George Box's Contributions to Time Series Analysis and Forecasting

Greta M. Ljung, Ph.D.

Lexington, MA

greta.ljung@verizon.net

October 19, 2019

Significant Contributions to Many Areas of Statistics

Key Areas:

- Design of experiments and response surface methodolgy
- Distribution theory, transformations, and non-linear estimation
- Time series analysis and forecasting
- Statistical inference, Bayesian methods, and robustness
- Quality and productivity improvement

Publications:

- 10 books; Over 200 papers

Recommended reading:

- "The Collected Works of George E.P. Box, Volumes I and II", George C. Tiao, Editor in Chief, (1884), Wadsworth
- "Box on Quality and Discovery" edited by Tiao *et al* (2000), Wiley

Collaboration with Gwilym Jenkins on Time Series Analysis

- Met Jenkins at Princeton in the fall of 1959
- Worked together during the 1960's and into the 1970's
- Early collaboration involved a problem in automatic process control. The yield of a chemical tended to fluctuate and the problem was to reduce the fluctuations by automatically adjusting temperature.
- Developing an adjustment scheme involved forecasting future deviations from target.

Paper on "Some Statistical Aspects of Adaptive Optimization and Control" published in JRSS B in 1962.

Included in *Breatkthrougs in Statistics, Volume II* edited by Samuel Kotz and Norman L.Johnson

Box and Jenkins (1970): Time Series Analysis and Forecasting

- One chapter devoted to the control problem.
- Two chapters on transfer function-noise models describing dynamic relationships between two or more time series.
- The rest of the book was devoted to modeling and forecasting of univariate time series using ARMA and ARIMA models.

Subsequent editions:

- A revised version appeared in 1976.
- Greg Reinsel was added as co-author in 1994. New chapter on intervention analysis, outlier detection, and missing values.
- The 2008 edition included a new chapter on multivariate methods and a new chapter on special topics such as volatility modeling, non-linear models, and long memory models.
- The fifth edition in 2016 updated earlier material and added new exercises, R code, new references, etc.

Box and Jenkins (1970)

Key features of Box and Jenkins (1970)

- The book provided a practical and unified approach to model buidling based on an iterative cycle of
  - Model specification
  - Parameter estimation
  - Diagnostic checking
- The book provided a model-based approach to forecasting
- Unlike earlier literature, the book covered stationary and non-stationary time series including seasonal time series

Emphasis on simplicity and parsimony:

"*Our goal will be to derive models possessing maximum simplicity and the minimum number of parameters consonant with representational adequacy*"

What Was the Initial Reaction to the Book?

George in interview with Daniel Pena (2001): "As I recall, what reaction there was, tended to be negative."

Some grumbling in the beginning:

- Mixed review by Kendall (1971)
- Negative forecasting results reported by Chatfield and Prothero (1973) in JRSS-A
- O.D. Anderson (1977): "Is Box-Jenkins a Waste of Time?"

Some comments by O.D.: "The methodology is of doubtful practical value"; "It relies too heavily on scarse expertise and is oversold"; "It will be superseded".

The Book Has Had a Profound and Lasting Impact

Close to 50,000 citations in Google Scholar; More than 13,000 since 2014

The book has had a big impact in economics and econometrics, in particular. Some reasons are:

- The models fit many macroeconomic time series
- Box and Jenkins showed how to develop and use these models
- The models have good forecasting performance. For example, the models were shown to perform well against large econometric models with hundreds and sometimes thousand of equations [e.g. Nelson (1972), Cooper (1972), Naylor, *et al* (1972), and Newbold and Granger (1974)]
- The models allow for stochastic as well as deterministic trends
- Energized researchers and resulted in many new developments

Some Testimonials:

- Granger (2003): "Being asked to preview the book in 1968 was one of ten lucky breaks in my professional career"
- Granger (1986): Forecasts based on these methods are usually very diffcult to beat by alternative methods, even when time-varying parameters, non-linearties, or structural constraints are introduced.

