U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

AD-A023 662

COMPARISON OF FORECASTS AND ACTUALITY

WISCONSIN UNIVERSITY

PREPARED FOR
ARMY RESEARCH OFFICE-DURHAM

MAY 1975

TECHNICAL REPORT NO. 402

May 1975

COMPARISON OF FORECASTS AND ACTUALITY

by

G. E. P. Box

and

G. C. Tiao
University of Wisconsin

DEPARTMENT OF STATISTICS

The University of Wisconsin
Madison, Wisconsin

TECHNICAL REPORT NO. 402

May 1975

COMPARISON OF FORECASTS AND ACTUALITY

by

G. E. P. Box
and
G. C. Tiao
University of Wisconsin

This research was supported in part by a grant from the American Petroleum Institute and in part by U. S. Army Research Office under Grant DA-ARO-D-31-124-72-G162.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR No. 402	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) COMPARISON OF FORECASTS AND ACTUALITY		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) G. E. P. Box and G. C. Tiao		8. CONTRACT OR GRANT NUMBER(s) Grant DA-ARO-D-31-124-72-G162
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics University of Wisconsin Madison, Wisconsin		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Durham, North Carolina		12. REPORT DATE May 1975
		13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this document is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

intervention analysis
time series
atmospheric ozone concentration
transfer function
forecast errors
linear model
level change in time series
parameter change in time series
autoregressive process

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In the analysis of time series data, it is frequently of interest to compare a set of forecasts made at some time point when a change in level is suspected to have taken place with actual observations. This paper discusses various statistical techniques for making such comparisons, and illustrates these techniques in terms of an actual example concerning the monthly averages of atmospheric ozone concentration of Azusa, California.

PRICES SUBJECT TO CHANGE

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

Comparison of Forecasts and Actuality

by

G. E. P. Box and G. C. Tiao
University of Wisconsin, Madison

1. Introduction

In recent work, Box and Jenkins (1968, 1970), methods for building stochastic and dynamic models were described and their application to forecasting was discussed. These methods were used by Tiao, Box and Hamming (1973), to build a stochastic model for the monthly average atmospheric ozone concentration at Azusa, California. The data consisted of 180 successive values from January 1956 to December 1970. The model was used to produce forecasts (from the origin December 1970) for the next 24 months. The 24 forecasts are compared in Figure 1 with what actually happened. This particular comparison is of interest because new automobile emissions standards were introduced at the end of 1970. These measures might have reduced ozone below levels expected if no new standards had been introduced. That such a reduction occurred is certainly plausible since most of the data actually observed fall below the forecasts made at the end of 1970. However it is of importance to make a more precise analysis. The object of this paper is to do this and to show how the methods we develop can be more generally useful and how they relate to earlier work, Box and Tiao (1965, 1973) on "Intervention Analysis."

Forecasts made in December 1970 of Ozone concentration at
Azusa California using the model

