International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Statistical Analysis of Long Term Trends in Atmospheric CO2 Concentration at Baseline Stations

Antonovsky, M.Y., Buchstaber, V.M. and Zubenko, A.A.

IIASA Working Paper

WP-88-122

December 1988

WORKING PAPER

STATISTICAL ANALYSIS OF LONG TERM TRENDS IN ATMOSPHERIC CO_2 CONCENTRATIONS AT BASELINE STATIONS

M.Ya. Antonovsky
V.M. Bukhshtaber
A.A. Zubenko

December 1988
WP-88-122

IIASA
International Institute for Applied Systems Analysis

STATISTICAL ANALYSIS OF LONG TERM TRENDS IN ATMOSPHERIC CO_2 CONCENTRATIONS AT BASELINE STATIONS

M.Ya. Antonovsky
V.M. Bukhshtaber
A.A. Zubenko

December 1988
WP-88-122

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

Foreword

Carbon dioxide is one of several greenhouse gases that can modify the earth's heat balance by absorbing outgoing radiation from the earth's surface, thereby increasing the amount of heat retained by the atmosphere (the so-called greenhouse effect). Changes in CO_2 are therefore of considerable importance.

In this paper, the long-term trends are assessed at four baseline stations - Mauna Loa (Hawaii), Barrow (Alaska), American Samoa and South Pole. The authors conclude that a parabolic model provides the best fit for the observed rates of CO_2 concentration growth over the last 20-30 years.

I welcome Prof. Antonovsky's initiative in tackling this very important problem.

Bo R. Döös
Leader, Environment Program

STATISTICAL ANALYSIS OF LONG TERM TRENDS IN ATMOSPHERIC CO_2 CONCENTRATIONS AT BASELINE STATIONS

M.Ya. Antonovsky, V.M. Bukhshtaber and A.A. Zubenko**

1. INTRODUCTION

An assessment of possible future climatic changes requires a knowledge of trends in CO_2 concentrations in the atmosphere. Trend analysis is difficult because of annual, seasonal and daily fluctuations in CO_2 concentrations. In this paper we use algorithms to remove the periodic components. Then we assess long-term trends at four baseline monitoring stations.

2. DATA SET

The data set consisted of NOAA/GMCC time-series of CO_2 concentrations over the period 1968-82. These data included mean-monthly CO_2 concentrations for a set of stations (see Table 1), but we are mainly interested in the four monitoring stations: Barrow, Alaska (1973-82), Mauna-Loa, Hawaii (1974-82), American Samoa (1976-82) and South Pole (1975-82). All four stations are located at an approximately similar longitude of 170°. The total number of data is 400. These data were obtained on tape from the Carbon Dioxide Information Center, Oak Ridge National Laboratory, USA. The tape also included mean weekly data for 1968-82 (11397 readings), and hourly data for Mauna-Loa for the years 1958-86 (CO_2 concentrations and meteorological observations of pressure, temperature, relative humidity...) (254208 data points).

* All-Union Research Institute of Physiotechnical and Radiotechnical Measurements, USSR.

Table 1: Coordinates of stations of date base demonstrations (network of stations NOAA/GMCC)

Station	Notation	Longitude	Altitude	Region
Amsterdam Is.	AMS	77°E	37°S	Indian Ocean
Ascension Is.	ASC	14°W	7°S	S. Atlantic
St.Croix Is.	AVI	64°N	17°N	Carribean Sea
Azors Islands	AZR	27°W	38°N	N. Atlantic
Barrow	BRW	156°W	71°N	Alaska
Cold Bay	CBA	162°W	55°N	Alaska
Cape Meares	CMO	124°W	45°N	Oregon
Cosmos	COS	75°W	12°S	Peru
Falkland Islands	FLK	60°W	52°S	S. Atlantic
Guam Is.	GMI	144°E	13°N	N. Pacific
Key Biscane	KEY	80°W	25°N	Florida
Kumukahi	KUM	158°W	22°N	Hawaii
Mould Bay	MBC	119°W	76°N	Canada
Mauna Kea	MKO	155°W	20°N	Hawaii
Mauna Loa	MLO	155°W	19°N	Hawaii
Niwot Ridge	NWR	105°W	40°N	Colorado
Palmer	PSA	64°W	64°S	Antarctic
Point Six	PSM	110°W	47°S	Montana
Seychelles	SEY	55°E	4°S	Indian Ocean
American Samoa	SMO	170°W	14°S	S. Pacific
Ammundsen Scott	SPO	24°W	89°S	Antartic
"Charly" Ocean	STC			N. Atlantic
"M" Ocean	STM	2°E	66°N	N. Atlantic

3. METHOD

We assume that time series of monthly values can be represented in the following form:

$$C(Y,M) = C_0 + C_1(Y) + C_2(M) + \varepsilon(Y,M) , \quad (1)$$

where C_0 is a typical concentration for a given station, C_1 is the yearly variation, C_2 is the monthly variation and ε is a random fluctuation associated with other factors. Equation (1) allows us to analyze seasonal variations and long-term trends. To minimize bias we use medians rather than mean values (Huber, 1981).

The procedure is as follows (see Figure 1):

1. The medians for each year are calculated and the results are subtracted from the yearly data. The medians themselves are added to C_1, which for the first iteration is assumed to be equal to zero.
2. In the year-month matrix, medians for each month are calculated and the results subtracted from the data for the corresponding month. The monthly C_2 is assumed to be equal to zero for the first iteration.
3. The computed median of yearly effects is subtracted from each yearly C_1 and added to its computed value for the first iteration.
4. The computed median of monthly effects is subtracted from each monthly C_2 and added to its computed value for the first iteration.
5. This process continues until changes in the deviation of residuals are less than 1% of the residuals in the previous iteration. The deviation is measured by the sum of absolute values of residuals.

We assume that the changes of monthly effects are a reflection of biospheric seasonal cycling and that fluctuations are caused by local CO_2 sources. A spectral analysis is used to study the strengths of these factors.

4. RESULTS

4.1. *Spectral analysis*

A spectral analysis of ten years of monthly values for the Barrow station is given in Figure 2. The amplitude of harmonica with given frequency are presented here for each frequency point at the abscissa-axis (the module of discrete Fourie transformation at a given point). The abscissa point 1/120 corresponds to the period of oscillation once in 120 months $|w = 1/T|$. This harmonic is characteristic of the trend, i.e., the harmonic is equal to the period of observation.

In Figure 2, there also exist harmonics with periods equal to 12, 6, 4 and 3 months. The subtraction of yearly effects from the mean monthly concentrations (Figure 3) resulted in the extraction of the harmonic with the 120 month period, i.e, the trend component. Therefore, the long-term trends in CO_2 concentrations are characterized by the yearly effect.

The amplitude spectrum of the monthly effects (Figure 5) and residuals (Figures 4 and 7) show the precision of the expansion. By the criteria of the maximum entropy, it also shows that the monthly effects contain all evident harmonics (with periods of 12, 6, 4 and 3 months). A comparison of the monthly spectrum for different stations confirms the conclusion of previous studies (see, for example, Gemon et al., 1986) regarding the growth of amplitudes of seasonal oscillations in the direction from south to north, as well as the opposite phase of oscillations for the northern and southern hemispheres. This fact suggests that the monthly effects reflect the seasonal biosphere cycle, the behavior of the residuals reflect the local sources and sins of CO_2.

Consideration of the amplitude spectrum of the time-series for various stations (see, for example, Figures 3 and 6) shows that the signal with maximal amplitude always has a 12-month period. All stations have a 6-month harmonic, and most stations have clear 4- and 3-month harmonics.

4.2. ***Longitudinal patterns***

The behavior of C_0 in expansion (1) (Figure 8) shows the tendency of a decrease in value in the direction from north to south (Figure 10). However, deviations from this tendency have been observed in some cases. It would appear that remote stations have the lowest values for a given latitude and, therefore, provide an opportunity to consider C_0 as a background level of CO_2 concentration.

The similarity in behavior of the yearly effects (Figure 9) for different stations allow the construction of a global model of CO_2 changes based on data from a single station.

In contrast to the algorithms of the time-series analysis, the algorithm of the two-factor analysis (Tukey, 1977) has a property of statistically stable expansion into background trends and seasonal components of concentration variations.

4.3. ***Long-term trends at Mauna-Loa***

An analysis of long-term trends in annual values of yearly effects was done on the longest time-series (1958-87) for the Mauna-Loa Observatory. Evaluation of such trends has been considered in a number of researches (Keeling, 1984, 1987; Gemon et al., 1986; Pearman and Hyson, 1981, and Antonovsky, 1986). For this purpose, Pearman and Hyson (1981) for example, used a cubic spline approximation with a 1-year time step.

It is worth mentioning that the presentation of a long-term series of observations as some spline is a way of smoothing of experimental data that are slowly changing. More precisely, let z_i - mean monthly concentrations in month

x_i, $i=1,\ldots,n$. In the class C^2 of twice differentiated functions $\sigma(t)$ let us consider two functionals:

$$F_1(\sigma) = \sum_{i=1}^{n} |\sigma(t_i) - x_i|^2 \tag{1}$$

the functional of the least square method, and

$$F_2(\sigma) = \int_a^b |\sigma''(t)|\, dt \tag{2}$$

the Sobolian functional. Its minimization corresponds to choosing the most smoothly changing function σ. Here t_i is on the time axis corresponding to the month x.

It is found that the cubic spline $\sigma^0(t)$ of the data (t_i, x_i) gives the minimum of the functional

$$F(\sigma) = \lambda_1 F_1(\sigma) + \lambda_2 F_2(\sigma),\ \lambda_1, \lambda_2 > 0,\ \lambda_1 + \lambda_2 = 1\ ,$$

i.e., $\sigma^0 I(t) = arg \min_{\sigma \in C^2} F(\sigma)$

Thus, in Pearman and Hyson (1981), a combination of quasiparametric and extremal approaches to the construction of a trend is used.

The non-linear regression (for functions of a shape $\alpha_1 \exp \alpha_2 t + \alpha_3$) BMDP3R gives the following best approximations:

$$c(t) = 312.546 \exp(0.0033t) + 0.0193 \tag{3}$$

The RMS deviation of the approximation from the observed curve is 0.9. The results are shown in Figure 11. An inflexion point on the residuals curve (Figure 12) suggests that a piecewise approximation with two exponents might give a better result. The inflexion point coincides, approximately, to the year 1969. A non-linear regression which was done for the periods 1958-1969 and 1970-1987 respectively, gives:

$$c(t) = \begin{cases} 314.4\ e^{0.0024t} + 0.001 \text{ if } t \in (1958, 1969) \,|\, D = 0.07 \,| \\ 308.1\ e^{0.0040t} + 0.245 \text{ if } t \in (1970, 1987) \,|\, D = 0.29 \,| \end{cases} \tag{4}$$

A polynomial regression BMDP5R gives the best approximation of all curves of the form $\alpha_1 t^3 + \alpha_2 t^2 + \alpha_3 t + \alpha_4$ as the following second order polynomial

$$c(t) = 0.02t^2 + 0.55t + 314.74 . \tag{5}$$

The RMS deviation for the regression (5) is equal to 0.15 (Figure 13). Now the residual behavior becomes rather random (Figure 14). The characteristics of all three models are given in Figure 15. Thus, the parabolic approximation is better than the exponential for the whole interval (1958-86).

When we say that one model is better than another, we should clarify in what sense. Suppose that by the method of nonparametric estimation over the interval $[0,T]$ (in the case of Mauna Loa, T = 30_{years}), we have found a trend, i.e., trend $\tau = \tau(t)$ is defined by points $\tau(t_1), \ldots, \tau(t_n)$. For construction of a parametric model, let us choose a family of functions $\sigma(t;\alpha, \ldots, \alpha_k)$, where σ is a given functional, and $\alpha_1, \ldots, \alpha_k$ are the parameters estimated during the analysis. In our case:

$$\sigma_1(t;\alpha_1, \ldots, \alpha_n) = \alpha_1 t^3 + \alpha_t t^2 + \alpha_3 t + \alpha_n ,$$
$$\alpha_2(t;\alpha_1,\alpha_2,\alpha_3) = \alpha_1 e^{\alpha_2 t} + \alpha_3 .$$

Let us choose $T_1 < T$ and let us consider the functional of the method of least squares (MLS):

$$F(T_1;\alpha_1, \ldots, \alpha_k) = \sum_{l=1}^{T_1} (\tau(t_l) - \sigma(t_l;\alpha_1, \ldots, \alpha_k))^2$$

Let us put:

$$(\alpha_1^*, \ldots, \alpha_k^*) = arg \min_{\alpha_1, \ldots, \alpha_k} F(T_1;\alpha_1, \ldots, \alpha_k) ,$$

i.e., we find $\alpha_1^*, \ldots, \alpha_k^*$ by the method of least squares.

In the framework of a given family of functions, we find the optimal number of parameters k with the aid of F-criteria as in BMDP.

As a result, we get a function of t (model):

$$\sigma(t;\alpha_1, \ldots, \alpha_k^*) .$$

The first method of comparing models.

We calculate residuals of predictions:

$$\frac{1}{T-T_1} \cdot \sum_{t=T_1\,1}^{T} (\tau(t_t) - \sigma(t_t;\alpha_1^*, \ldots, \alpha_k^*))^2$$

This residual is a function of T_1. We say that model $\sigma_1(t;\alpha_1^*, \ldots, \alpha_{k1}^*)$ is better than model $\sigma_2(t;\alpha_1^*, \ldots, \alpha_{k2}^*)$ if the residuals of predictions by model $\sigma_1(;)$ are less than residuals of predictions by model $\sigma_2(;)$.

The second method of comparing models.

Let us define the residuals of predictions in the form:

$$\Delta(T_1) = \frac{1}{T} \sum_{t=1}^{T} (\tau(t_t) - \sigma(t_t;\alpha_1^*, \ldots, \alpha_k^*))^2$$

We say that model $\sigma(t;\alpha_1, \ldots, \alpha_k)$ is well on the level ε, if $\Delta(T_1) \leq \varepsilon$, where ε is a given number. Let T_q^* be the first value for that model $\sigma_q(t;\alpha_1, \ldots, \alpha_{kq})$, q =1,2, is well on the level ε.

