Tellus (1978), *30*, 193–200

A statistical study of a composite isotopic paleotemperature series from the last 700,000 years

By ERIK LUNDTANG PETERSEN and SØREN E. LARSEN,[1] *Meteorology Section, Risø National Laboratory, DK-4000 Roskilde, Denmark*

(Manuscript received December 28, 1976; in final form October 21, 1977)

ABSTRACT

An attempt is made to identify a stochastic model that makes the best fit to a generalized temperature curve covering the last 700,000 years. A search is made for the model within the family of auto-regressive, integrated, moving average models. All the models presented forecast a decline in temperature during the next 5000 years.

1. Introduction

With the fast-growing interest in climate and climate changes, the number and quality of time series indicating the past climate have much increased.

Recently, a study was presented by Emiliani and Shackleton (1974) containing the time series shown in Fig. 1, which displays an indicator of the temperature variations in portions of the northern hemisphere over the last 700,000 years.

In the present paper we have used methods from stochastic forecast and control theory to model this time series in terms of linear stochastic models. In the description of the procedure we have intended to illustrate both the method and characteristic statistical aspects of the data. Finally we have used the assigned models to forecast future behaviour of the series.

Acceptance of these forecasts as descriptive for future climate development must be based on both acceptance of the series as descriptive for the climatic past and on the adequacy of the applied models. Of those problems we shall here be concerned with the problems of assigning the models to the data. Furthermore only models based on the apparent memory in the series will be discussed. By excluding models containing a forcing term, e.g. Milankovitch's insolation theory

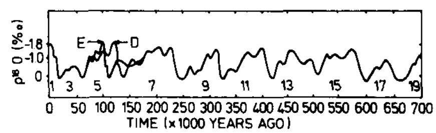

Fig. 1a. Two alternative composite isotopic paleotemperature series from the last 700,000 years (from Emiliani and Shackleton, 1974). The series are based on the ratio between O^{18} and O^{16} in piston cores from the Pacific and Atlantic Oceans and the Caribbean Sea.

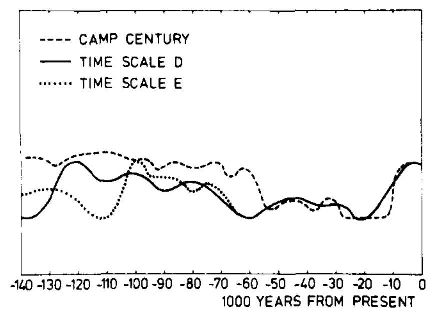

Fig. 1b. The last 140,000 years of the series from Fig. 1a together with a corresponding series from Greenland (from Johnsen et al, 1972). The curves are scaled so that they approximately coincide at the present and at the minimum of the last Ice Age.

[1] Presently on leave of absence at Department of Atmospheric Sciences, University of Washington.

(Monin, 1974), we avoid the difficulties associated with assigning the form of the forcing, as discussed by Kukla (1975) and Hays et al. (1976), and moreover we avoid linking our results to any specific physical model.

Principally, it would, of course, be more desirable to apply the technique used here on several different series at the same time to obtain an estimate of the correctness of the resulting model. However, this would be considerably more difficult, since the applied mathematical method was developed to handle series where the uncertainty is associated with the series values, while the time axis involved is assumed to be well known. On the contrary, for paleoclimatic series, the greatest uncertainty is associated with the time axis, and we have therefore chosen, as a compromise, to study one series, but a series for which the time axis has been established as a composite of many overlapping individual series (See Emiliani and Shackleton, 1974).

Mason (1976) reviewed some of the problems of reconstructing and interpreting historical records, and discussed to what extent the past may provide a guide to future changes. Further, he discussed some possible causes of climatic change, and described the promises and the limitations of numerical modelling for understanding and predicting the evolution of ther climate of the earth.

As the emphasis in the present study is on the statistical model building, readers interested in the representativeness of the series as climatic indicator and who wish to see the study in perspective are referred to the paper presenting the series and to Mason's text.

2. The data

Two alternative composite temperature series covering the last 700,000 years were presented by Emiliani and Shackelton (1974). The series are generated by comparing a large number of individual series based on analysis of the ratio between O^{18} and O^{16} in piston cores from the Pacific and the Atlantic Oceans and the Caribbean Sea. Their Fig. 4 displays the two different series, denoted D and E. These series correspond to two alternative scalings of the time axis. Both series were determined for this study from a magnified photograph of the figure for each 5000 years as mean values over 5000 years, giving 141 points per series. The resulting series are shown on Fig. 1a. For simplicity of presentation we have here chosen to present only analysis pertaining to one of the series. In Fig. 1b the last 140,000 years of the two series are compared with the O^{18}/O^{16} series from Camp Century in Greenland, covering the same time (Johnsen et al., 1972). From Fig. 1b it is seen that time scale D corresponds better to the Camp Century series time scale than does time scale E. Therefore we chose series D as the one for which to present the analysis.