- Diebold (1995): "The Past, Present, and Future of Macroeconomic Forecasting" described the large impact of Box and Jenkins (1970) and noted that "many of Box-Jenkins insights started literatures that grew explosively"

Autoregressive-Moving Average (ARMA) Models

Autoregressive model of order p, or AR(p) model:

$$y_t = \phi_1 y_{t-1} \cdots + \phi_p y_{t-p} + a_t$$

Mixed ARMA(p, q) model:

$$y_t = \phi_1 y_{t-1} \cdots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

In practice, $p + q$ is often ≤ 2

Measure of dependence in the series: $\rho_k = \text{Corr}(y_t, y_{t-k})$

Second-order Stationarity: The expected values, variances, and covariances of $y_1, \ldots, y_n$ is constant over time.

Stationarity condition: Roots of $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p = 0$ lie outside the unit circle.

Stationary Time Series; Example

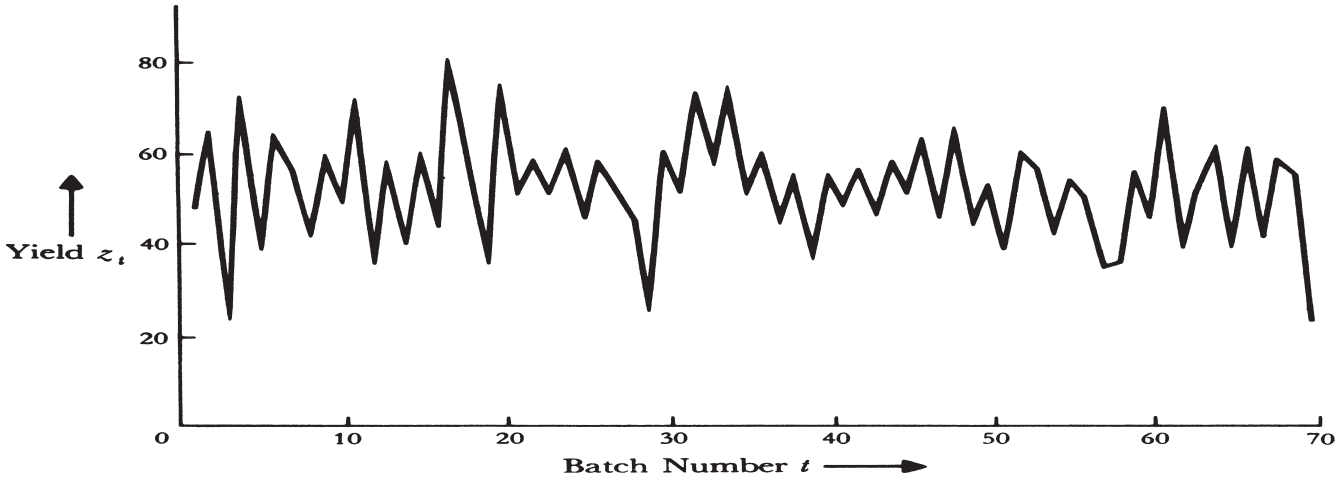


Figure: Yield of a Chemical Process

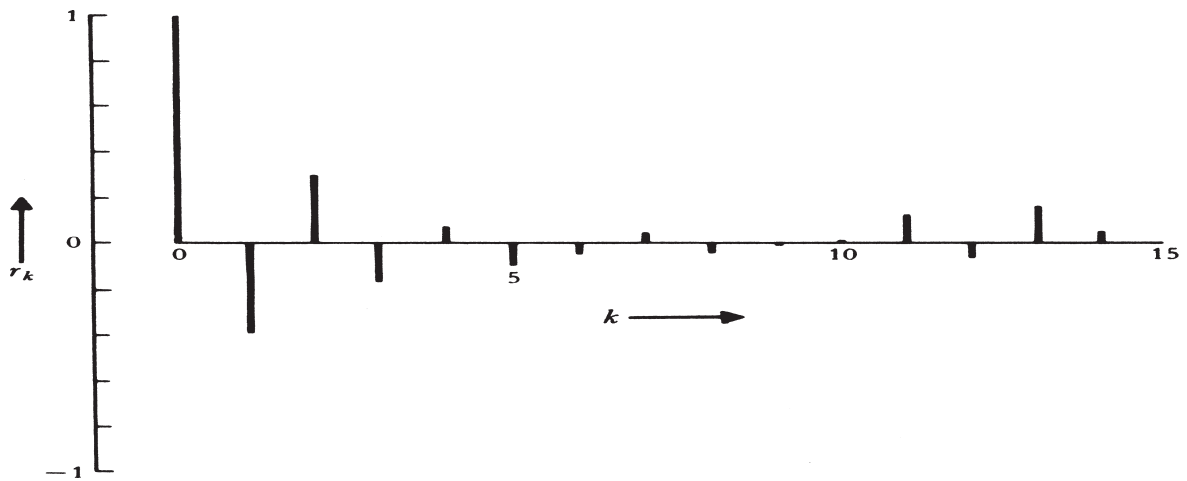
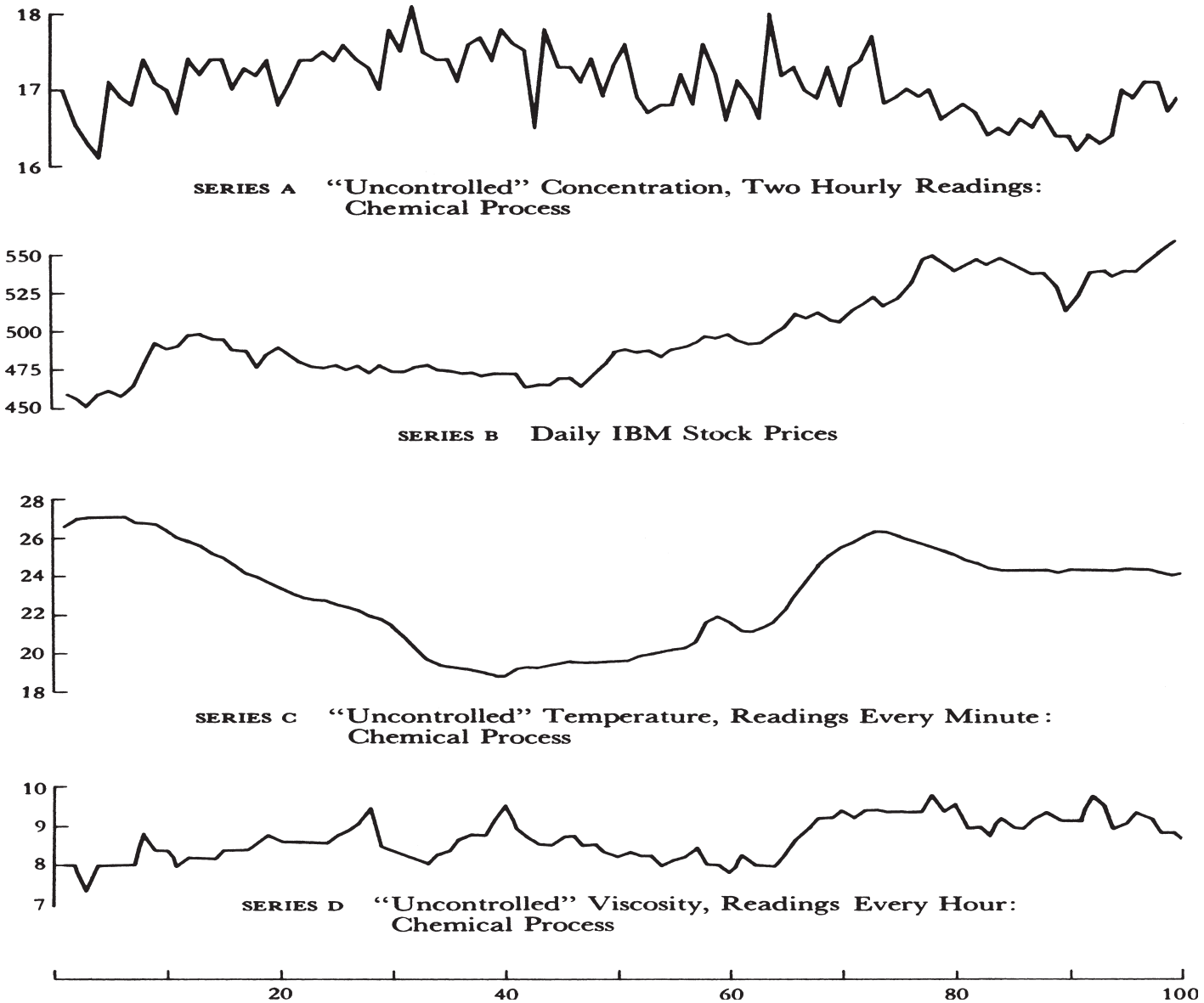