$$(1-B^{12})z_t = (1 + 0.15B)(1 - 0.91B^{12})a_t$$

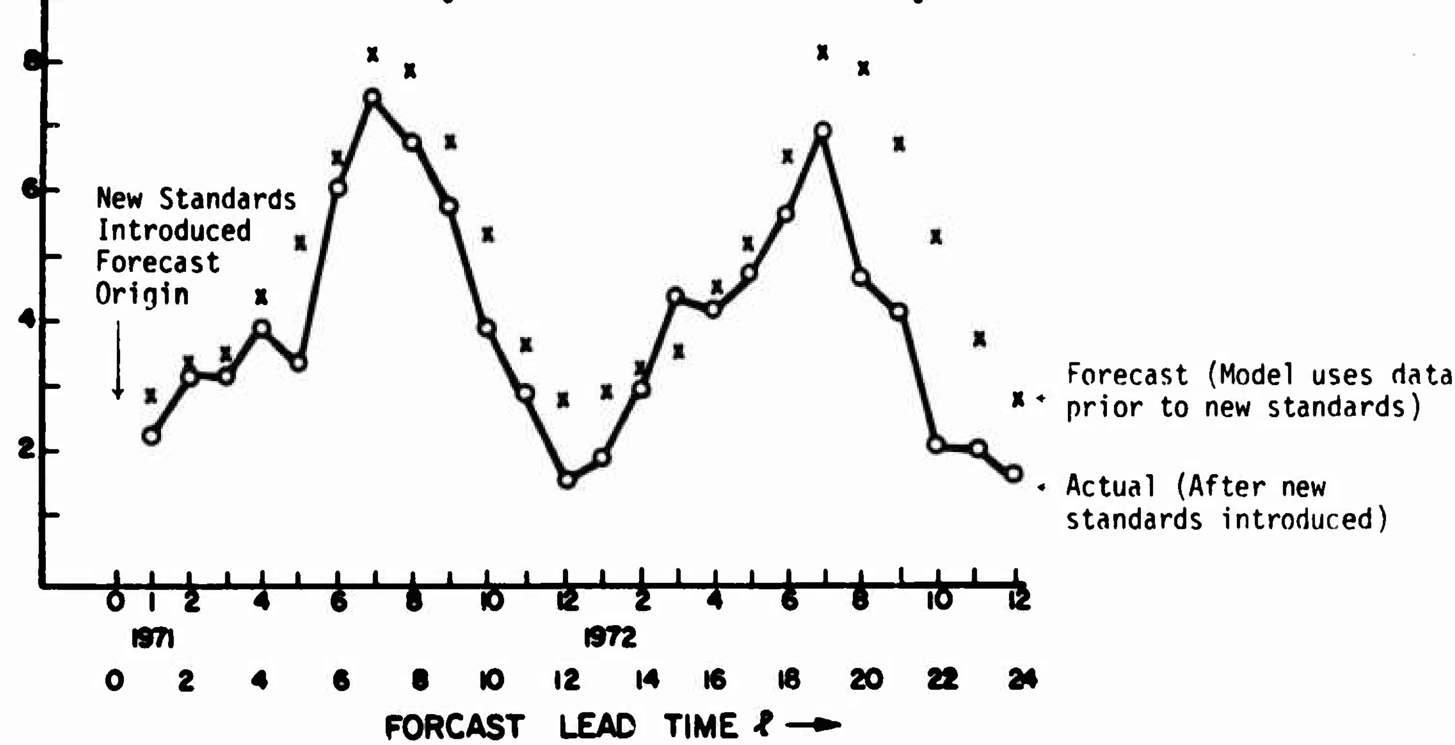

Figure 1

2. <u>A time series model for the ozone data.</u>

Following notation and methodology used and more fully explained in the references mentioned above we denote a time series (e.g., the monthly ozone data) by the sequence $\ldots z_{t-1}, z_t, z_{t+1} \ldots$. We also define a <u>white noise</u> series $\ldots a_{t-1}, a_t, a_{t+1} \ldots$ as a sequence of independently and normally distributed random <u>shocks</u> with mean zero and variance σ^2. Serially dependent values z_t of the time series are supposed to be generated from the random shocks a_t by a <u>linear filtering</u> operation

$$z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \cdots \tag{1}$$

where the ψ_j are fixed constants which typify the filter "memory". If we define the back shift operator B such that

$$Ba_t = a_{t-1} \text{ whence } B^k a_t = a_{t-k},$$

then (1) may be written

$$z_t = (1 + \psi_1 B + \psi_2 B^2 + \ldots)a_t \tag{2}$$

or

$$z_t = \psi(B)a_t \quad \text{where } \psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots$$

and $\psi(B)$ is called the transfer function of the filter.

Following ideas which originated with Yule (1927) and Yaglom (1955) the transfer function is often parsimoniously parameterized in terms of a difference equation. Iterative methods for building such a model, when applied to the ozone data, resulted in a representation of the form

$$(1-B^{12})z_t = (1 - \theta_1 B)(1 - \theta_2 B^{12})a_t \; . \tag{3}$$

Maximum likelihood estimates of the parameters were

$$\hat{\theta}_1 = -0.15 \quad (0.07), \qquad \hat{\theta}_2 = 0.91 \quad (0.04), \qquad \hat{\sigma}^2 = 1.00$$

where the numbers in the brackets are the corresponding estimated standard errors. This model was used to obtain the forecasts in Figure 1.

Thus, for the ozone data the fitted transfer function is

$$\psi(B) = \frac{(1 + 0.15B)(1 - 0.91B^{12})}{1 - B^{12}} . \tag{4}$$

We can alternatively define z_t as a linear function of previous observations z_{t-1}, z_{t-2}, ... plus a random shock so that

$$z_t = \pi_1 z_{t-1} + \pi_2 z_{t-2} + \ldots + a_t \tag{5}$$

or

$$(1 - \pi_1 B - \pi_2 B^2 + \ldots)\, z_t = a_t$$

or

$$\pi(B) z_t = a_t \text{ where } \pi(B) = 1 - \pi_1 B - \pi_2 B^2 - \ldots .$$

Now using (2) and (5)

$$\pi(B)\psi(B) = 1.$$

For the model in (3) the first 24 ψ and π weights are given in Table 1.