3. Summary of the method of analysis

An appropriate statistical model for the time series was searched for within the general family of ARIMA (Auto Regressive Integrated Moving Average) models (Box and Jenkins, 1970). Denoting the series by Z_t and the generating white noise by a_t, $t = 1, 2, 3 \ldots N$, the general model may be written:

$$\phi_P(B^s)\,\phi_p(B)\,\nabla_s^b\nabla^c Z_t = \theta_Q(B^h)\,\theta_q(B)\,a_t \tag{1}$$

In (1) the notation is as follows:

B is the backshift operator:

$$BZ_t = Z_{t-1}$$
$$B^s Z_t = Z_{t-s} \tag{2}$$

Operators such as ∇_s^b and ∇^c are defined by:

$$\nabla_s^b = (1 - B^s)^b$$
$$\nabla^c = (1 - B)^c \tag{3}$$

The operator $\phi_p(B)$ is an autoregressive operator of order p:

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 \ldots - \phi_p B^p \tag{4}$$

The operator $\theta_q(B)$ is a moving average operator of order q.

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 \ldots - \theta_q B^q \tag{5}$$

The operators ∇_s, $\theta_Q(B^h)$, and $\phi_P(B^s)$ are called seasonal and can often be used with advantage if a series displays periodic (seasonal) behaviour with periods s or h. The problem of assigning models to a given time series can generally be separated into three parts: model identification, model fitting and model checking.

The fitting procedure is most straightforward: When a type of model has been tentatively specified (that is the values of P, p, q, Q, s, h, b, and c in (1) have been guessed), the remaining parameters in the model are fitted to the data by either linear or non-linear methods.

Model identification involves a general comparison between the statistical properties of the given series with those of the models.

Model checking involves the same comparison as in the identification procedure, but now in more precise terms. Furthermore, checking implies studies of the statistical properties of the residuals that are the estimates of the generating white noise process, a_t, in (1).

In this study we have used the following statistical indicators: Autocorrelation coefficients, partial correlation coefficients, power spectra, and the one-step prediction error. For given stationary statistical models, these quantities can be calculated straightforwardly from the model parameters (Box and Jenkins, 1970). A number of different methods exist for their estimation from raw data and residuals (x_t, $t = 1 \ldots N$), for which reason the calculation schemes used here are summarized below:

Mean value, $\bar{x}$

$$\bar{x} = \frac{1}{N} \sum_{t=1}^{N} x_t \tag{6}$$

Autocovariance of lag, C_k:

$$C_k = \frac{1}{N} \sum_{t=1}^{N-k} (x_t - \bar{x})(x_{k+t} - \bar{x}) \tag{7}$$

Autocorrelation coefficient of lag k, ρ_k:

$$\rho_k = C_k/C_0 \tag{8}$$

Power spectrum of frequency $f (= 1 \ldots (N-1)/2)$, $S(f)$:

$$S(f) = \frac{2}{N} \left| \sum_{t=1}^{N} x_t \exp(i2\pi f t) \right|^2 \tag{9}$$

The partial correlation coefficient of lag k, p_k, is equal to the kth coefficient, ϕ_k, obtained if an autoregressive process of order k is fitted to the data.

Equation (1) shows that the predictability of any given process one step ahead into the future is limited by the standard deviation of the generating white noise process, σ_a. Therefore estimates of σ_a are important measures of the predictive efficiency of a given model.

For a given series Z_t, $t = 1 \ldots N$, and a given fitted model, σ_a, can be estimated from the squared sum of deviations between the predicted value, $\hat{Z}_t$, and the actual values Z_t. We have used two estimates of the sum of squares:

$$\hat{S} = \sum_{t=\alpha}^{N} (Z_t - \hat{Z}_t)^2 \tag{10}$$

$$S = \sum_{t=1}^{N} (Z_t - \hat{Z}_t)^2 \tag{11}$$

In (10) the α-1 first Z_t's have been used to generate the first predicted element $\hat{Z}_\alpha$. In (11) $\hat{Z}_t$, $t \leqslant \alpha$, have been generated by letting $(Z_t, a_t) = (0, 0)$ for $t \leqslant 0$.