We say that model $\sigma(t;\alpha_1^*, \ldots, \alpha_k^*)$ is better than model $\sigma_2(t;\alpha_1^*, \ldots, \alpha_{k1}^*)$ on the level ε, if $T_1^* < T_2^*$.

The calculation by both methods has shown that the long-term trend of series of concentration of CO_2 on observatory Mauna-Loa is better described by a parabolic model than an exponential one.

Let us consider the forecast ability of a parabolic model. As can be seen from Figures 16 and 16', a parabola obtained from the data for the first 5-year period deviates over the complete 28-year period. However, a parabola obtained from data for the first 15-years gives an almost exact forecast for the next 13-years (Figures 17 and 17'). Therefore, the observed present CO_2 concentration was possible to predict in 1973 on the basis of the parabolic model, i.e., the observed rates of CO_2 concentration growth are still the same as in the previous 10-years.

Let us make an analogous investigation of the temperature time-series for the Mauna-Loa station over the period 1958-1986. An amplitude spectrum of the temperature time-series (Figure 18) shows the presence of trend and two harmonics in a period of 12-months and of approximately 3-years. Those harmonics of more than a 12-month period disappear after subtraction of the yearly effects (Figure 19). This means that the curve of the yearly effects contains information on global behavior. The behavior of the monthly effects evidently reflects seasonal (winter-summer) temperature oscillations (Figure 20). The large year-to-year variability does not allow construction of a simple analytical model. Nevertheless, it is possible to say that, over the past 10 years, the mean annual temperature has increased. The second order polynomial which gives the best approximation is as follows:

$$c(t) = 0.01t^2 + 0.22t + 0.99 \; (D = 0.35) \,. \tag{6}$$

The amplitude spectrum for the yearly temperature effects before and after subtraction of the parabolic curve (6) can be seen in Figures 22 and 23. The comparison shows that the parabola can be considered as a model of trend because its subtraction deletes the harmonic with the maximal period, leaving all other harmonics unchanged. A shift in temperature parabola (Figure 21) in comparison with the parabola of CO_2 growth might be explained as a result of the lag between rising CO_2 concentrations and rising global temperatures.

5. CONCLUSIONS

A method was developed for finding a long-term tendency in the atmospheric concentration of CO_2. In particular, it was shown that on a long-time interval the polynomial approach is better than the exponential one, which means that the prediction force of a parabolic model is stronger than an exponential one. The constancy of the rate of growth of the statistically stable characteristic of the main direction of a series of concentrations of CO_2 for the Mauna Loa station during the last 28-years was also shown.

More particular results were: the amplitude of the seasonal oscillation on graphs of the monthly effect considerably increases from the South Pole to the North Pole; to obtain the effects of annual trends, a 12-month interval is a statistically stable interval; oscillation of the monthly effects for the northern and southern hemisphere has opposite phases.

The main problem of the possible existence of a statistically confident correlation between long-term trends in the observation series of concentrations of CO_2 and the main climatic variable temperature remains open.

ACKNOWLEDGEMENTS

The authors wish to express their recognition of D. Keeling, Scripps Institution of Oceanography, for his pioneer work at Mauna Loa. They would also like to thank Oak Ridge National Laboratory for providing a tape of the data used in this analysis.

The authors would also like to thank Professors Bo Döös and Ted Munn for their advice and support.

REFERENCES

Antonovsky, M.Ya. (1986) ***The modern assessment of carbon dioxide and other trace gas roles in climate variations.*** WMO/TD No.151, December. Env. Poll. Mon. & Res. Prog. No.45.

BMPD (BioMeDical Package), Statistical Software Manual, 1440 Supulveda Blvd., Palo Alto, Ca., USA.

Gemon, R.H., J.T. Peterson, W.D. Komyr (1986) ***Atmospheric CO_2 concentrations.*** NOAA, GIMS, Flask and Continuous Network. CDIAC, NDP 005, Oak Ridge Nat. Lab., Tenn., USA.

Huber, P.J. (1981) ***Robust Statistics,*** Wiley & Sons, NY.

Keeling, C.D. (1984) ***Atmospheric and oceanic measurement needed for establishing data base in the potential effect of CO_2 induced climatic changes of Alaska.*** Proceedings of University of Alaska.

Keeling, C.D. (1987) ***Hourly calibration atmospheric CO_2 concentration 1958-1986***. Mauna Loa Observatory. CDIAC NDP 043, Oak Ridge Nat. Lab., Tenn., USA.

Pearman, G., P. Hyson (1981) ***The annual variation of atmospheric CO_2 concentration observed in the northern hemisphere***. Journal of Geophysical Research. **86**: 40:9839-9843.

Tukey, J.W. (1977) ***Exploratory Data Analysis.*** Reading, M.A.: Addison-Wesley.

	1	2	3	4	5	6	7	8	9	10	11	12	
1858	$\varepsilon(1858,1)$	.	.	.	.	.	.	.	.	.	.	$\varepsilon(1958,12)$	$C_1(1958)$
1959	.	.	.	.	.	.	.	.	.	.	.	.	$C_1(1959)$
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	. .	.
1987	$\varepsilon(1987,1)$	.	.	.	.	.	.	.	.	.	.	$\varepsilon(1987,12)$	$C_1(1987)$
	$C_2(2)$	$C_2(2)$	.	.	.	.	.	.	.	.	.	$C_2(12)$	C_0

Figure 1: Decomposition of matrix of data on effects by method of median smoothing.

c_0 - typical (characteristic) value for a given station;
c_1 - year effect (variation);
c_2 - month effect (variation);
ε - random fluctuation associated with other factors.

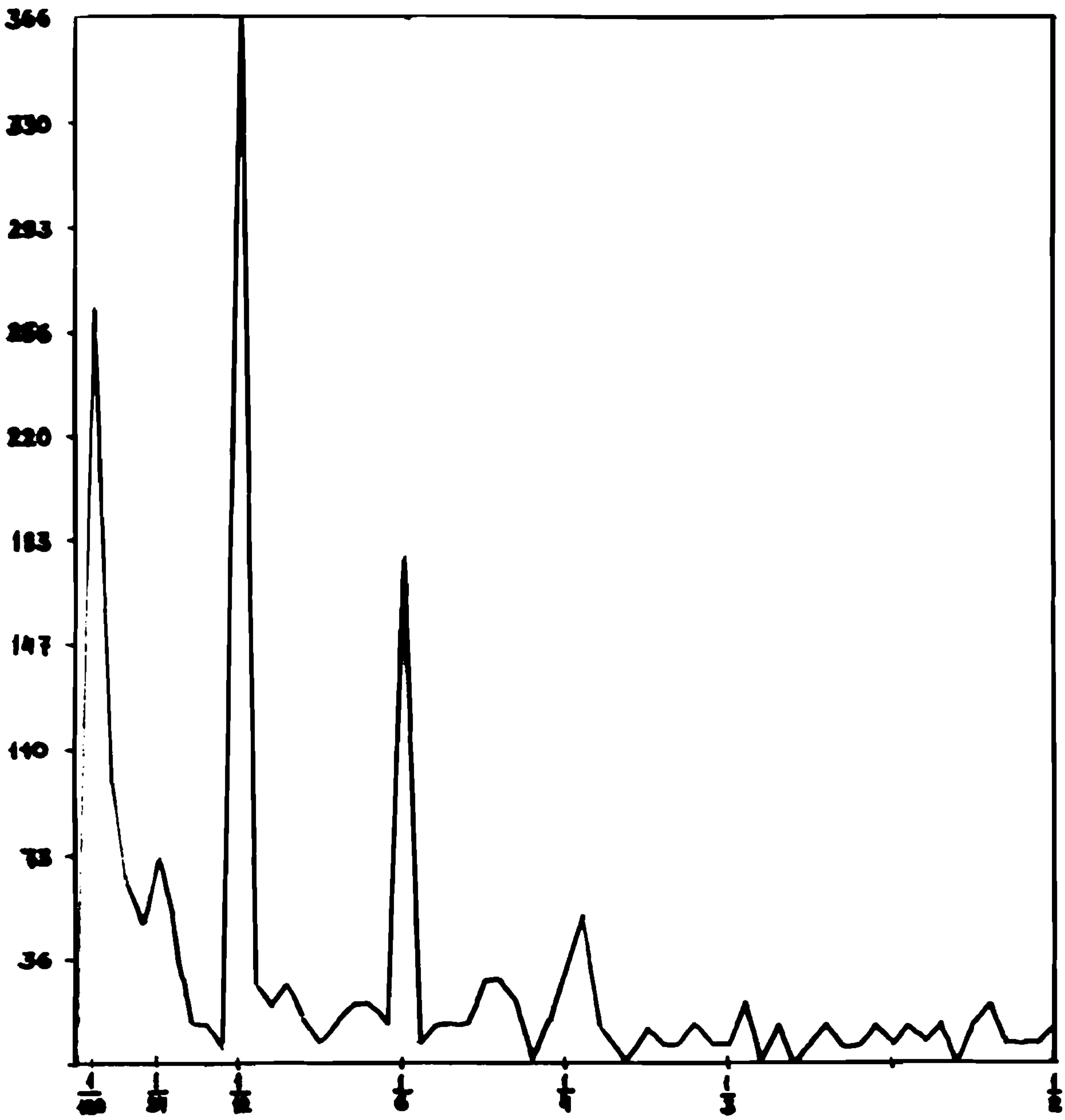

Figure 2: Amplitude spectrum of mean-monthly concentration of CO_2, Barrow station, 1973-1982.

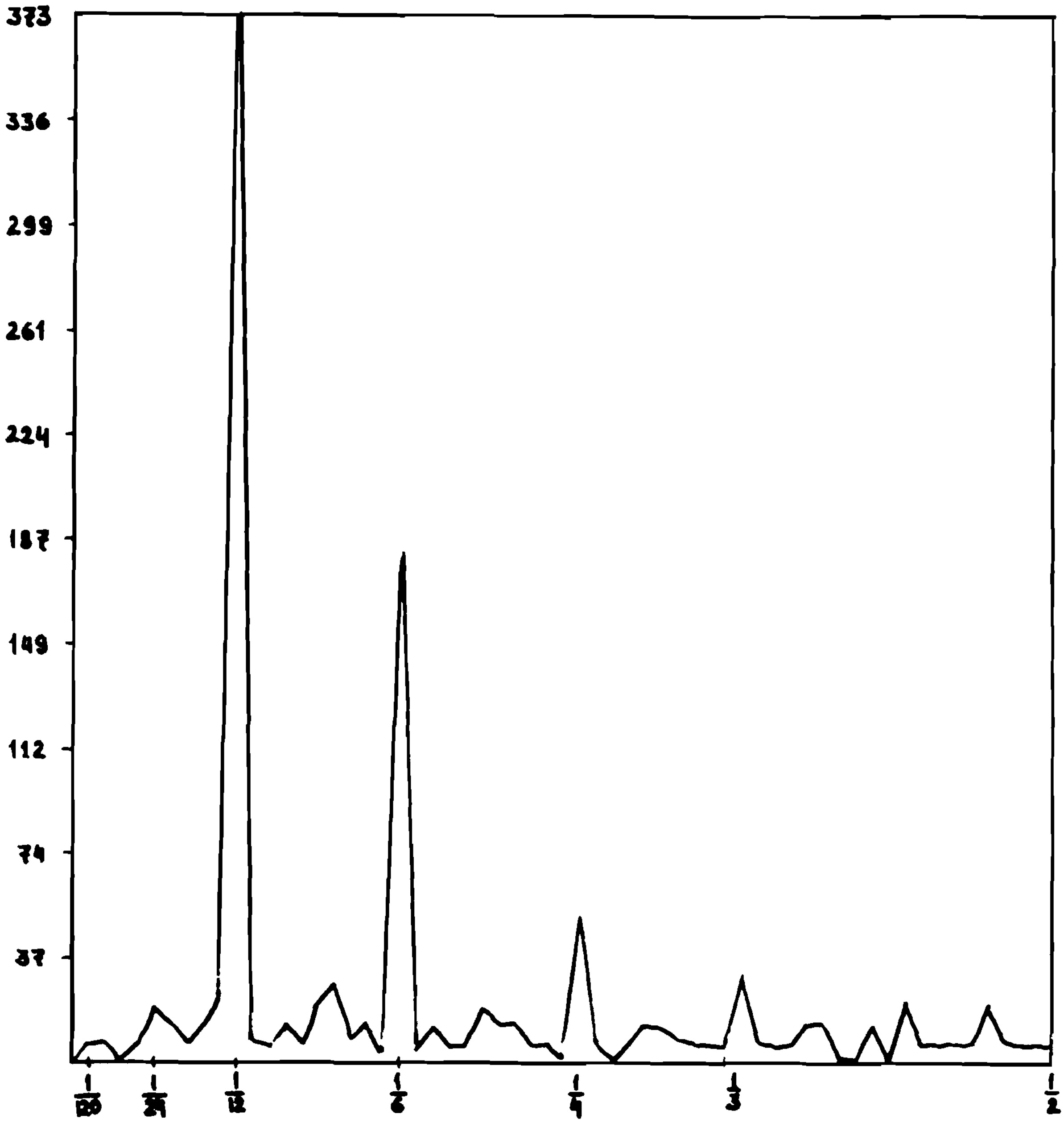

Figure 3: Amplitude spectrum of mean-monthly concentration of CO_2 with subtracted year effect, Barrow station, 1973-1982.