3. Comparison of forecast and actuality

The following theory applies exactly if the model is precisely known and approximately if the model parameters are estimated. Some discussion of the approximation is given in the appendix.

The minimum mean square error forecast made at origin T of $z_{T+\ell}$ is denoted by $\hat{z}_T(\ell)$, where $\ell = 1, 2, \ldots$, is called the lead time. It is readily shown that the lead ℓ forecast error $e_T(\ell) = z_{T+\ell} - \hat{z}_T(\ell)$ is given by

Table 1

ℓ	ψ_ℓ	$-\pi_\ell$	$e_T(\ell)$	$a_{T+\ell}$	$\hat{x}_{1\ell}$	$\hat{x}_{2\ell}$
1	.15	.15	-.35	-.35	1.0000	0
2	0	-.0225	.26	.31	.8500	.
3	.	.0034	-.09	-.14	.8725	.
4	.	-.0005	.20	.22	.8691	.
5	.	.0001	-2.54	-2.57	.8696	.
6	.	0	-.18	.21	.8696	1
7	.	.	-.73	-.76	.8696	.8500
8	.	.	-1.15	-1.04	.8696	.8725
9	.	.	-.95	-.79	.8696	.9691
10	.	.	-1.07	-.95	.8696	.8696
11	0	0	-.30	-.16	.8696	-.1304
12	.09	.09	-1.10	-1.08	.8696	.0196
13	.0135	.1230	-.96	-.77	.7796	-.0029
14	0	-.0185	-.04	.05	.7931	.0004
15	.	.0028	1.11	1.11	.7910	-.0001
16	.	-.0004	.40	.22	.7913	0
17	.	.0001	-1.34	-1.14	.7913	0
18	.	0	-.68	-.49	.7913	1.9100
19	.	.	-1.33	-1.19	.7913	1.6235
20	.	.	-3.15	-2.87	.7913	1.6665
21	.	.	-2.55	-2.03	.7913	1.6600
22	.	.	-2.87	-2.47	.7913	1.6610
23	.	.	-1.20	-.80	.7913	-.2491
24	0	0	-1.10	-.88	.7913	.0374

$$e_T(\ell) = a_{T+\ell} + \psi_1 a_{T+\ell-1} + \ldots + \psi_{\ell-1} a_{T+1} . \qquad (6)$$

From (5) it follows that (i) $a_{T+1} = z_{T+1} - \hat{z}_T(1)$ so that the random shocks $\ldots a_{T-1}, a_T, a_{T+1} \ldots$ are the one step forecast errors made at origins $\ldots$ T-2, T-1, T, ...; (ii) since these a_t's are independent, the variance of the lead ℓ forecast error is

$$\mathrm{Var}\{e_T(\ell)\} = (1 + \psi_1^2 + \ldots + \psi_{\ell-1}^2)\sigma^2 . \qquad (7)$$

Now consider the forecasts made for lead times $\ell = 1, 2, \ldots, m$ and the corresponding forecast errors $\underset{\sim}{e}' = [e_T(1), \ldots, e_T(m)]$. Also, let $\underset{\sim}{a}' = [a_{T+1}, \ldots, a_{T+m}]$. Then from (6) the transformation which converts the random shocks $\underset{\sim}{a}$ to the forecast errors $\underset{\sim}{e}$ is

$$\underset{\sim}{e} = \underset{\sim}{\psi}\, \underset{\sim}{a} \qquad (8)$$

where the $m \times m$ matrix $\underset{\sim}{\psi}$ and its inverse $\underset{\sim}{\psi}^{-1} = \underset{\sim}{\pi}$ are:

$$\underset{\sim}{\psi} = \begin{bmatrix} 1 & & & & & \\ \psi_1 & 1 & & & & \\ \cdot & \psi_1 & \cdot & & & \\ \cdot & \cdot & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & \cdot & 1 & \\ \psi_{m-1} & \psi_{m-2} & \cdot & \cdot & \psi_1 & 1 \end{bmatrix} \quad \underset{\sim}{\psi}^{-1} = \begin{bmatrix} 1 & & & & \\ -\pi_1 & 1 & & & \\ -\pi_2 & -\pi_1 & \cdot & & \\ \cdot & & & \cdot & \\ \cdot & & & & \\ \cdot & & & 1 & \\ -\pi_{m-1} & -\pi_{m-2} & \cdots & -\pi_1 & 1 \end{bmatrix}$$

That π is the inverse of ψ is readily confirmed by successively equating coefficients in the identity $\psi(B)\pi(B) = 1$.