From (10) the unbiased estimate of σ_a is

$$\hat{\sigma}_a = \left(\frac{\hat{S}}{N - \alpha - \beta - 1} \right)^{1/2} \tag{12}$$

β being the number of adjustable parameters in the model.

From (11) a biased and an unbiased estimate of σ_a can be obtained:

$$\hat{\sigma}'_a = \left(\frac{S}{N} \right)^{1/2} \tag{13}$$

$$\hat{\hat{\sigma}}_a = \left(\frac{S}{N - \beta - 1} \right)^{1/2} \tag{14}$$

As further guidelines to model selection, we made use of two criteria (15) and (16) developed for stationary processes by Akaike (1974) and Parzen (1974), respectively.

$$\mathrm{AIC}(\beta) = \ln \hat{\sigma}'^2_a + \frac{2\beta}{N} \tag{15}$$

$$\mathrm{CAT}(p) = 1 - \frac{\hat{\sigma}^2_\infty}{\hat{\hat{\sigma}}^2_a} + \frac{P}{N} \tag{16}$$

According to both criteria, the model to use is the model that minimizes (15), respectively (16).

Equation (15) can be used for stationary models containing both autoregressive and moving average operators (4, 5). Equation (16) provides an estimate of the order, p, of the autoregressive operator that, given the finite data series, will describe as well as possible the behaviour of the autoregressive operator of infinitely high order, which is postulated to be the true operator for the series.

In (16) $\hat{\sigma}_\infty$ is an estimate of the prediction error for the true infinite autoregressive model estimated from the theoretical expression:

$$\ln \sigma_\infty^2 = \int_{-1/2}^{1/2} \ln (S(f))\, df \tag{17}$$

where the integral of the logarithm of the true power spectrum is estimated from (9) as described in Jones (1974).

4. Model identification

Figs. 2 to 5 display different statistical aspects of the raw data series. An autoregressive process of order p will give rise to autocorrelation coefficients, ρ_k, that in general can be described as a mixture of exponentials and sinusoidals as function of the lag number, k. The partial correlation coefficients for such a process will vanish for $k > p$. Therefore Fig. 2 indicates that an autoregressive process of order 3, AR(3), will be a good guess for the series.

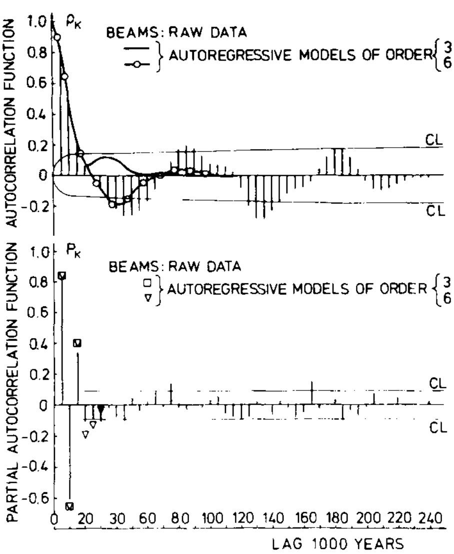

Fig. 2. Autocorrelation functions and partial autocorrelation functions for the temperature series and for two autoregressive models fitted to the series. Also shown are 95% confidence limits, assuming the series to be white noise (see e.g. Box and Jenkins, 1970).

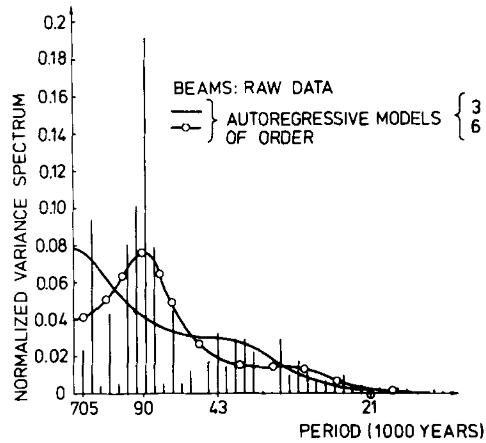

Fig. 3. Periodiogram of the temperature series and spectra of two autoregressive models fitted to the series. The horizontal axis is linear in frequency. The average periods of the orbital elements are indicated: Eccentricity, obliquity, and precession with periods of about 90,000, 43,000, and 21,000 years, respectively.

