had all 56 sub-series passed the portmanteau test.  
346 Seven series had one sub-series failing to pass and 1 series had two sub-series  
347 failing to pass at the 99% level. Provided that the increment interval ( $\tau$ ) can be  
348 chosen over a wide range of intervals, this means the approach can adequately  
349 describe the long-term variability of the temperature anomaly series; keeping  
350 in mind there may exist a few sub-series with the memory longer than two  
351 months. The sixth column shows the mean value of the coefficient  $\Theta_1$  obtained  
352 over 56 sub-series. The main amount of series lead to high mean values (from  
353 0.946 to 0.971). Only two stations (Kluchi and Ashghabat) show considerably  
354 weaker coefficient. The range between the values obtained for sub-series of one  
355 and the same station is similar to that shown in Figure ?? for the long CET  
356 record. On the basis of the examined temperature series the model does not  
357 depend essentially on the series length in terms of  $\Theta_1$ .

358 Different situation takes place if comparing the coefficients  $\Theta_0$ . The model  
359 ( ??) represents a stochastic trend (Box and Jenkins 1976) and a non-zero  $\Theta_0$

360 should be included if its component appears in every sub-series with the same  
361 sign and magnitude. In this case an additional deterministic trend component  
362 can be expected. Sample values for the average increment are used to estimate  
363 the  $\Theta_0$  coefficient. During the fitting process the values of  $\overline{x_\tau} = \tau^{-1} \sum_{j=1}^{\tau} \overline{x_j}$ ,  
364 where  $\tau = 56$ , computed for every dataset are shown in the seventh column of  
365 Table ???. The column shows that the average increment values obtained for  
366 14 series are mainly positive (11 from 14 occasions). They show the average  
367 trend over 56 day time interval. The 56 day increment mean values for sub-  
368 series can give an approximate answer to the first criterion. The mean values  
369 for different sub-series do not show any unanimity. In the series showing an  
370 average warming there exists about 10 to 20 sub-series showing some cooling.  
371 This means that including a deterministic trend term is not justified.

## 372 5 Conclusions

373 To represent the temporal variability in various climate related time series  
374 it is important to take into consideration that the model type may depend  
375 on the time interval between the consecutive terms in the initial series. The  
376 daily series representing TSI at the TOA and local air temperature in the me-  
377 teorological stations show strong short-range non-stationarity that weakens  
378 considerably together with increasing time scale (e.g Kärner 2005). Quantify-  
379 ing the variability by means of the structure function is a simple tool enabling  
380 us to select an appropriate interval. Modeling the variability of sub-series over  
381 the chosen time interval enables us to represent the long-range variability in  
382 the series.

383 Long-range temporal variability in total solar irradiance (TSI) and station

384 based surface air temperature anomaly time series is examined by means of  
385 the non-stationary ARIMA (0,1,1) models. The models are fitted to sub-series  
386 over 56 day time step. The interval was chosen on the basis of two solar rotation  
387 synodic period. The interval ensures the influence of short-range variability in  
388 the fitted model is minimized, and the model enables us to more precisely  
389 describe long-range variations in TSI and temperature, respectively.

390 Brightening of the sun with increasing sunspot number and its influence to  
391 the TSI is described by Foukal et al. (2006). The fitted model shows that the  
392 long-term variations in TSI can be described by means of random shock type  
393 increments. It also shows that due to a stochastic (and sample dependent)  
394 nature of the trend in TSI series the difference between the estimates by  
395 Fröhlich and Lean (1998a,b) and Willson and Mordvinov (2003) has a marginal  
396 value in terms of the model type. The versions are equal in terms of the model.

397 The same model type appears to be appropriate for representing variability  
398 in long daily surface air temperature time series. Comparison of the obtained  
399 values for the coefficient  $\Theta_1$  shows that the temperature models explain consid-  
400 erably higher part of the variance than the TSI models. This is an indication  
401 that there holds a strong negative feedback in the climate system capable  
402 to smooth the influence of non-stationary forcing. Modeling of 14 time series  
403 shows that the local climate (in terms of the air temperature) may be better  
404 balanced than it can be expected on the basis of the TSI fluctuations.