Figure: Sample Autocorrelation Function of the Yield Data

Examples of Some Non-stationary Time Series



Accomodating Non-Stationarity Through Differencing

Methods used for model building typically assume **stationarity**.

Many non-stationary time series be made stationary by **differencing**:

$$w_t = y_t - y_{t-1}$$

If y_t follows an ARMA(p, q), the model for w_t follows an ARIMA$(p, 1, q)$ model.

Random Walk model: $y_t = y_{t-1} + a_t$
Differencing gives: $w_t = a_t$

The random walk model can be written: $y_t = y_0 + \sum_{i=1}^{i=t} a_i$

The summation of the a's gives rise to a **stochastic trend**

Random walk with drift: $y_t = \beta + y_{t-1} + a_t = y_0 + \beta t + \sum_{i=1}^{i=t} a_i$

Model Building Procedure

Iterative 3-step Procedure

1. **Model identification:** Determine suitable values for (p, d, q). Useful tools:
   - Autocorrelation function
   - Partial-autocorrelation function
   - AIC and other criteria
2. **Parameter estimation**.
   - Conditional and unconditional least squares
   - Exact maximimum likelihood method
3. **Model checking**
   - Plots of residuals from the fitted model, Q-Q plots, etc.
   - Examination of the ACF and the PACF of the residuals
   - Several other tools available (including the "Ljung-Box" test)

Is Differencing Needed?

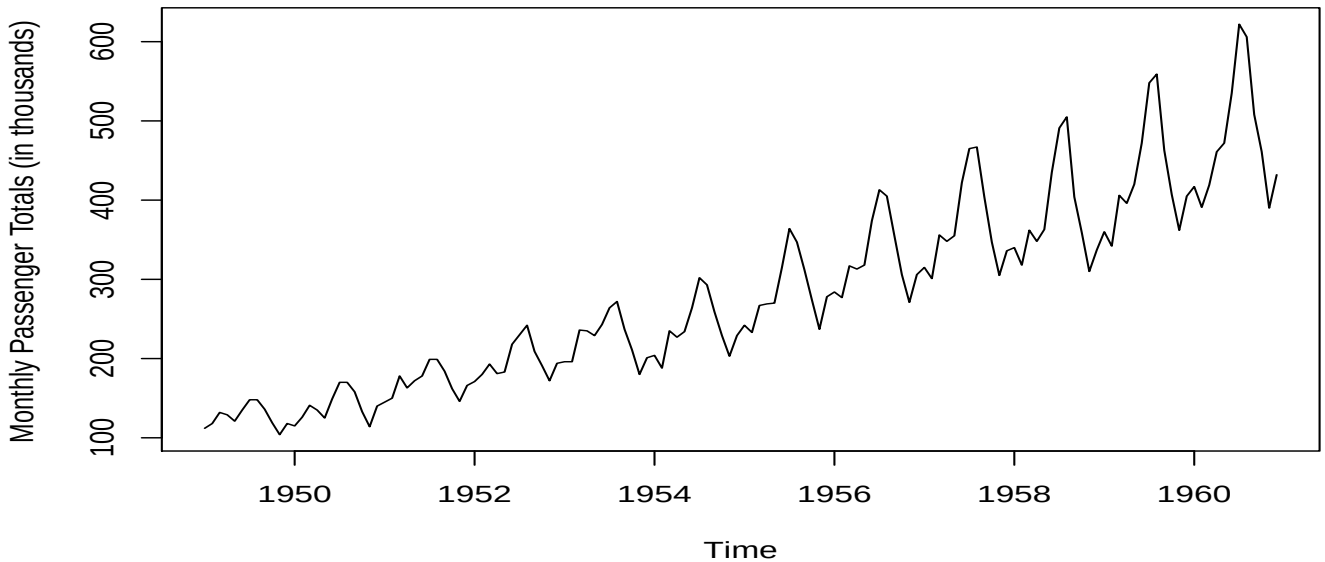
Essential tool: Visual inspection of time series plots and graphs of the sample autocorrelation and partial autocorrelation function