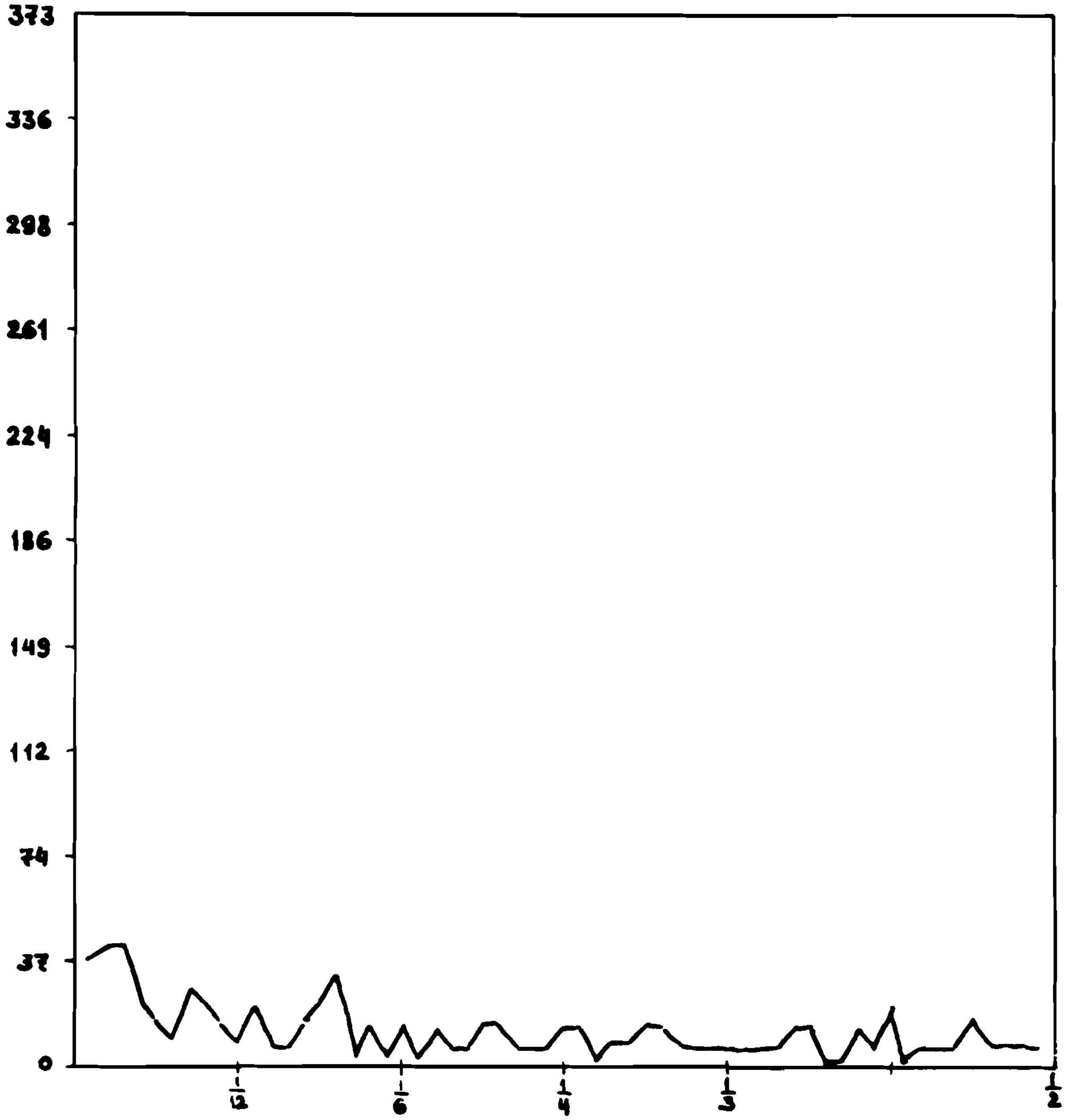

Figure 4: Amplitude spectrum of residuals of concentrations of CO_2 , Barrow station, 1973-1982.

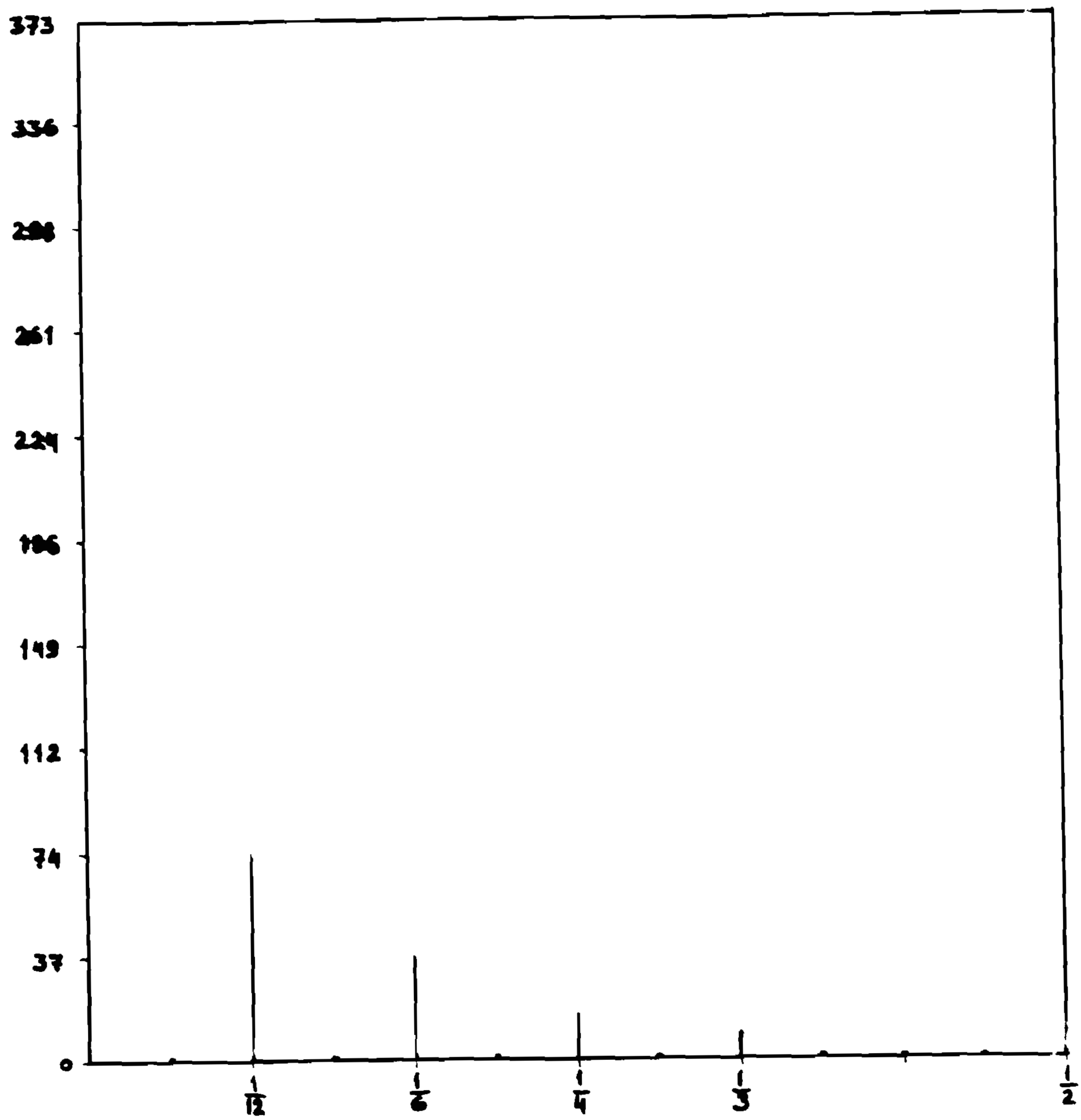

Figure 5: Amplitude spectrum of monthly effects of concentration, Barrow station, 1973-1982.

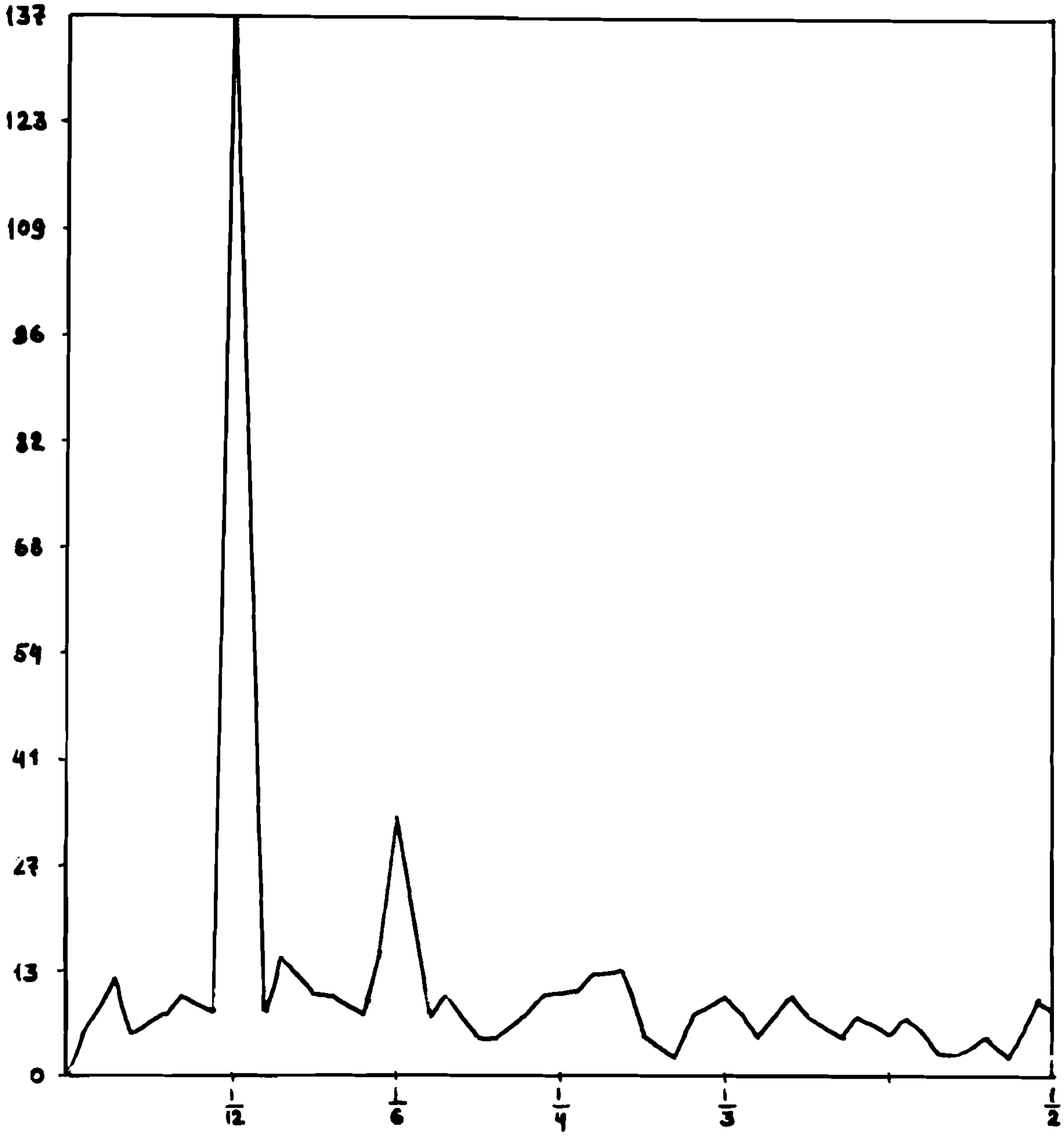

Figure 6: Amplitude spectrum of mean monthly effects of concentration, of CO_2 with subtracted year effect, NWR (Niwot Ridge).

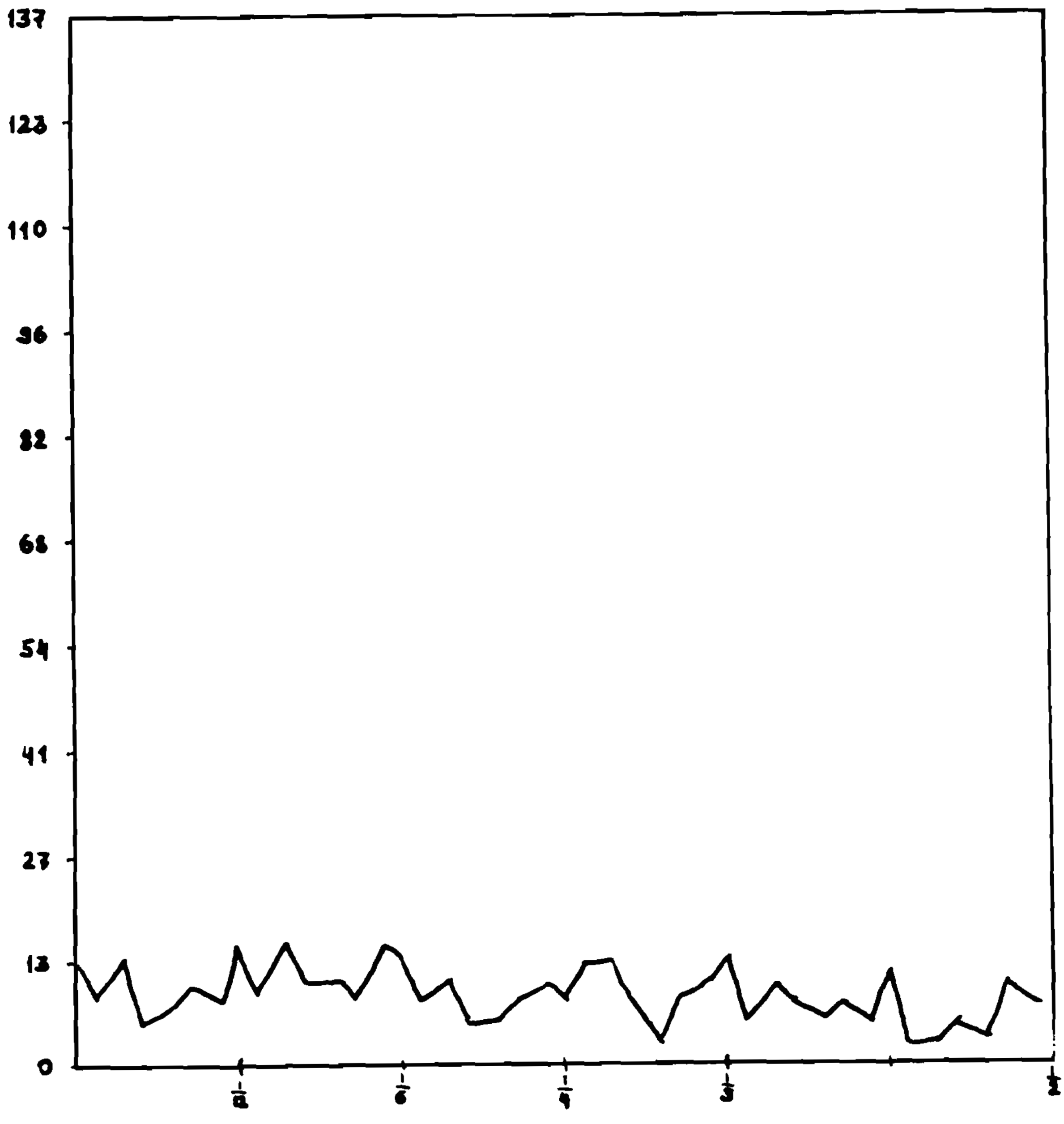

Figure 7: Amplitude spectrum of residuals of concentrations of CO_2, NWR (Niwot Ridge) station, 1976-1982.