405 Generally, the same empirical model type for two series does not mean any  
406 inevitable dependence between the series. But in case if one series presents  
407 the main external forcing to the climate system and other a local response in  
408 various places of the system, at least some dependence is expected. The current

409 comparison shows that in terms of the model type the surface air temperature  
410 variability imitates closely that for the TSI. This may mean that during the  
411 analysed time interval no other forcing to the climate system was capable to  
412 change the random walk type variability established by the varying activity of  
413 the rotating sun. The result should inspire more detailed examination of the  
414 dependence of various climate series on the short-range TSI variations.

#### 415 *Acknowledgments*

416 Support was provided by the Estonian Science Foundation Grant 6814. Con-  
417 structive critique by Jaan Pelt is gratefully acknowledged.

#### 418 **References**

- 419 Box, G.E.P., Jenkins, G., 1976. Time Series Analysis, Forecasting and Con-  
420 trol, (revised edn). 575 pp. Holden-Day, San Francisco, CA.
- 421 Dewitte, S., Crommelinck, D., Mekaoui, S., Joukoff, A., 2004. Measurements  
422 and uncertainty of the long-term total solar irradiance trend. Solar Physics,  
423 224, 209-216.
- 424 Foukal, P., Fröhlich, C., Spruit, H., Wigley, T.M.L., 2006. Variations in solar  
425 luminosity and their effect on the Earth's climate. Nature, 443, 161-166. doi:  
426 10.1038/nature05072
- 427 Fröhlich, C. 2006. Solar irradiance variability since 1978. Revision of the  
428 PMOD Composite During Solar Cycle 21. Space Science Reviews, 125, 53-  
429 65, doi: 10.1007/s11214-006-9046-5
- 430 Fröhlich C., Lean J., 1998a. The Sun's Total Irradiance: Cycles, Trends and  
431 Related Climate Change Uncertainties since 1978, Geophysical Research Let-

432 ters, 25, 4377-4380.

433 Fröhlich C., Lean J., 1998b. Total solar irradiance variations: The construction  
434 of a composite and its comparison with models. In: Deubner, F.L., Christensen-  
435 Dalsgaard, J. and Kurtz D. (eds): IAU Symposium 185: New Eyes to See Inside  
436 the Sun and Stars. pp 89-102. Kluwer Academic Publ. , Dordrecht,  
437 The Netherlands.

438 Gordon, A.H., 1991. Global warming as a random walk. *Journal of Climate*,  
439 4, 589-597.

440 Harvey, D.I., Mills, T.C., 2003. Modeling trends in Central England temper-  
441 atures, *Journal of Forecasting*, 22, 35-47.

442 Kärner O., 1996. Global temperature deviations as a random walk. *Journal of*  
443 *Climate*, 9, 656-658.

444 Kärner O., 2002. On nonstationarity and antipersistence in global temperature  
445 series. *Journal Geophysical Research*, D107, 4415-4426. doi:10.1029/2001JD002024

446 Kärner O., 2005. Some examples on negative feedback in the Earth climate  
447 system. *Central European Journal of Physics* 3, 190-208.

448 Klein Tank, A.M.G. and Coauthors, 2002. Daily dataset of 20th-century sur-  
449 face air temperature and precipitation series for the European Climate As-  
450 sessment. *International Journal of Climatology*, 22, 1441-1453.

451 Lovejoy S., Schertzer, D., 1986. Scale invariance in climatological tempera-  
452 tures and the local spectral plateau. *Annales Geophysicae*, 4B, 401-410.

453 Mandelbrot B.B., and J.W. Van Ness, 1968. Fractional Brownian motions,  
454 fractional noises and applications, *SIAM Review*, **10**, 422-437.

455 McKittrick, R., 2002. Inference about trends in temperature data after con-  
456 trolling for serial correlation and heteroskedastic variance. *Proceedings of the*  
457 *Russian Geographical Society*, 164, 16-24.

458 Monin A.S. and Yaglom A.M., 1975. *Statistical Fluid Mechanics*, Vol 2., 683

459 pp, MIT Press Boston Mass.

460 Parker, D.E., Legg, T.P., Folland, C.K., 1992. A new daily Central England  
461 Temperature Series, 1772-1991. *International Journal of Climatology*, 12, 317-  
462 342.

463 Tong, H., 1993. *Non-linear Time Series*, 564 pp. Clarendon, Oxford.

464 Webster, P.J., Stephens, G.L., 1984. Cloud-radiation interaction and the cli-  
465 mate problem. In: Houghton J. (ed), *The Global Climate*, pp 63-78, Cambridge  
466 University Press.