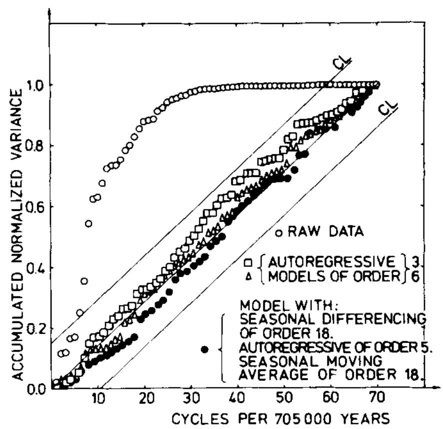

Fig. 4. Accumulated normalized periodograms for the temperature series and for the residuals of the fitted models. Also shown are the 95% confidence limits assuming white noise.

This guess is supported by Fig. 5, where the family of indicators derived from the variance of the residuals (12, 13, 14, 15, 16) is plotted as a function of the order, p, of an autoregressive process fitted to the series. From this figure it is seen that both AIC(p) and CAT(p), eqs. (15) and (16), have a very distinct minimum for $p = 3$.

It may therefore be concluded that an AR(3) model is supported by many clues from the data.

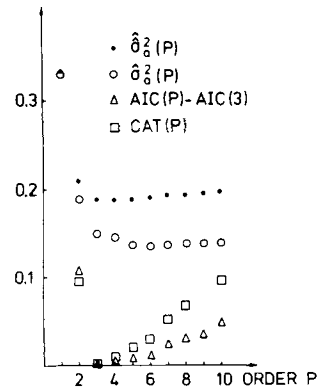

Fig. 5. Various indicators derived from the variance of the residuals plotted as function of the order p of autoregressive processes fitted to the temperature series.

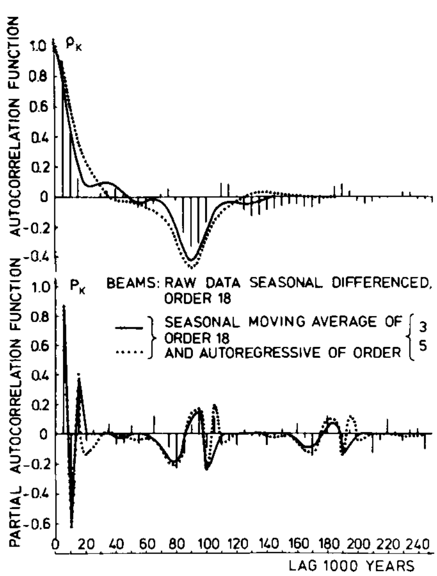

Fig. 6. Autocorrelation functions and partial autocorrelation functions for the seasonally differenced temperature series and for two models fitted to this series.

However, as will be discussed later, such a model fails to describe the periodic appearance of the series, which is best illustrated by the autocorrelation coefficient, ρ_k, on Fig. 2. From this figure it is seen that ρ_k, after an initial fall off, tends to repeat itself with a period of about 90,000 years (corresponding to a lag number period of about 18) with remarkably constant amplitude up to about 250,000 years, where the periodicity seems to break down.

Such behaviour can, of course, be modelled by increasing the order of the autoregressive operator, but it would demand very large values of p (of the order of 40) with a correspondingly large uncertainty on the estimates (cf. Fig. 5). It would therefore yield only a little more information on the structure of the series. Instead, we have chosen the more direct approach of postulating a periodicity of 90,000 years by applying the seasonal differencing operator (3) with s equal to 18.

The autocorrelation and partial correlation coefficients for the seasonally differenced series,

$$W_t = \nabla_{18} Z_t \tag{18}$$

are shown on Fig. 6. From the behaviour of the partial correlation coefficients it is guessed that the autoregressive part of the model is of the order of 3, while the large negative value of the autocorrelation coefficients at lag 18 (corresponding to 90,000 years) indicates a moving average operator coefficient at this lag.

Thus the resulting model is a combination of (18) and (19):

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3) W_t = (1 - \theta_{18} B^{18}) a_t. \tag{19}$$

Before proceeding with the estimation and checking procedures, we shall qualitatively discuss use of the model in (19) somewhat further.

The differencing operator in (19) implies that, apart from exhibiting a periodic pattern, the model is instationary in the sense that the level is set free to move over time spans of the order of 90,000 years (Box and Jenkins, 1970).