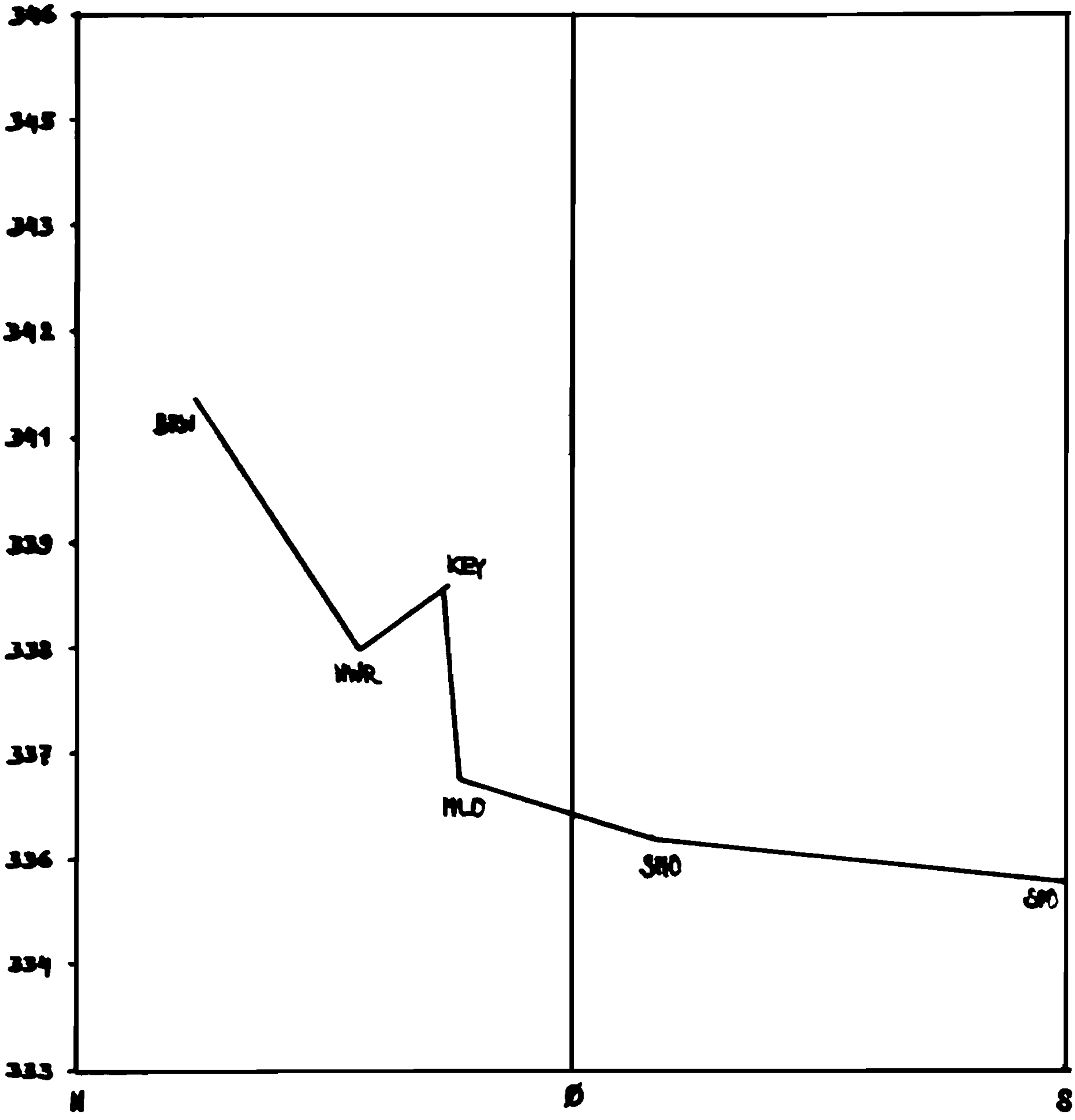

Figure 8: Dependence of characteristic value of concentration of CO_2 from altitude of the stations of Global Monitoring of climate change NOAA/GMCC, 1968-1982.

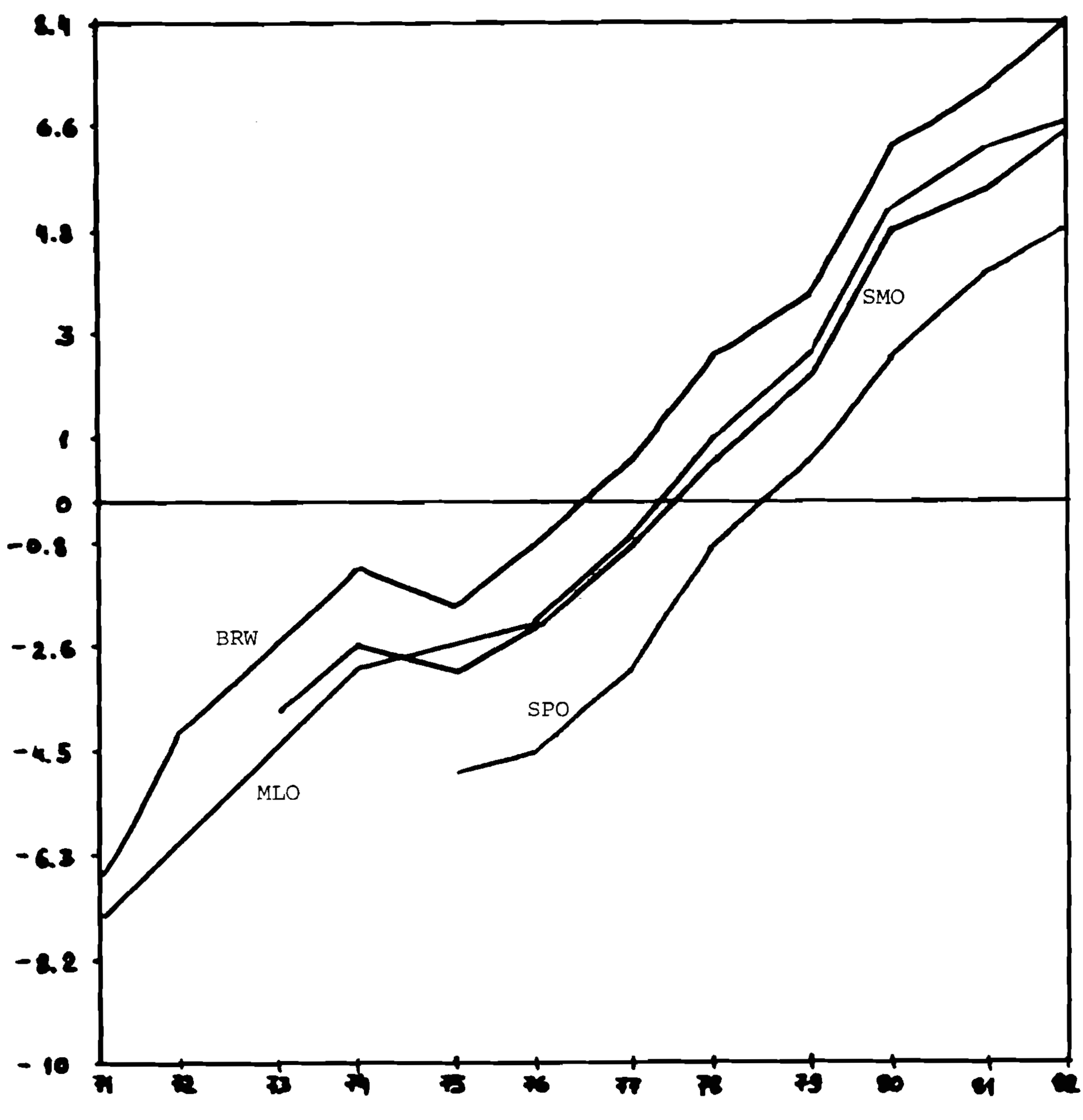

Figure 9: Increase of year effects for four stations of continuous monitoring: BRW, MLO, SMO, SPO stations, 1971-1982.

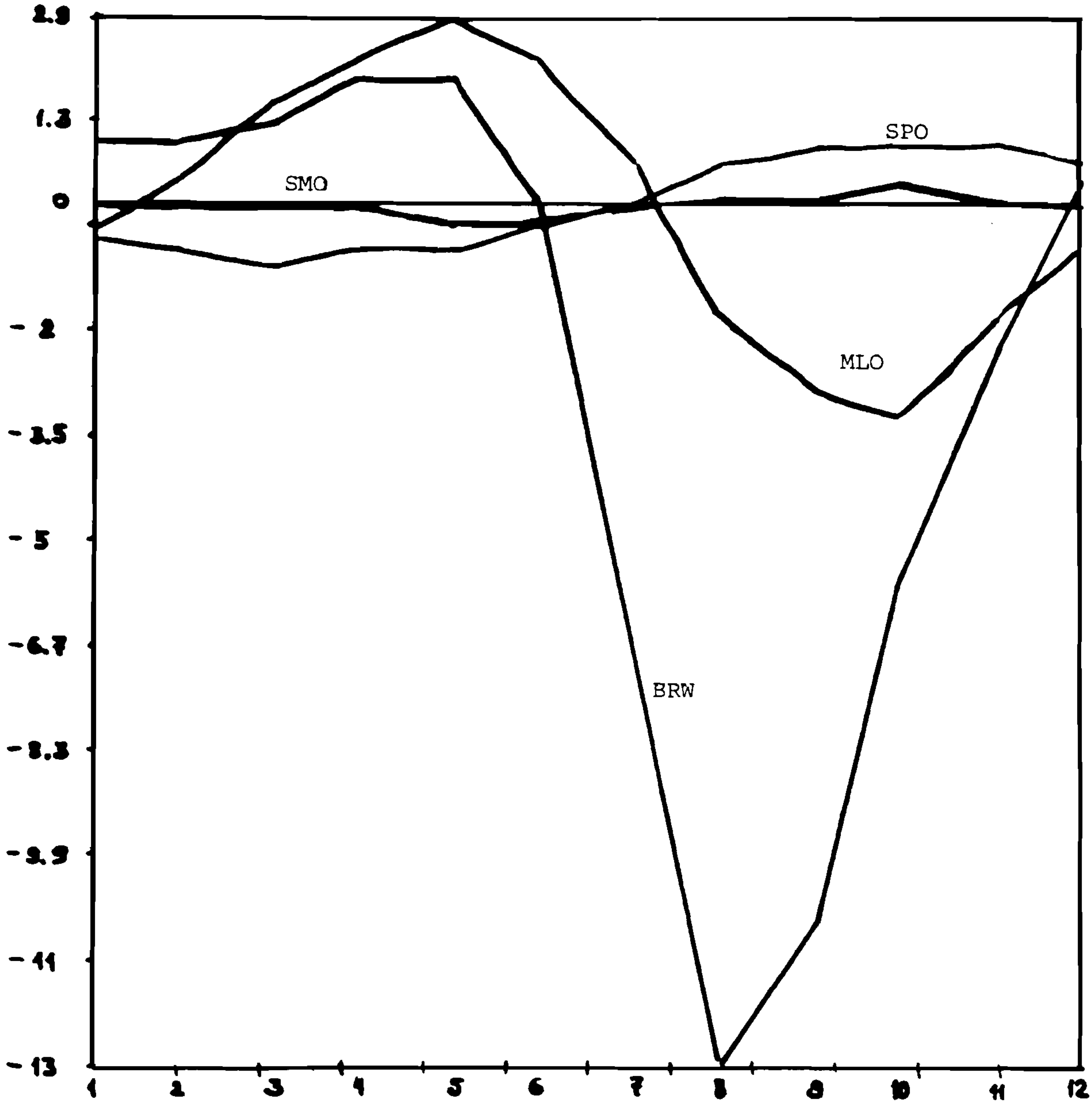

Figure 10: Behavior of mean-month effects of concentration of CO_2 for four stations of continuous monitoring: BRW, MLO, SMO, SPO, 1971-1982.