George Box: "You can see a lot by just looking".

Formal tests for differencing, or for unit roots, have been discussed extensively in the literature:

- Dickey and Fuller (1979, 1981)
- Dickey and Pantula (1987)
- Phillips and Perron (1988) and Perron (1988)
- Elliott, Rothenberg, and Stock (1996)
- Schmidt and Phillips (1992)
- And many more

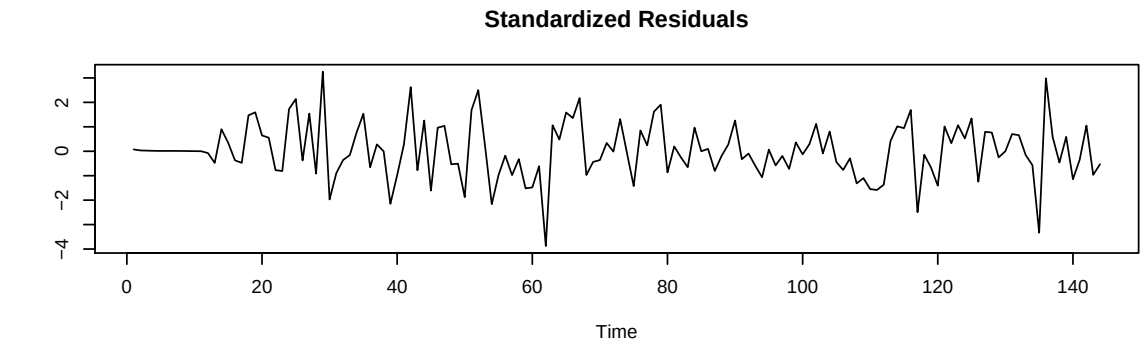
Example: Totals of International Airline Passengers

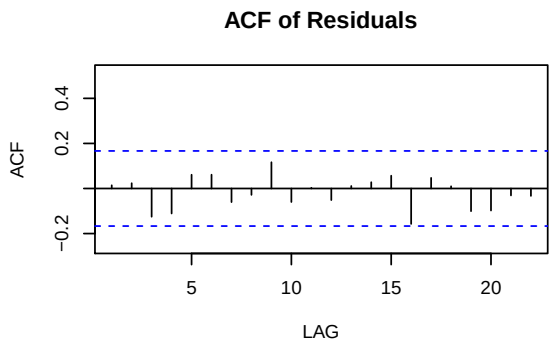


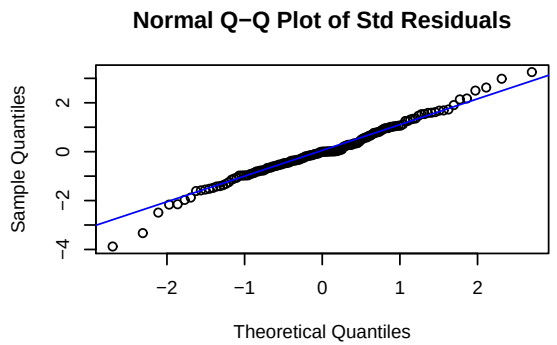
Modeling: After log transformation and double differencing, a model with two MA parameters provides a good fit and produces good forecasts.