Now the $m \times m$ covariance matrix for the vector $\mathbf{e}$ is $\mathbf{V} = E(\mathbf{e}\,\mathbf{e}') = \psi\psi'\sigma^2$. It follows that if the original model is appropriate during the period $T+1, \ldots, T+m$, then $Q = \mathbf{e}'\mathbf{V}^{-1}\mathbf{e}$ is distributed as χ^2 with m degrees of freedom, where $V^{-1} = \pi'\pi/\sigma^2$. If, on the other hand, the model differs from that previously experienced then we may expect the $e_T(\ell)$'s to be inflated.

Now rather than compute Q from the $e_T(\ell)$'s it is easier to employ the identity

$$Q = \mathbf{e}'\mathbf{V}^{-1}\mathbf{e} = \mathbf{e}'\pi'\pi\mathbf{e}/\sigma^2 = \mathbf{a}'\mathbf{a}/\sigma^2 , \tag{9}$$

whence

$$Q = \sigma^{-2}\sum_{\ell=1}^{m} a_{T+\ell}^2 \tag{10}$$

is the standardized sum of squares of the one step ahead forecast errors, $a_{T+1}, \ldots, a_{T+m}$. Thus, as we suggested in our joint paper with Hamming (1973), an overall test of the appropriateness of the model could be achieved by referring Q to a χ^2 table with m degrees of freedom. Further, this is equivalent to the appropriate test applied to all the lead ℓ forecast errors $e_T(\ell)$, $\ell = 1, \ldots, m$.

Since in practice σ^2 is estimated from n data values to which, say, p parameters have been fitted, a closer approximation would refer $\hat{Q}/m$, where $\hat{Q} = \hat{\sigma}^{-2} \sum_{\ell=1}^{m} a_{T+\ell}^2$, to an F table with m and n-p degrees of freedom. However, when n is large this refinement would make little difference to the result which is in any case approximate.

For the ozone data we find that $\hat{Q} = 36.01$ which is close to the 5% value of χ^2 with 24 degrees of freedom and suggests that the deviations from the model are real.

Components of χ^2

The test based on Q is an overall test having, like all such tests,

(i) the advantage that it is unnecessary to be specific about the nature of the feared discrepancy,

(ii) the disadvantage that the test lacks sensitivity (or power) when compared with a specified test which assumes that we have guessed correctly about what to be afraid of.

We now illustrate how, where appropriate, the Q statistic may be analyzed into components which correspond with specific alternatives.

4. Changed in model defined in terms of the z_t's

One way in which model changes may be defined is in terms of changes in the z_t's. Suppose that a change at time T has resulted in an additional component in the z_t's which at time $T+\ell$ is of the form

$$\beta_1 x_{1\ell} + \ldots + \beta_k x_{k\ell} \quad , \qquad k \leq m . \tag{11}$$

For example,

(i) if $k = 1$ with $x_{1\ell} = 1, \ell = 1, \ldots, m$, the model allows for a possible change in level of β_1 ;

(ii) if $k = 2$ with $x_{1\ell} = 1$, $x_{2\ell} = \ell$, $\ell = 1, \ldots, m$, the model allows for a possible change in level of β_1 and a change in slope of β_2;

(iii) alternatively, x_1, x_2, etc. could be genuine exogenous variables which have previously had no effect on the system.

In general, the errors $\underset{\sim}{e}$ may contain a deterministic component $\underset{\sim}{X}\underset{\sim}{\beta}$ where $\underset{\sim}{\beta}' = (\beta_1, \ldots, \beta_k)$ and $\underset{\sim}{X}$ is the $m \times k$ matrix $\{x_{j\ell}\}$, $(j = 1, \ldots, k;\ \ell = 1, \ldots, m)$. We may then write

$$\underset{\sim}{e} = \underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{\epsilon} \qquad (12)$$

where $\underset{\sim}{\epsilon}$ has mean zero and covariance matrix $\underset{\sim}{V}$.