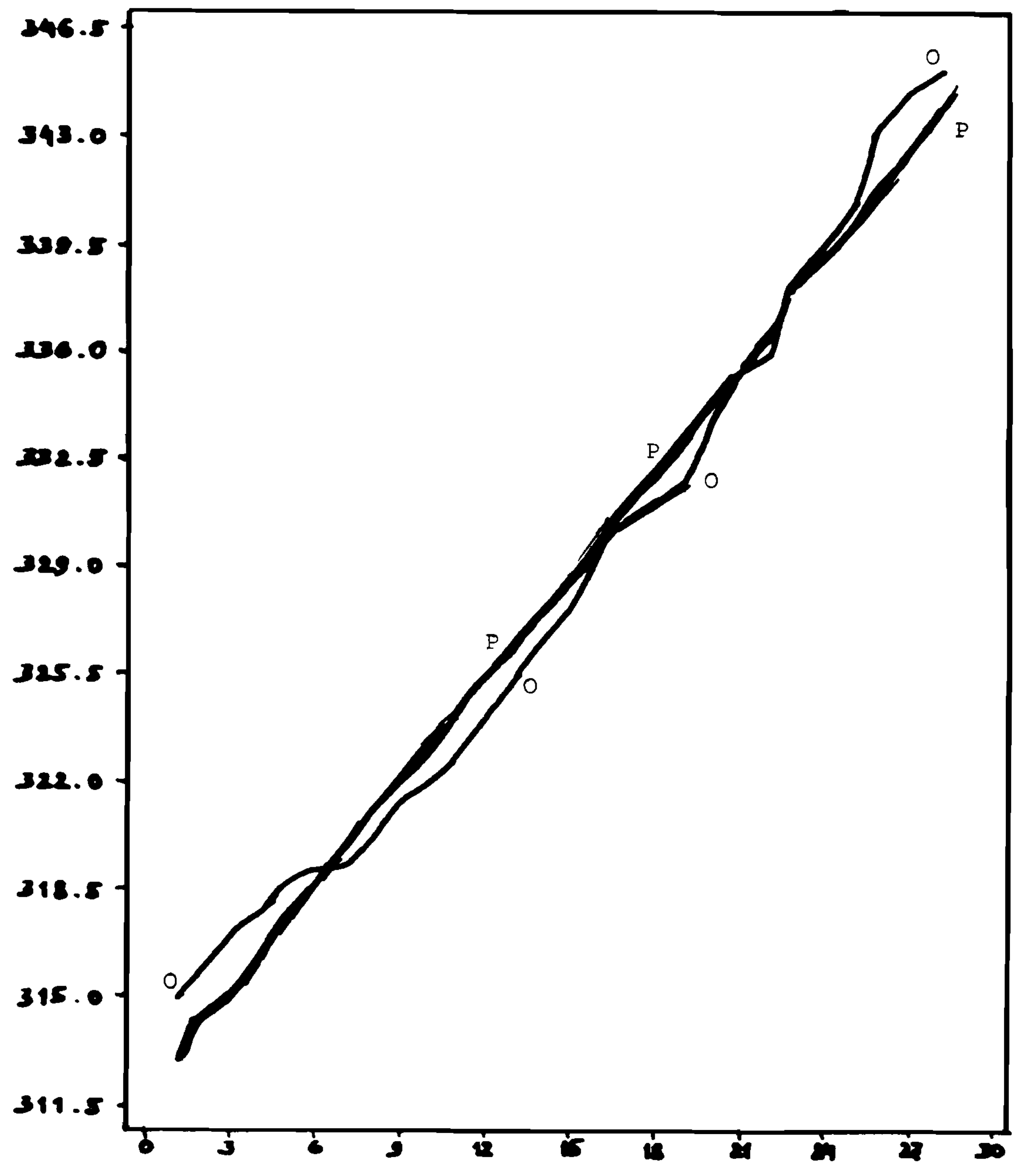

Figure 11: Approximation of a curve of year effects of concentration of CO_2 (o-observed) by exponential function (p - prediction) on interval of time 1958-1985.

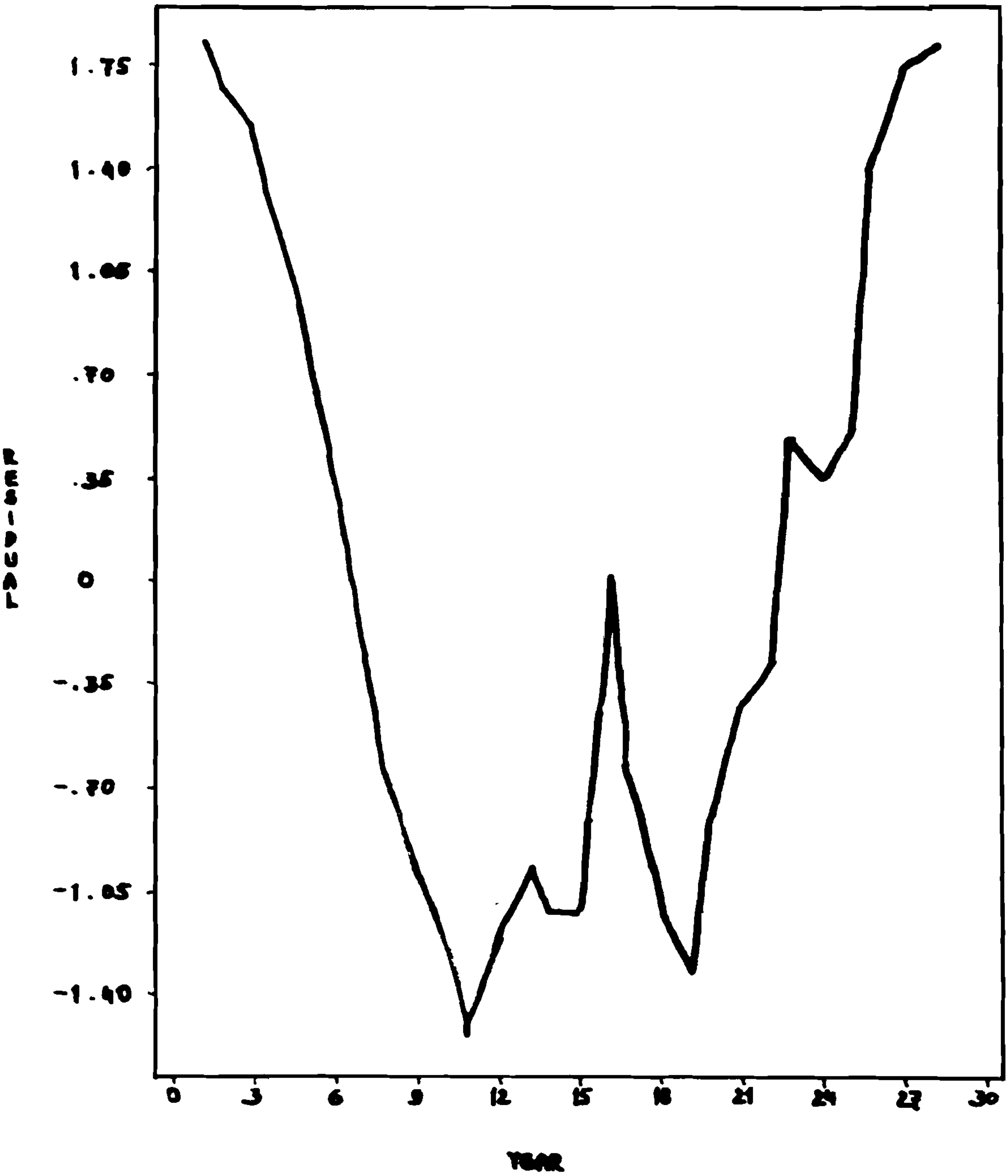

Figure 12: Behavior of residuals of exponential approximation of curve of year effects of concentrations of CO_2 during 1958-1985.

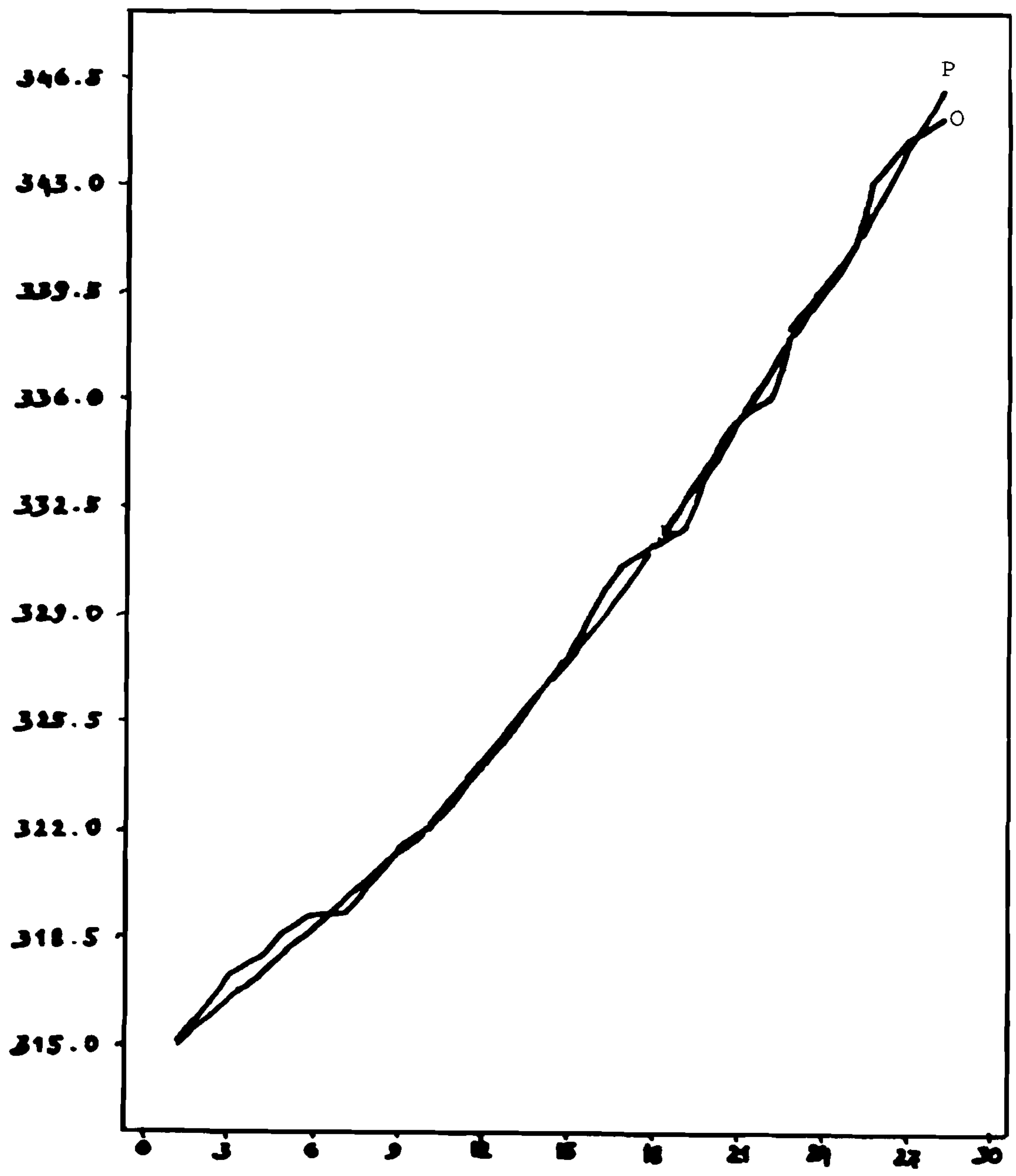

Figure 13: Approximation of curve of year effects of concentration of CO_2 (o-observed) by polynomial function (p-predicted) on time interval 1958-1985.

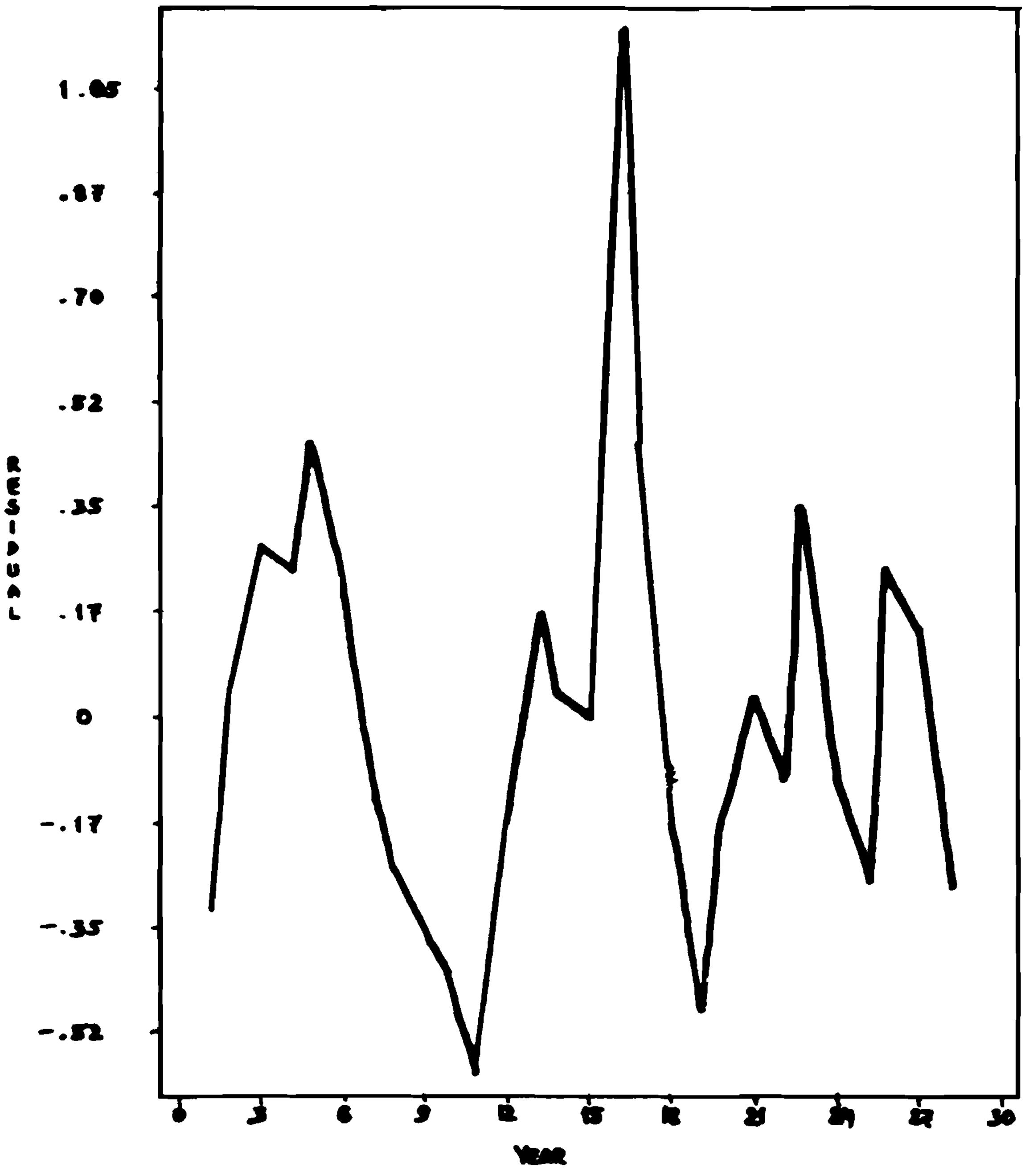

Figure 14: Behavior of residual of polynomial approximation of curve of year effects of concentrations of CO_2 during 1958-1985.