467 Willson, R.C., 1997. Total solar irradiance trend during solar cycles 21 and  
468 22. *Science*, 277, 1963-1965.

469 Willson, R.C., Mordvinov, A.V., 2003. Secular total solar irradiance trend  
470 during solar cycles 21-23. *Geophysical Research Letter*, 30, 1199.

471 doi: 10.1029/2002GL016038

472 Figure captions.

473

474 1. Local standard deviation (in  $\text{Wm}^{-2}$ ) over the consecutive non-overlapping  
475 56 day long time intervals. Upper panel: For  $I_{P\text{MOD}}$  and its daily increments.  
476 Lower panel: For the Central England Temperature series over the same time  
477 interval 1976-2006

478 2. Growth rate of  $D(\tau)$  calculated for the same 1978-2006 interval. Upper  
479 panel: For three TSI versions. Lower panel: For three surface air temperature  
480 anomaly series.

481 3. Change of correlation between the consecutive increments as the increment  
482 interval increases from 1 to 100 days. Upper panel: For three TSI versions.  
483 Lower panel: For three air temperature series.

484 4. Frequency distribution of the coefficient  $\Theta_1$ , fitted to 56 sub-series of  $I_{P\text{MOD}}$   
485 series (upper panel) and for  $I_{ACRIM}$  series (lower panel).

486 5. Frequency distribution of the coefficient  $\Theta_1$ , fitted to 56 sub-series of  $I_{IRMB}$   
487 series (upper panel) and two versions of the CET series (lower panel)

488 Table captions.

489

490 1. Version name, series length, mean value, st. dev. ( $s_X$ ), st. dev. for daily  
491 increments  $s_x$ , mean and st dev for 56 day increments, and mean coefficient  
492  $\Theta_1$  calculated over the common interval

493 2. Station data, series length, fraction of passed test, mean coefficient, and  
494 mean increment for 56-day sub-series

Table 1

Version name, series length, mean value, st. dev. ( $s_X$ ), st. dev. for daily increments  $s_x$ , mean and st dev for 56 day increments, and mean coefficient  $\Theta_1$  calculated over the common interval

Version	Years	n(days)	Mean $\text{Wm}^{-2}$	$s_X$	$s_x$	$\overline{x_{56}}$	$s_{56}$	$\overline{\Theta_1}$
PMOD	1978-2004	9436	1366.04	.570	.183	-.0035	.512	.751
ACRIM	1978-2004	9436	1366.25	.717	.239	-.0082	.558	.693
IRMB	1978-2004	9436	1366.49	.731	.227	-.0042	.546	.689

Table 2

Station data, series length, fraction of passed test, mean coefficient, and mean increment for 56-day sub-series

Station, Country	LAT	LON	Years	$fr_{56}$	$\overline{\Theta}_1$	$\overline{x_{56}}$
Stockholm, SE	59.35	18.05	1756-2006	0.98	0.969	0.0009
Central England, GB	52.42	-1.83	1772-2005	0.98	0.970	0.0023
Prague, CZ	50.09	14.42	1775-2005	1.00	0.962	-0.0012
Bologna, IT	44.48	11.25	1814-2002	0.96	0.959	0.0024
Zagreb, HR	45.82	15.99	1861-2006	1.00	0.964	0.0046
Armagh, GB	54.35	-6.65	1865-2001	1.00	0.965	-0.0028
St. Petersburg, RU	59.96	30.30	1881-2006	0.98	0.951	0.0069
Archangelsk, RU	64.50	40.73	1881-2006	1.00	0.968	-0.0011
Lisbon, PT	38.72	-9.15	1901-2006	0.98	0.954	0.0052
Davos, CH	46.81	9.85	1901-2003	0.98	0.947	0.0078
Tomsk, RU	56.43	85.01	1884-1995	1.00	0.946	0.0068
Verhoyansk, RU	67.55	133.38	1926-1995	0.98	0.971	0.0045
Kluchi, RU	56.30	160.80	1931-1995	0.98	0.916	0.0279
Ashgabat, Turkm	38.00	58.4	1951-1995	1.00	0.912	0.0187