Now after premultiplying (12) by $\underset{\sim}{\pi}$ we have

$$\underset{\sim}{\pi}\underset{\sim}{e} = \underset{\sim}{\pi}\underset{\sim}{X}\underset{\sim}{\beta} + \underset{\sim}{\pi}\underset{\sim}{e}$$

or

$$\underset{\sim}{a} = \dot{\underset{\sim}{X}}\underset{\sim}{\beta} + \underset{\sim}{\alpha} \qquad (13)$$

where $\dot{\underset{\sim}{X}} = \underset{\sim}{\pi}\underset{\sim}{X}$, and $\underset{\sim}{\alpha} = \underset{\sim}{\pi}\underset{\sim}{\epsilon}$ is normally distributed with $E(\underset{\sim}{\alpha}) = \underset{\sim}{0}$ and $E(\underset{\sim}{\alpha}\underset{\sim}{\alpha}') = \underset{\sim}{I}_m \sigma^2$. The least squares estimate of $\underset{\sim}{\beta}$ is $\hat{\underset{\sim}{\beta}} = (\dot{\underset{\sim}{X}}'\dot{\underset{\sim}{X}})^{-1}\dot{\underset{\sim}{X}}'\underset{\sim}{a}$ and the model yields the following analysis of variance table.

Source	Sum of Squares	D. F.
Added component	$\hat{\beta}'\dot{X}'\dot{X}\hat{\beta}$	k
Residual	$(a-\dot{X}\hat{\beta})'(a-\dot{X}\hat{\beta})$	m-k

Example

(i) A natural but somewhat naive hypothesis is that the intervention at the end of 1970 will simply change the level of ozone z_t and hence the level of the $e_T(\ell)$'s. To test this hypothesis, we consider the model

$$e_T(\ell) = \beta_1 + \epsilon_\ell \tag{14}$$

or, by setting $x_{1\ell} = 1$, for $\ell = 1, \ldots, 24$, equivalently $e_T(\ell) = \beta_1 x_{1\ell} + \epsilon_\ell$ which transforms to

$$a_{T+\ell} = \beta_1 \dot{x}_{1\ell} + \alpha_\ell \tag{15}$$

with

$$\dot{x}_{1\ell} = 1, \qquad \ell = 1$$

$$\dot{x}_{1\ell} = 1 - \sum_{j=1}^{\ell-1} \pi_j, \qquad \ell = 2, \ldots, 24 .$$

The values for $e_T(\ell)$, $a_{T+\ell}$ and $\dot{x}_{1\ell}$ are given in Table 1. Using these values we have

$$\hat{\beta}_1 = \Sigma \dot{x}_{1\ell} a_{T+\ell} / \Sigma \dot{x}_{1\ell}^2 = -0.9035 . \tag{16}$$

The corresponding analysis of variance table is

Source	Sum of Squares	D. F.	Mean Square
Level change	13.70	1	13.70
Residual	22.32	23	.97
$\hat{\sigma}^2$		166	1.00

from which it appears that the hypothesis of no change in level is in fact discredited by the data.

(ii) A slightly more sophisticated analysis might take account of the following facts:

(a) Ozone levels are highly seasonal and are at their highest in the "summer" months June-October. It is only during this period that the new emission standard would be expected to make much difference.

(b) The number of cars fitted with devices required by the new standards would be roughly twice as high in the second year as in the first.

To take account of these facts, let us define, in addition to $x_{1\ell}$ as given in (15), a column vector $\underset{\sim}{x}_2$ with 24 elements such that $x_{2\ell} = 1$, for $\ell = 6, 7, 8, 9, 10$; $x_{2\ell} = 2$, for $\ell = 18, 19, 20, 21, 22$; and $x_{2\ell} = 0$ elsewhere.

The model $e_T(\ell) = \beta_1 x_{1\ell} + \beta_2 x_{2\ell} + \epsilon_\ell$ now transforms to

$$a_{T+\ell} = \beta_1 \dot{x}_{1\ell} + \beta_2 \dot{x}_{2\ell} + \alpha_\ell \tag{17}$$

where $\dot{x}_{2\ell}$ are given in the last column of Table 1. The corresponding analysis of variance table is

Source		Sum of Squares		D. F.		Mean Square
Due to x_1 & x_2	x_2 alone	19.52	17.01	2	1	17.01
	Extra due to x_1		2.51		1	2.51
	Residual		16.50		22	0.75
	$\hat{\sigma}^2$				166	1.00

which suggests that the model $e_T(\ell) = \beta_2 x_{2\ell} + \epsilon_\ell$ is sufficient to explain the data.