SHAPE OF THE MODEL	STANDARD DEVIATIONS
$c(t) = 312.057\, e^{0.0034t} + 0.031$	1.315
$c(t) = \begin{cases} 314.4\, e^{0.0024t} + 0.001, \text{ if } t \sim 58-69r. \\ 308.1\, e^{0.0040t} + 0.245, \text{ if } t \sim 70-85r. \end{cases}$	0.070 0.289
$c(t) = 0.02t^2 = 0.55t + 317.74$	0.141
$c(t) = \begin{cases} 0.01t^2 + 0.65t + 314.62, \text{ if } t \sim 58-69r. \\ 0.02t^2 + 0.45t + 315.99, \text{ if } t \sim 70-85r. \end{cases}$	0.072 0.152

Figure 15: **The models of the behavior of year effects of concentrations of CO_2 Mouna-Loa station, 1958-1985.**

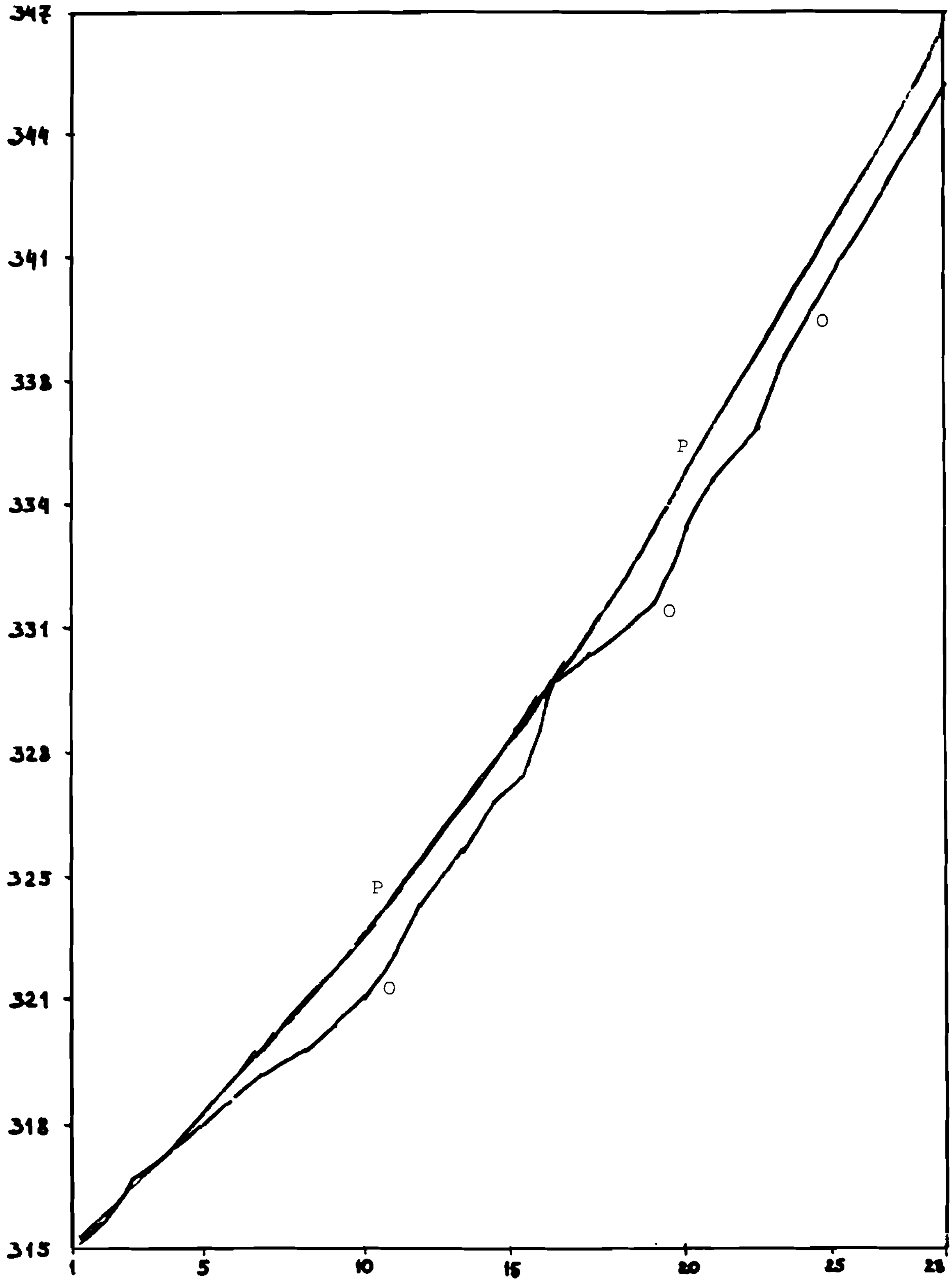

Figure 16: Prediction of the behavior of year effects of concentration of CO_2 by polynomial model, constructed by data for the first 5-years.

p1 = 314.293215 p2 = 0.002561 p3 = 0.000732

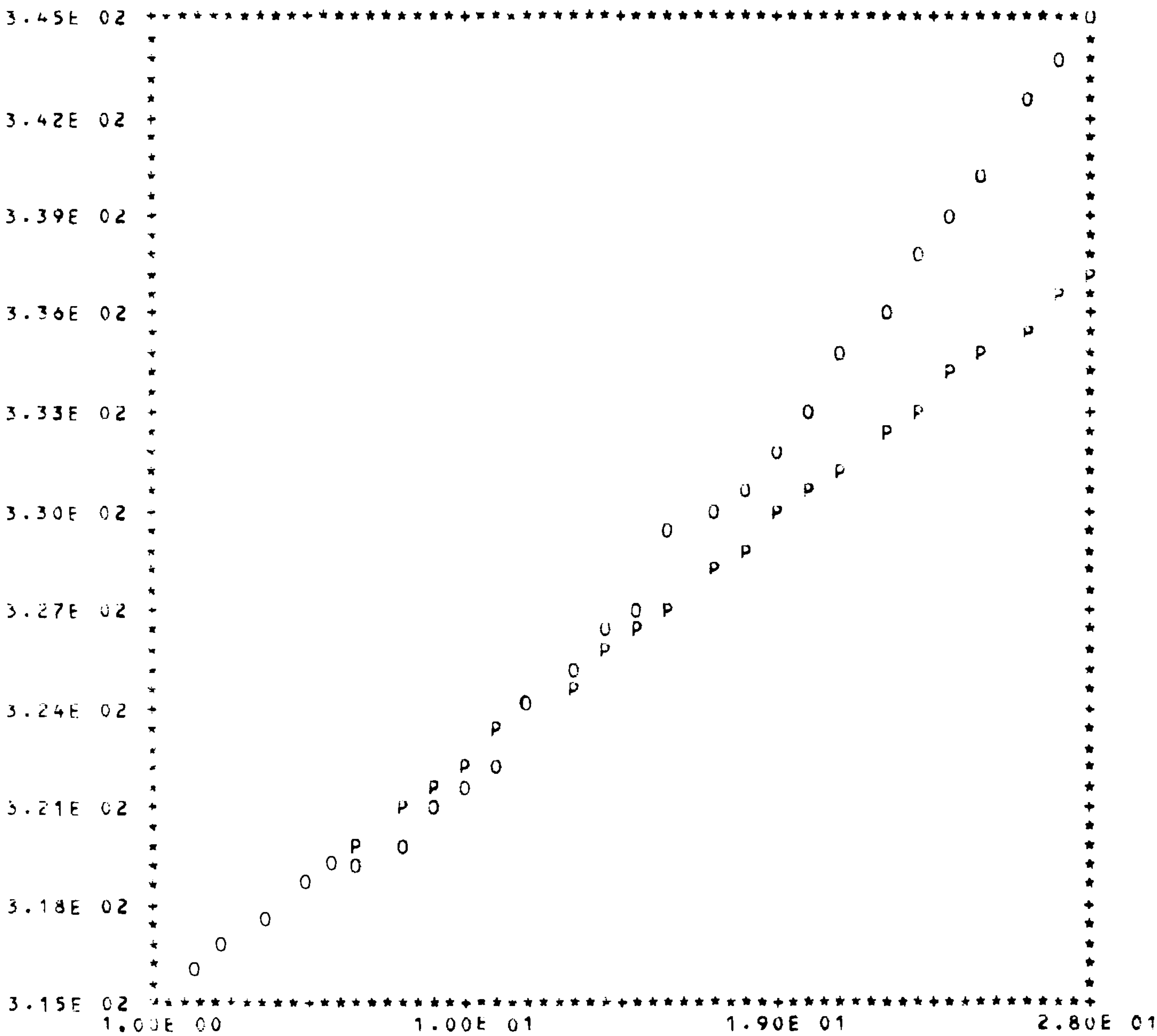

Figure 16': Graph of observed and predicted functions $\alpha_1 \exp \alpha_2 t + \alpha_3$ for 28 years. Prediction of the behavior of year effects of concentration of CO_2 with exponential model, constructed by data for first 5 years.

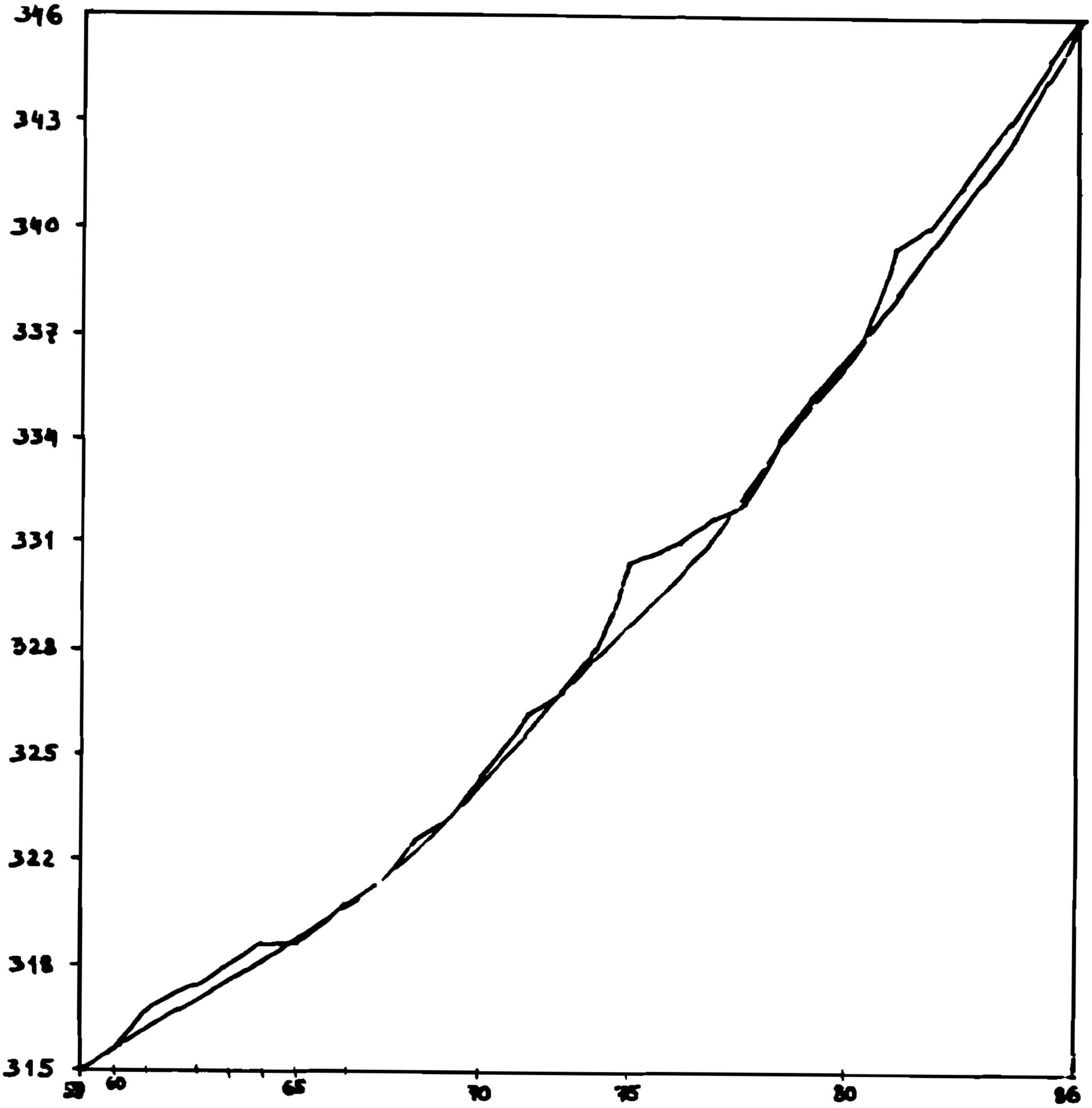

Figure 17: Prediction of the behavior of year effects of concentration of CO_2 polynomial model, constructed by data for first 15 years.