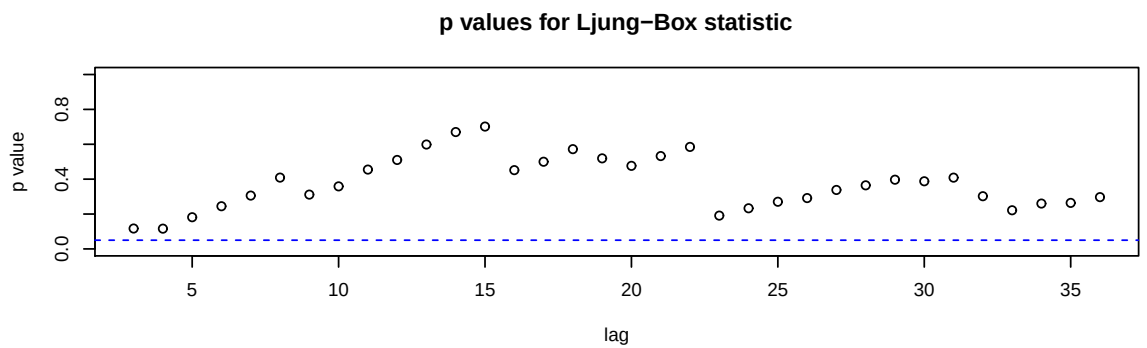
R Command: sarima(log.AP,0,1,1,0,1,1,S=12) in astsa package