5. <u>Changes in the parameters of a time series model</u>

As an alternative we may desire to entertain the possibility that at some time T one or more of the parameters of a time series model has changed. Let us assume that a time series model

$$\phi_o(B)\, z_t = \theta_o(B)\, a_t, \tag{18}$$

where $\phi_o(B) = 1 - \phi_{1o}B - \dots - \phi_{po}B^p$ and $\theta_o(B) = 1 - \theta_{1o}B - \dots - \theta_{qo}B^q$,

has been identified, fitted, and checked from data obtained prior to time T and is being used to make forecasts after time T. Let $a_{o,\,T+\ell}$ be the value of the shock at time $T+\ell$ computed from forecasts made for the model (18). Then $a_{o,\,T+1}, a_{o,\,T+2}, \dots, a_{o,\,T+m}$ may be computed from

$$a_{o,\,T+\ell} = \frac{\phi_o(B)}{\theta_o(B)}\, z_{T+\ell}, \quad \ell = 1, \dots, m. \tag{19}$$

Now suppose that at time T+1 the parameters may have changed from values $(\underset{\sim}{\phi}_o, \underset{\sim}{\theta}_o) = (\phi_{1o}, \ldots, \phi_{po}, \theta_{1o}, \ldots, \theta_{qo})$ to values $(\underset{\sim}{\phi}, \underset{\sim}{\theta}) = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q)$.

By linearly expanding $a_{T+\ell} = \frac{\phi(B)}{\theta(B)} z_{T+\ell}$ with respect to $(\underset{\sim}{\phi}, \underset{\sim}{\theta})$ at $(\underset{\sim}{\phi}_o, \underset{\sim}{\theta}_o)$ and rearranging we have approximately

$$a_{o,T+\ell} \doteq \sum_{i=1}^{p} (\phi_i - \phi_{io}) w_{i\ell} + \sum_{j=1}^{q} (\theta_j - \theta_{jo}) w_{(p+j)\ell} + a_{T+\ell} \tag{20}$$

where $$-w_{i\ell} = \left. \frac{\partial a_{T+\ell}}{\partial \phi_i} \right|_{(\underset{\sim}{\phi}_o, \underset{\sim}{\theta}_o)}, \qquad -w_{j\ell} = \left. \frac{\partial a_{T+\ell}}{\partial \theta_j} \right|_{(\underset{\sim}{\phi}_o, \underset{\sim}{\theta}_o)} .$$

This has the same form of a linear model and so allows us to proceed as before.

<u>Example</u>

The derivatives may be calculated from the changes that occur in the $a_{T+\ell}$'s when small changes are made in the parameters. Thus, suppose we wish to entertain the possibility that the discrepancies in the forecasts of Figure 1 are produced by changes in the parameters θ_1 and θ_2 at time T+1 (January, 1971) in the model (3), then we can calculate derivatives in the manner illustrated in Table 2 where $-\dot{x}_{3\ell} \doteq -w_{1\ell} = \left. \frac{\partial a_{T+\ell}}{\partial \theta_1} \right|_{\underset{\sim}{\theta}_o}$,

$-\dot{x}_{4\ell} \doteq -w_{2\ell} = \left. \frac{\partial a_{T+\ell}}{\partial \theta_2} \right|_{\underset{\sim}{\theta}_o}$, and $\underset{\sim}{\theta}_o = (\theta_{1o}, \theta_{2o})$.

Alternatively, they may be computed recursively as described in Box and Jenkins (1970, p. 235).