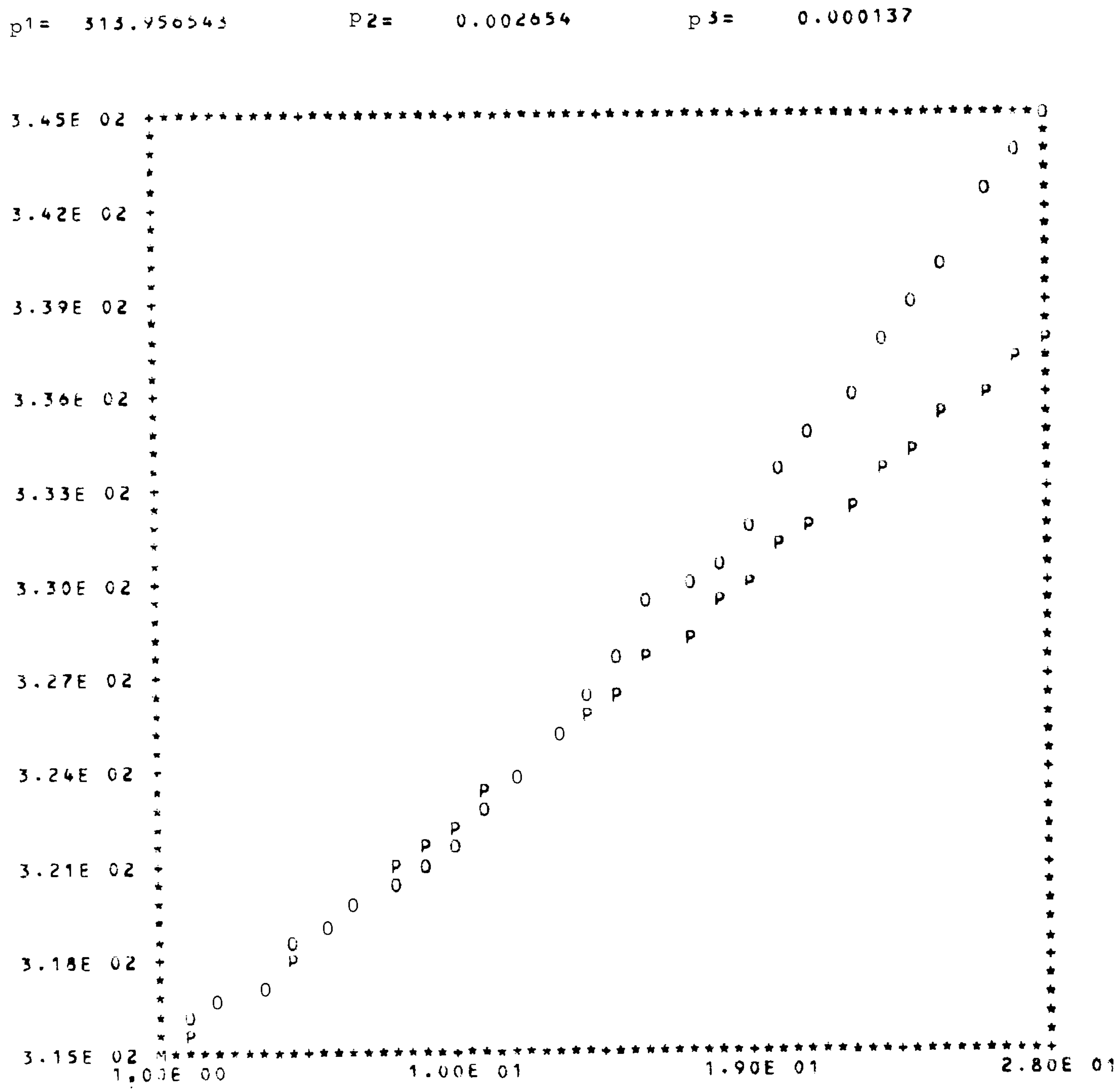

Figure 17': Graph of observed and predicted functions of $\alpha_1 \exp \alpha_2 t + \alpha_3$ for 28 years. Prediction of the behavior of year effects of concentration of CO_2 with exponential model, constructed by data for first 15 years.

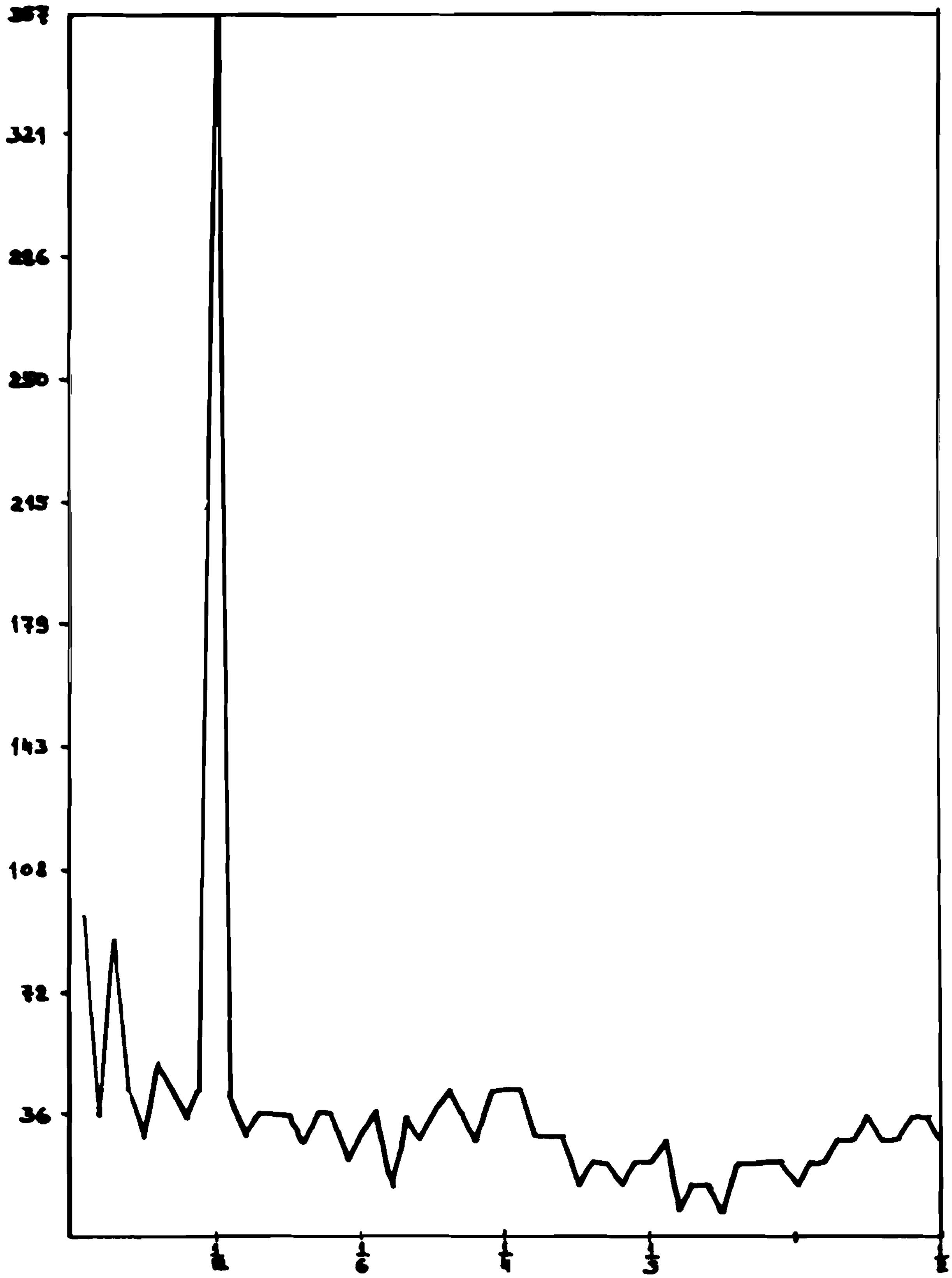

Figure 18: Amplitude spectrum of mean-monthly temperatures. Mauna-Loa stations, 1958-1985.

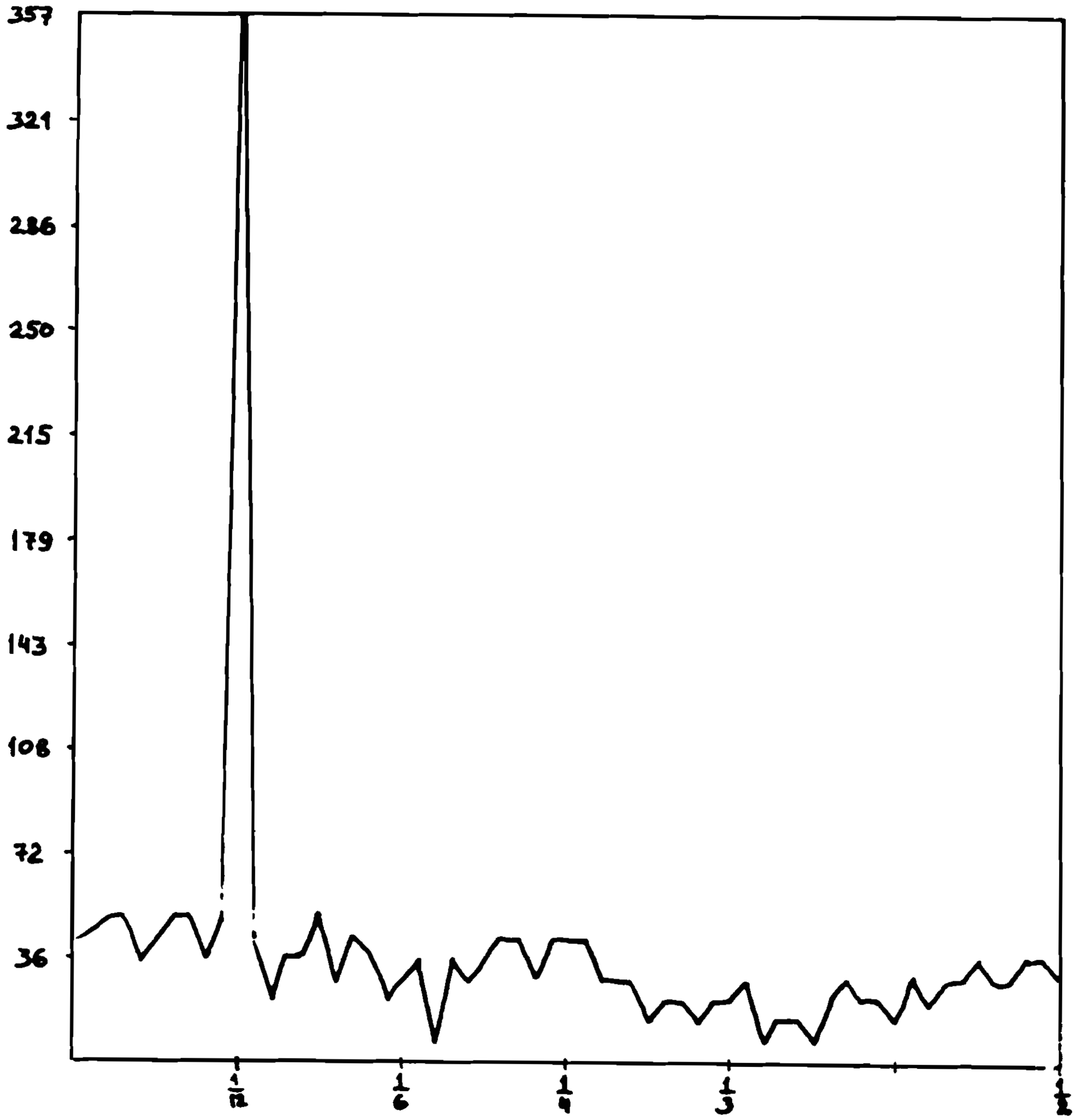

Figure 19: Amplitude spectrum of mean-monthly temperatures with substructed year effect, Mauna Loa, 1958-1985.