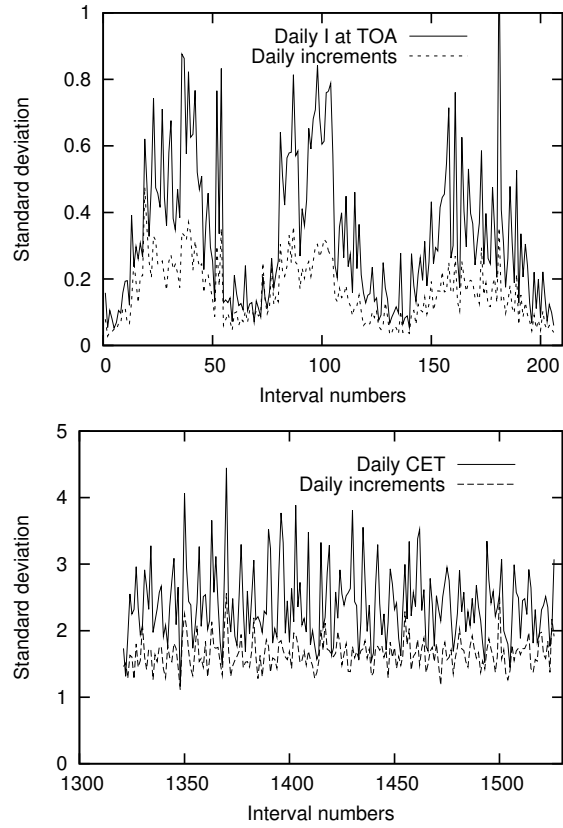


Fig. 1. Local standard deviation (in  $\text{Wm}^{-2}$ ) over the consecutive non-overlapping 56 day long time intervals. Upper panel: For  $I_{PMOD}$  and its daily increments. Lower panel: For the Central England Temperature series over the same time interval 1976-2006

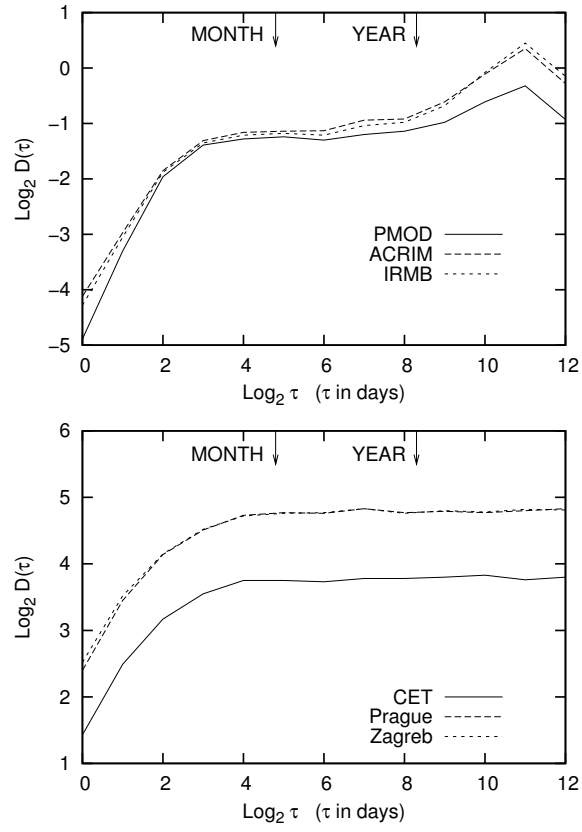


Fig. 2. Growth rate of  $D(\tau)$  calculated for the same 1978-2006 interval. Upper panel: For three TSI versions. Lower panel: For three surface air temperature anomaly series.