Modeling Diagnostic: Logarithm of Airline Data

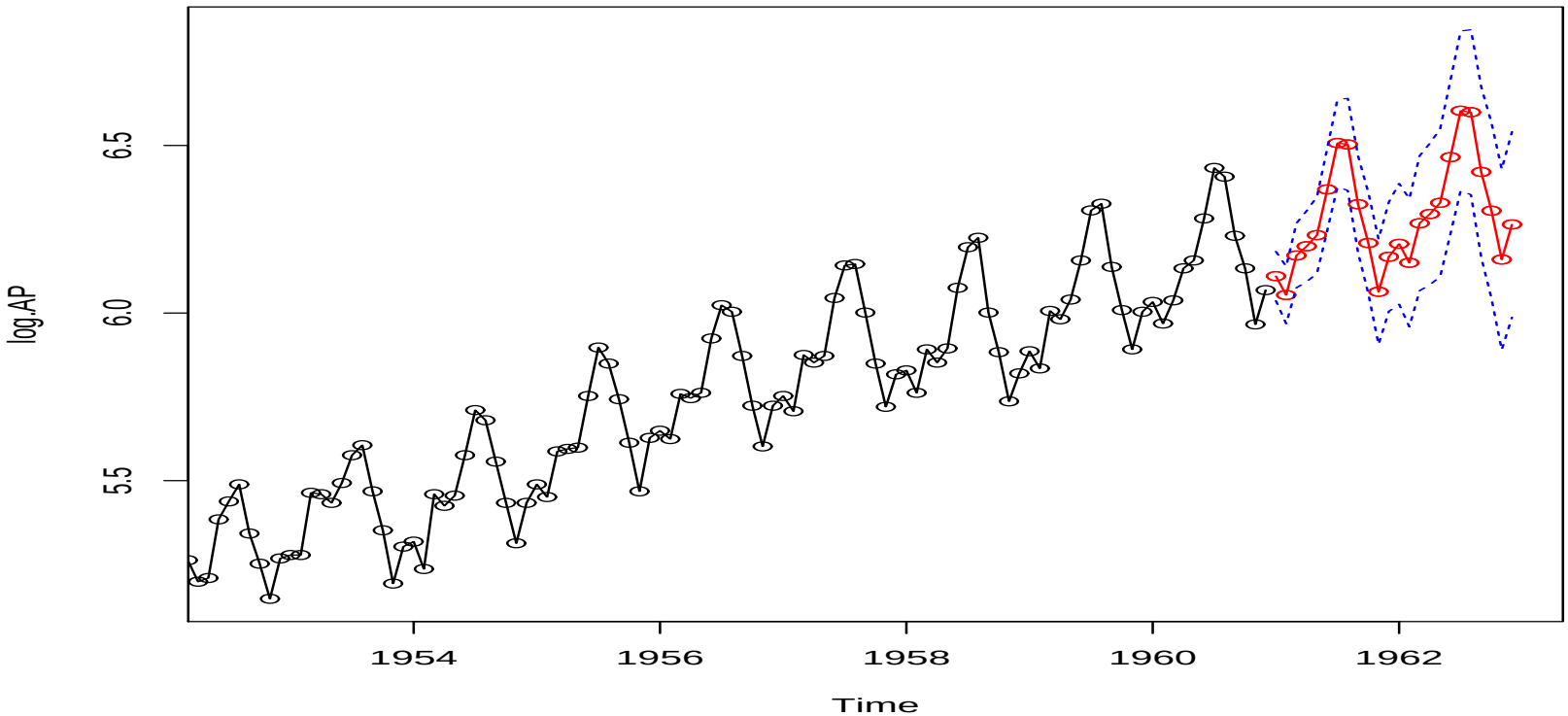








Two-Year-Ahead Forecasts Using Two Parameter Model



R command for forecasting 24 months ahead:
sarima.for(log.AP,24,0,1,1,0,1,1,12)

Seasonal Time Series Analyzed by Chatfield and Prothero (JRSS B, 1973): Disappointing Forecasts

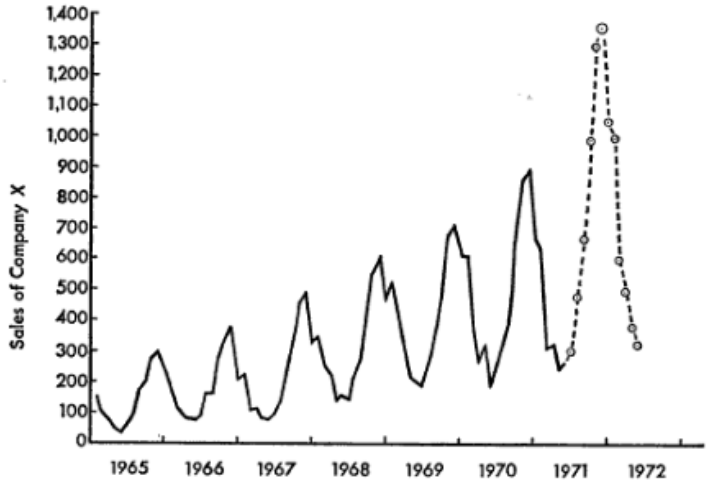


FIG. 1. Sales of Company *X*, January 1965–May 1971 and initial forecasts from May 1971 using model *A*.

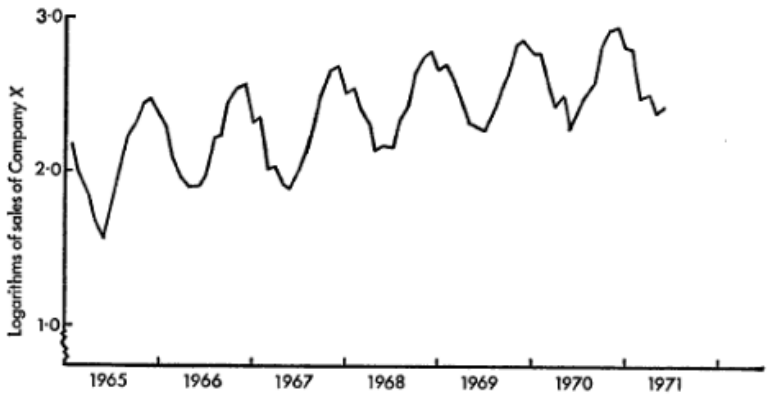


FIG. 2. Logarithms of Sales of Company *X*, January 1965–May 1971.