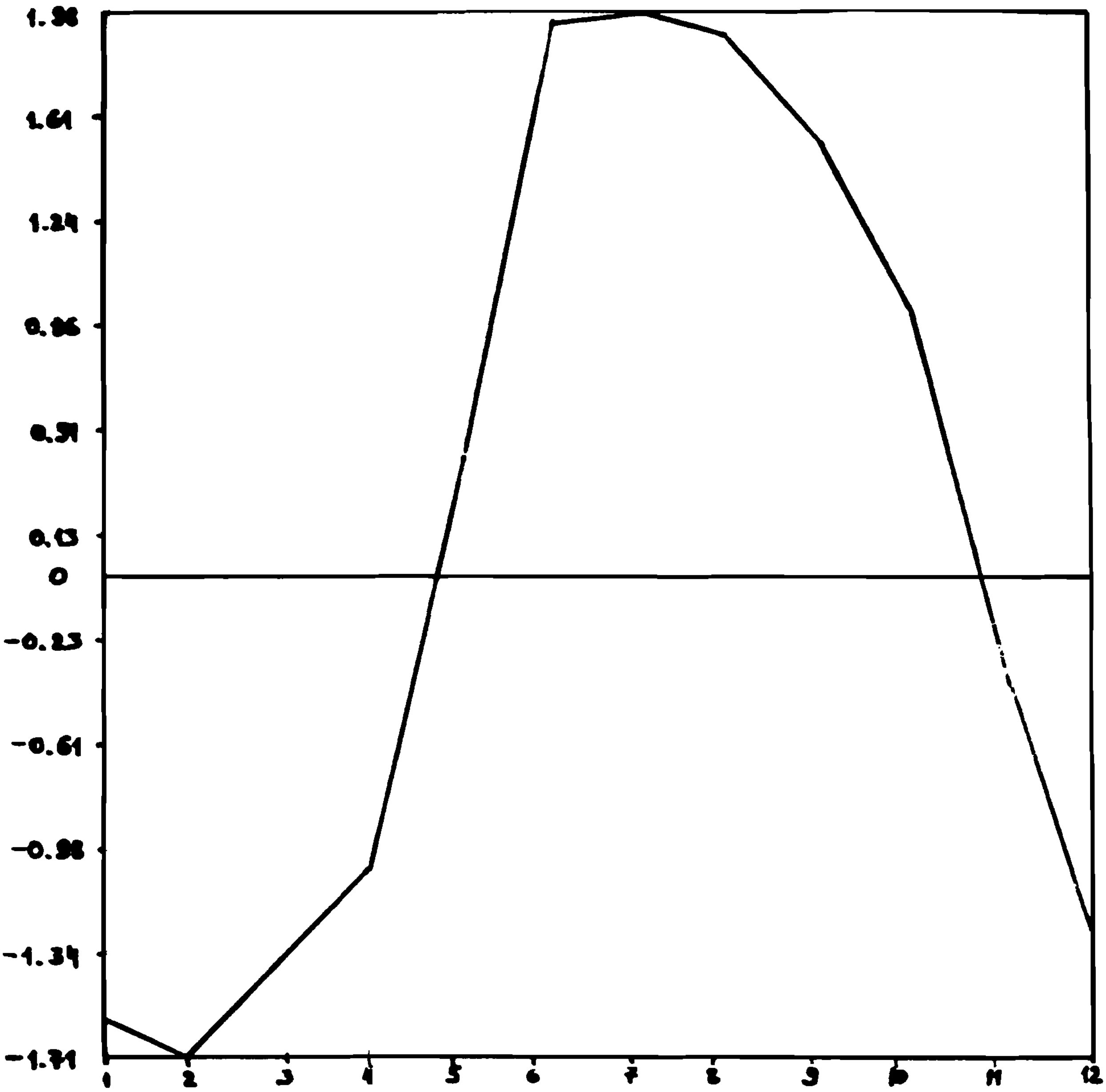

Figure 20: Behavior of mean-monthly effects of the temperature Mauna-Loa station, 1958-1985.

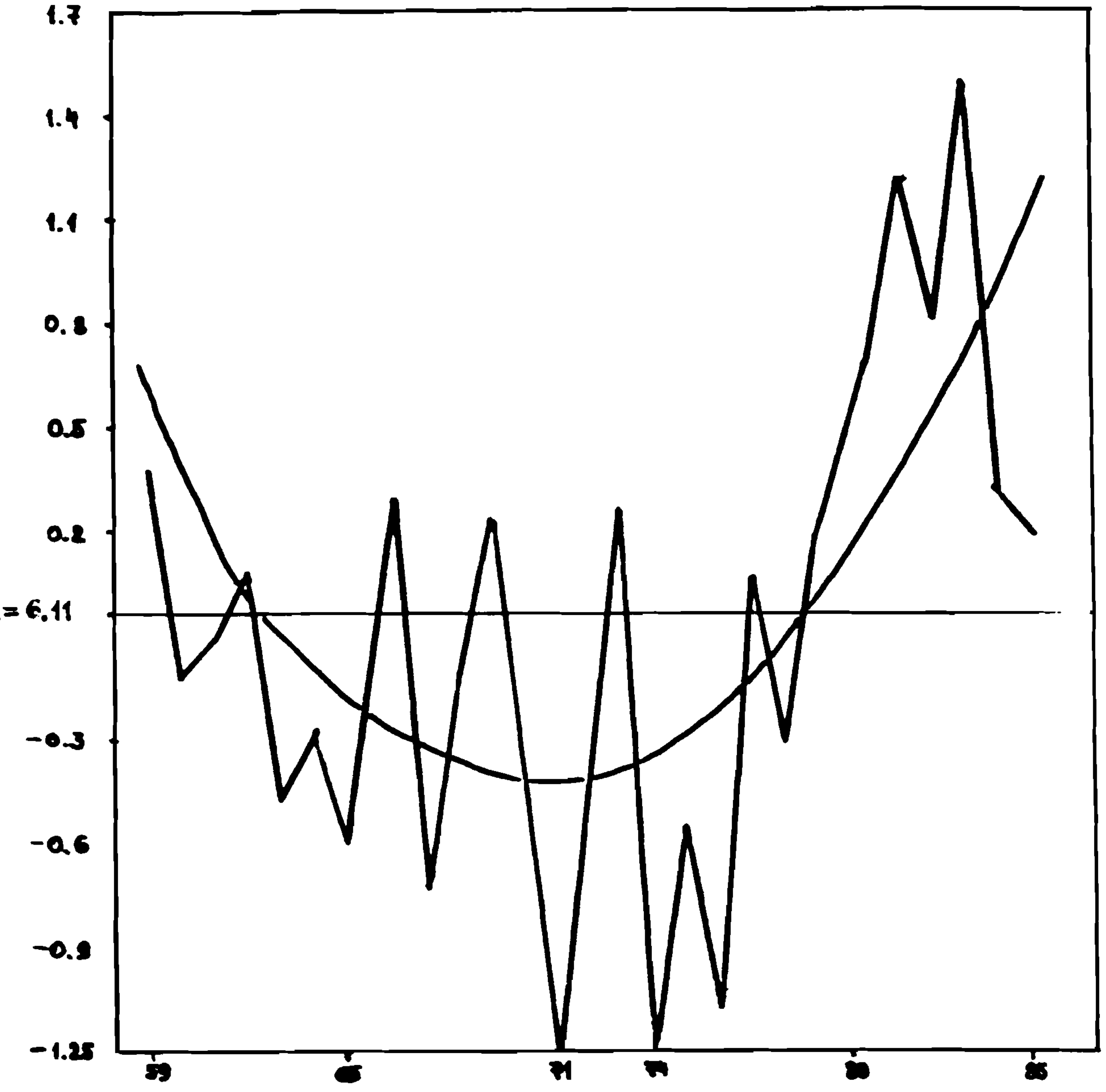

Figure 21: Behavior of year effects of mean-monthly temperature. Mauna Loa station, 1958-1985.

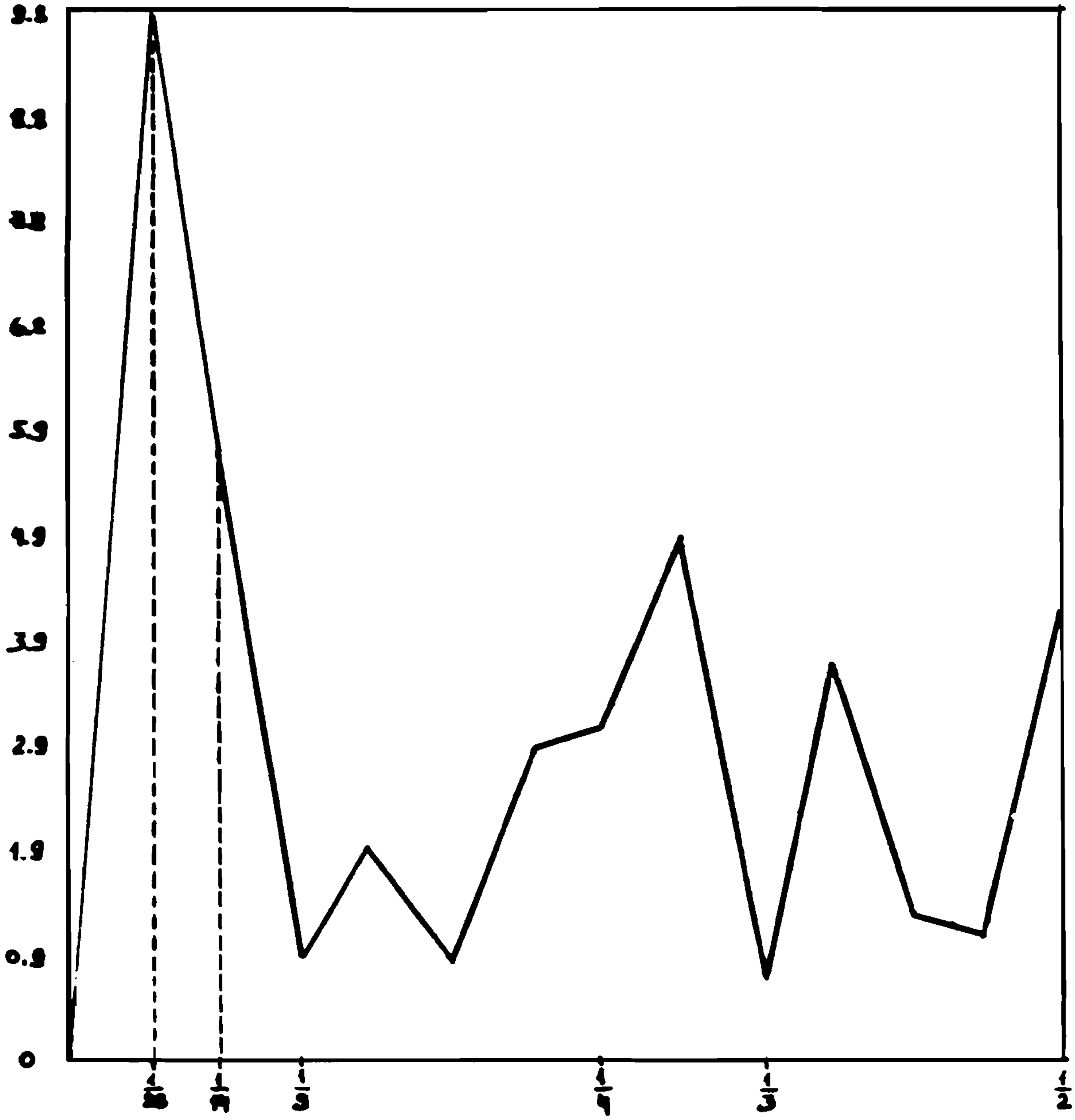

Figure 22: Amplitude spectrum of year effects of temperature, Mauna Loa station, 1958-1985.

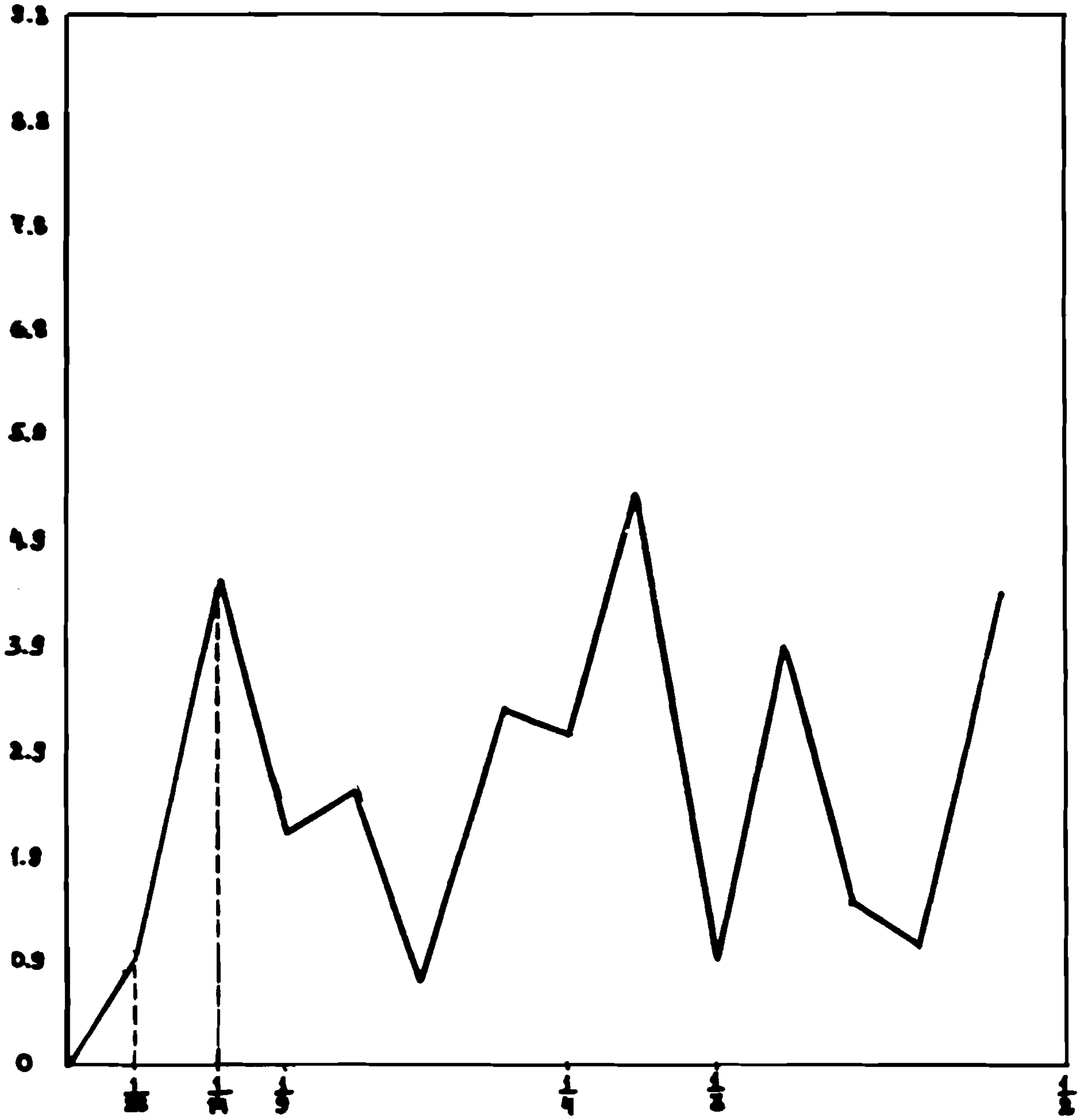

Figure 23: Amplitude spectrum of year effects of temperature with substructed parabolic trend, Mauna Loa station, 1958-1985.