A physical explanation for the periodicity may be found in Milankovitch's theory of ice ages (Monin, 1974), which gives a variation of the incoming solar radiation with a period of roughly 90,000 years for the last million years. This periodicity, which is due to the variation of the eccentricity of earth's orbit, is found in many climate series, e.g. Hays et al. (1976). However, the eccentricity does not vary strictly periodically, and the approximately 90,000 years can be considered only an average period for the time interval considered here.

The discussion above might indicate that the differencing operator in (19), with its emphasis on a strict periodicity, places too much constraint on the model, even though it allows for instationarities in the series. In spite of this reservation we shall proceed with this model, (19), as a possible alternative to the autoregressive models, also entertained.

5. Model checking and fitting

The numerical values of the coefficients in the models considered were determined by computer programs developed by Jones (1974) and Box and Jenkins (1970), respectively. We shall not discuss details in these programs, but proceed by describing the checking of the models discussed in Section 3 and the additional models introduced as a consequence of the shortcomings of the first models revealed by the checking procedure.

The behaviour of the fitted models is compared with the raw data on the displayed figures and the models are summarized in Table 1.

Figs. 2 and 3 show that the AR(3) model describes the short-time memory of the data satisfactorily, while the behaviour of the series over longer time scales is unsatisfactorily modelled. As discussed in Section 3, this aspect of the model can be improved by increasing the order of the autoregressive operator (4). We have illustrated this by displaying the behaviour of an AR(6) model, this particular order being chosen because it gives a minimum value for the standard deviation of the noise, as calculated from (12).

Since the ARIMA (3, 18, 18) model given by (18) and (19) is instationary, it has no autocorrelation or power spectrum associated with it. Instead, on Fig. 6 we have compared model correlation coefficients for the stationary part of the model (19) with those of the raw data differenced as given by (18). As seen from this figure, the model describes the data quite well in terms of correlation coefficients. However, by varying the length of the autoregressive operator in the model it was found that the smallest σ_a (12) for these types of model was associated with an ARIMA (5, 18, 18) model which was therefore included in the study too.

The next checking procedure makes use of the fact that, for a proper model, the residual should be a sample from a white noise process. Fig. 4 shows integrated power spectrum for the raw data as well as for the residuals for some of the models. On this figure white noise plots as the straight diagonal line, and it is seen that all the residuals stay sufficiently close to the diagonal to be accepted as white noise.

6. Forecast

Fig. 7 shows the last 90,000 years of the analysed series and a forecast for the next 90,000 years as predicted by different models. It is inherently given by the structure of the models that

Table 1. *The table summarizes the statistics of the different models. The residual standard deviations are measured relative to the series standard deviation and are calculated through formulae (12) and (14). The CAT and AIC criteria are described in (15, 16). The constants appearing in the models AR(3) and AR(6) are corrections for series means*

AR(3): $(1 - 1.65B + 1.21B^2 - 0.49B^3)Z_t = a_t + 0.02$

AR(6): $(1 - 1.72B + 1.40B^2 - 0.57B^3 - 0.04B^4 + 0.08B^5 + 0.03B^6)Z_t = a_t + 0.03$

ARIMA (3, 18, 18): $(1 - 1.61B + 1.14B^2 - 0.37B^3)\nabla_{18}Z_t = (1 - 0.58B^{18})a_t$

ARIMA (5, 18, 18): $(1 - 1.69B + 1.28B^2 - 0.50B^3 - 0.02B^4 + 0.09B^5)\nabla_{18}Z_t = (1 - 0.74B^{18})a_t$

	Residual stand. dev.,			
Model	$\hat{\sigma}_a$	$\hat{\hat{\sigma}}_a$	AIC(β)-AIC(3)	CAT
AR(3)	0.39	0.40	0.00	0.00
AR(6)	0.37	0.40	0.01	0.03
ARIMA (3, 18, 18)	0.45	0.42	0.10	
ARIMA (5, 18, 18)	0.40	0.41	0.08	

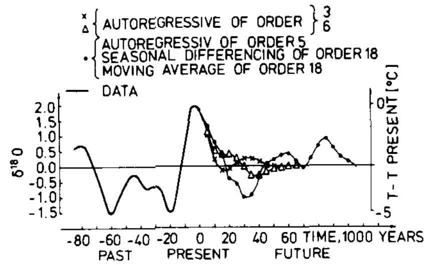

Fig. 7. Development of the temperature in the future based on three models fitted to the temperature series. Predictions are made in steps of 5000 years. The temperature range represents an average over the northern hemisphere (Emiliani and Shackleton, 1974).

the stationary ones tend towards the mean value as the memory dies out, while the seasonally differenced models tend towards a periodic pattern that repeats itself with a period of 90,000 years.