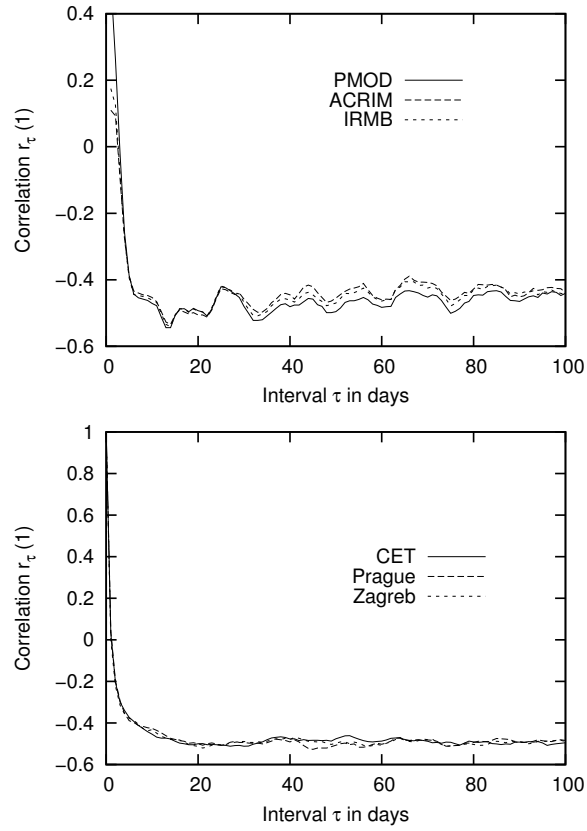


Fig. 3. Change of correlation between the consecutive increments as the increment interval increases from 1 to 100 days. Upper panel: For three TSI versions. Lower panel: For three air temperature series.

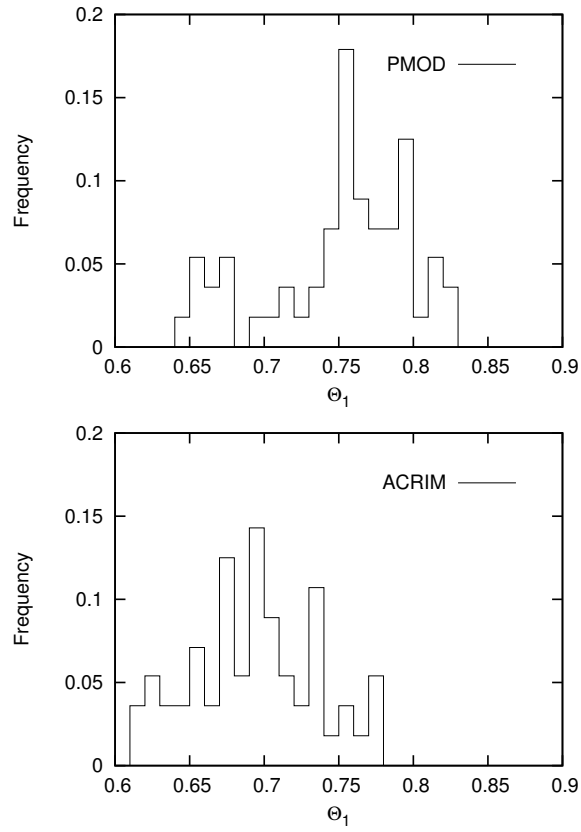


Fig. 4. Frequency distribution of the coefficient  $\Theta_1$ , fitted to 56 sub-series of  $I_{PMOD}$  series (Upper panel) and for  $I_{ACRIM}$  series.

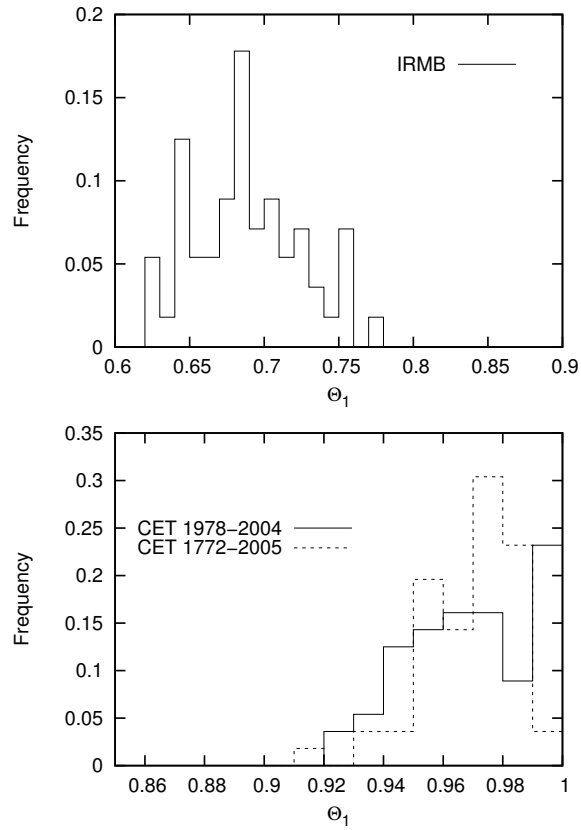


Fig. 5. Frequency distribution of the coefficient  $\Theta_1$ , fitted to 56 sub-series of  $I_{IRMB}$  series (Upper panel) and two versions of the CET series (Lower panel).