Table 2

ℓ	Value of parameters (θ_1, θ_2) (-.15, .91) $a_{T+\ell}$	(θ_1, θ_2) (-.14, .91) $a_{T+\ell}$	(θ_1, θ_2) (-.15, .92) $a_{T+\ell}$	$\dot{x}_{3\ell} = \frac{\text{row 1-row 2}}{.01}$	$\dot{x}_{4\ell} = \frac{\text{row 1-row 3}}{.01}$	Residuals after fitting by iterated least squares
1	-.3510	-.3600	-.3900	.9000	3.90	.195
2	.3092	.3069	.3525	.2300	-4.33	-.176
3	-.1403	-.1369	-.1001	-.3400	-4.02	.023
4	.2182	.2164	.3327	.1800	-11.45	.088
5	-2.5730	-2.5706	-2.7117	-.2400	13.87	-2.368
6	.2011	.1750	.2280	2.6100	-2.69	1.134
7	-.7554	-.7497	-.7907	-.5700	3.53	-1.419
8	-1.0340	-1.0423	-.9798	.8300	-5.42	-.529
9	-.7945	-.8037	-.7792	.9200	-1.53	-.627
10	-.9471	-.9538	-.8983	.6700	-4.88	-.244
11	-.1611	-.1697	-.1281	.8600	-3.30	-.381
12	-1.0779	-1.0783	-1.1282	.0400	5.03	-.578
13	-.7711	-.7819	-.8106	1.0800	3.95	-.358
14	.0491	.0429	.0921	.6200	-4.30	-.131
15	1.1072	1.1086	1.1428	-.1400	-3.56	1.222
16	.2134	.2243	.3209	-1.0900	-10.75	-.451
17	-1.1437	-1.1431	-1.2970	-.0600	15.33	-.096
18	-.4967	-.5082	-.4700	1.1500	-2.67	-.619
19	-1.1854	-1.1888	-1.2255	.3400	4.01	-.743
20	-2.8662	-2.8776	-2.8267	1.1400	-3.95	-2.481
21	-2.0342	-2.0613	-2.0281	2.7100	-.61	-.772
22	-2.4652	-2.4818	-2.4298	1.6600	-3.54	-1.759
23	-.8061	-.8285	-.7773	2.2400	-2.88	.218
24	-.8820	-.8869	-.9390	.4900	5.70	-.412

Fitting the equation

$$a_{o,T+\ell} \doteq (\theta_1-\theta_{1o})\dot{x}_{3\ell}+(\theta_2-\theta_{2o})\dot{x}_{4\ell} + a_{T+\ell} \qquad (21)$$

by least squares, first estimates of the adjustments to θ_1 and θ_2 are as follows:

$\theta_1 - \theta_{1o}$	$\theta_2 - \theta_{2o}$
-63 (.18)	-.07 (.03)

yielding adjusted values for (θ_1, θ_2) of $(-.78, .84)$ as compared with the previously estimated values of $(-.15, .91)$. The analysis of variance is as follows:

Source	Sum of Squares	D. F.	Mean Square
Change in parameters	14.75	2	7.37
Residual	21.24	22	0.97
$\hat{\sigma}^2$		166	1.00

The analysis suggests that if we exclude the previously considered possibilities in which a model change was associated with the level or slope of the $z_{T+\ell}$'s and allow instead only the possibility that the parameters have changed, then we find what appears to be a significant alteration particularly for θ_1.

This analysis must be treated with some caution because the $a_{T+\ell}$'s are in fact non-linear functions of the parameters θ_1 and θ_2 and the applicability of the linear analysis supposes that the linear approximation (21) is adequate over what has turned out to be a rather substantial adjustment for θ_1. However, iterated least squares yields final values for (θ_1, θ_2), the parameters after January 1970, of (-.58, .67) again indicating a substantial change.

The right hand column of Table 2 shows the residuals after the iterated least squares calculation. These strongly suggest that the trend has not been fully allowed for by the parameter adjustment, leading us to perform an analysis in which $\dot{x}_{2\ell}$, $\dot{x}_{3\ell}$ and $\dot{x}_{4\ell}$ are all included. The results are as follows:

Source	Sum of Squares	D. F.	Mean Square
Due to x_2	17.01	1	17.01
Extra due to possible change in parameters	4.04	2	2.02
Residual	14.91	21	.71
$\hat{\sigma}^2$		166	1.00

We are finally led to the conclusion, therefore, that the data may be fully accounted for on the simple and realistic hypothesis that the emissions from the new car engines resulted in the cumulative production of less ozone in the summer months, with no detectable change in the values of the stochastic parameters θ_1 and θ_2.

6. Relation with intervention analysis

In our earlier work, (1965, 1973), methods were described for estimating the effect of an "intervention" at a known point in a time series.

For example, in Los Angeles County in early 1960 events occurred (which may jointly be called the intervention) which might have been expected to reduce the level of ozone in Downtown Los Angeles. One expected effect was a change in level of the pollutant ozone. In our joint paper with Hamming (1973), a time series model was built for data from 1955 to 1970 which adequately allowed for a possible step change at the start of 1960. The size of the step and its standard error were estimated and it was possible to show that a substantial reduction in the pollutant almost certainly did occur at this time.

This situation differed from that discussed here in that the parameters of the time series model were estimated from substantial quantities of data available after, as well as before, the intervention.

However, these earlier methods could perfectly well be applied to examples like the present one and the results will be essentially similar if the period after the intervention is short. We believe the present procedure is worth separate consideration because of its simplicity and intuitive appeal. It is very natural to learn about a system by comparing a set of forecasts made at some point of possible change with actuality.