The figure displays how all the models discussed agree upon the forecast for the next 5000 year time step by forecasting a decreasing temperature. From Emiliani and Shackleton (1974) we shall take the temperature difference between a cold and a warm period to be $\Delta T \simeq 5$ °C for the part of the northern hemisphere, from which the series is obtained (see Fig. 7). Based on this scaling of the series, forecasts for the mean temperature for the first 5000 years for the different models are then given in Table 2.

7. Conclusion and discussion

From the point of view of model building, the present series has the disadvantage of including variations of many different time scales, of which the longest here discussed (90,000 years) is so long compared to the series length that it is difficult to represent it in a statistically reliable way.

From a statistical point of view there is no doubt therefore that, for the purpose of a one-step-ahead forecast, an autoregressive model of order 3 is most suitable, a fact that is clearly illustrated by Table 1 and Fig. 5. However, as discussed in the preceding section, this model clearly fails to reproduce the behaviour of the series on longer time scales.

The other three models here discussed are investigated in an attempt to introduce these longer time scales into the modelling. They can therefore be used to forecast further time steps into the future, but of course with less certainty. However, for the "immediate" future (5000–15,000 years) all models considered agree in their forecasts, indicating a future cooling trend over the time scales considered here. This, of course, being under the assumption that the time series used here is representative for the past climate and that the mechanisms controlling the climate are still the same as they have been for the last 700,000 years.

8. Acknowledgements

The authors acknowledge the help of Mr Jørgen Christensen and wish to thank Professor H. Panofsky for helpful discussions concerning Milankovitsh's theory. Finally, thanks are due to Professor W. Dansgaard and Mr S. J. Johnsen for providing us with the Camp Century series.

Table 2. *The table displays the expected change in mean temperature for the next 5000 years assuming the temperature scale shown in Fig. 7. Also given are the 95% confidence limits assuming the models to be true*

Model	Expected change in mean temperature for next 5000 years °C	95% confidence limits on temperature change °C
AR(3)	−1.3	−0.1, −2.4
AR(6)	−1.2	−0.1, −2.3
ARIMA(5, 18, 18)	−0.9	0.5, −2.3
ARIMA(5, 18, 18)	−0.9	0.2, −2.1

REFERENCES

Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, Vol. AC-19, *6*, 716–723.

Box, G. E. P. and Jenkins, G. M. 1970. *Time series analysis, forecasting and control*. San Francisco: Holden-Day.

Emiliani, C. and Shackleton, N. J. 1974. The Brunhes epoch: isotopic paleotemperatures and geochronology. *Science 183*, 511–514.

Hays, J. D., Imbrie, J. and Shackleton, N. J. 1976. Variations in the earth's orbit: pacemaker of the ice ages. *Science 194*, 1121–1132.

Johnsen, S. J., Dansgaard, W. Clausen, H. B. and Langway, C. C. Jr. 1972. Oxygen isotope profiles through the Antarctic and Greenland ice sheets. *Nature 235*, 429–434.

Jones, R. H. 1974. Estimation of the innovation generalized variance of a multivariate stationary time series. Technical Report No. 8, State Univ. of New York at Buffalo, Dept. of Comp. Science.

Jones, R. H. 1974. Identification and autoregressive spectrum estimation. Technical Report No. 8, State Univ. of New York at Buffalo, Dept. of Comp. Science.

Kukla, G. J. 1975. Missing link between Milankovitch and Climate. *Nature 253*, 600–603.

Mason, B. J. 1976. Towards the understanding and prediction of climate variations. *Quart. J. R. Met. Soc. 102*, 473–498.

Monin, A. S. 1974. *Earth's rotation and climate*. Delhi, India: Radok-Radhakrishna.

Parzen, E. 1974. Some recent advances in time series modelling. *IEEE Transactions on Automatic Control*, Vol. AC-19, *6*, 723–730.

СТАТИСТИЧЕСКОЕ ИЗУЧЕНИЕ СОСТАВНЫХ РЯДОВ ИЗОТОПНЫХ ПАЛЕОТЕМПЕРАТУР ЗА ПОСЛЕДНИЕ 700 000 ЛЕТ

Делается попытка построения стохастической модели, наилучшим образом удовлетворяющей обобщенной температурой кривой за последние 700000 лет. Делается поиск модели в семействе авторегрессивных, интегрированных моделей со скользящим средним. Все найденные модели дают прогноз падения температуры на последующие 5000 лет.