Appendix

Effect of Parameter Estimation Errors

In this paper we make the approximation that the model parameters may be treated as if they were exactly known, when in fact they must be estimated from a preliminary time series. (In the example considered preliminary time series consists of 180 data values at Azusa from January 1956 to December 1970.)

Some idea of the effect of the approximation in the special case of an auto-regressive process may be obtained as follows. Suppose $\hat{\phi}_1, \ldots, \hat{\phi}_p$ are estimated values based on n preliminary observations and $\phi_1, \phi_2, \ldots, \phi_p$ are the true values of an autoregressive process $a_t = \phi(B)z_t$, where $\phi(B) = (1 - \phi_1 B - \ldots - \phi_p B^p)$.

Now if the $\hat{a}_t$'s are estimated shocks such that

$$\hat{a}_t = (1 - \hat{\phi}_1 B - \ldots - \hat{\phi}_p B^p) z_t = \hat{\phi}(B) z_t$$

then

$$\hat{a}_t = a_t - \frac{\alpha_1 + \alpha_2 B + \ldots + \alpha_p B^{p-1}}{1 - \phi_1 B - \ldots - \phi_p B^p} a_{t-1} .$$

where $\alpha_i = \phi_i - \hat{\phi}_i, \quad i = 1, \ldots, p.$

It follows that, conditional on $\hat{\phi}_1, \ldots, \hat{\phi}_p$,

$$\frac{E(\hat{a}_t^2 \mid \hat{\phi}_1, \ldots, \hat{\phi}_p)}{\sigma^2} = 1 + \delta$$ where δ is the coefficient of B^0 in the expansion of

$$\frac{(\alpha_1 + \ldots + \alpha_p B^{p-1})(\alpha_1 + \ldots + \alpha_p B^{-(p-1)})}{\phi_p(B)\,\phi_p(B^{-1})} ,$$

that is

$$\delta = \underset{\sim}{\alpha}' \underset{\sim}{\Gamma}(\gamma)\underset{\sim}{\alpha}$$

where

$$\underset{\sim}{\alpha}' = (\alpha_1, \ldots, \alpha_p) \ ,$$

$$\underset{\sim}{\Gamma}(\gamma) = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdot & \cdot & \cdot & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \cdot & \cdot & \cdot \\ \cdot & \gamma_1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \gamma_0 & \gamma_1 \\ \gamma_{p-1} & \cdot & \cdot & \cdot & \gamma_1 & \gamma_0 \end{bmatrix}$$

and $\sigma^2\gamma_j = \text{cov}(z_t, z_{t-j})$.

Thus, $E(\hat{a}_t^2 | \hat{\phi}_1, \ldots, \hat{\phi}_p) = (1 + \underset{\sim}{\alpha}' \Gamma(\gamma)\underset{\sim}{\alpha})\,\sigma^2$.

Using the results in Box and Jenkins (1970, p. 241), we find that to order n^{-1},

$$\frac{E(\hat{a}_t)^2}{\sigma^2} = 1 + p/n \ .$$

Thus, there will be an inflation factor in the value of χ^2 produced by sampling errors in the ϕ's but this will be small if n is large compared with p.

References

Box, G. E. P. and Jenkins, G. M (1968) "Some recent advances in forecasting and control, I". Applied Statistics, 17, 91.

Box, G. E. P. and Jenkins, G. M. (1970). Time Series Analysis, Forecasting and Control, San Francisco: Holden-Day.

Box, G. E. P. and Tiao, G. C. (1965). "A change in level of a non-stationary time series." Biometrika, 52, 181.

Box, G. E. P. and Tiao, G. C. (1973). "Intervention analysis with applications to economic and environmental problems". Tech. Report #335, Department of Statistics, University of Wisconsin, Madison (to appear in J. Amer. Statist. Assoc.)

Tiao, G. C., Box, G. E. P. and Hamming, W. J. (1973) "Analysis of Los Angeles photochemical smog data: a statistical overview" Tech. Report #331, Department of Statistics, University of Wisconsin, Madison (to appear in J. Air Pollution Control Assoc.)

Yaglom, A. M (1955). "The correlation theory of processes whose n th difference constitute a stationary process". Matem. Sb., 37 (79), 141.

Yule, G. U. (1927) "On a method of investigating periodicities in disturbed series, with special reference to Wölfer's sunspot numbers", Phil. Trans., A 226, 267.