Figure: Sales of Company X

Standard Regression Model

Two or more variables:
Y_t = Output variable
X_t = Input variable

Standard regression model:

$$Y_t = v_0 + v_1 X_t + a_t$$

Some issues for time series data:

- The errors may not be uncorrelated. As a result, standard inference procedures based on the t and F distributions may not be valid.
- A change in the input variable may not take effect immediately. Delayed effects may be present.

Generalization: Transfer Function - Noise Model

Transfer function model:

$$Y_t = v_0 + v_1 X_t + v_2 X_{t-2} + \dots$$

More parsimonious representation:

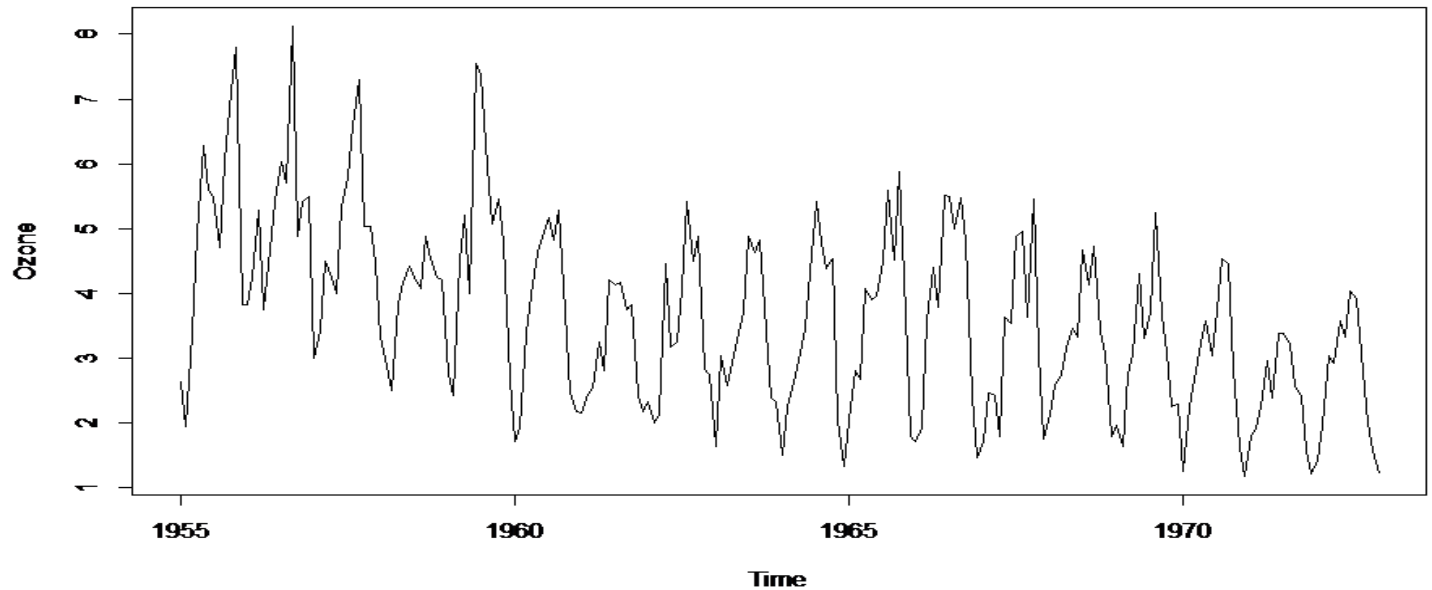
$$Y_t - \delta_1 Y_{t-1} - \cdots - \delta_r Y_{t-r} = X_{t-b} + \omega_1 X_{t-b-1} + \cdots + \omega_s X_{t-b-s}$$

Transfer function - noise model: A noise term that follows an ARMA(p, q) model is added to this model

Model building involves finding suitable values for r, b, s, p, and q

Interesting Application: Intervention Analysis

LA Ozone Data, 1955-1978; Box and Tiao, JASA (1974)



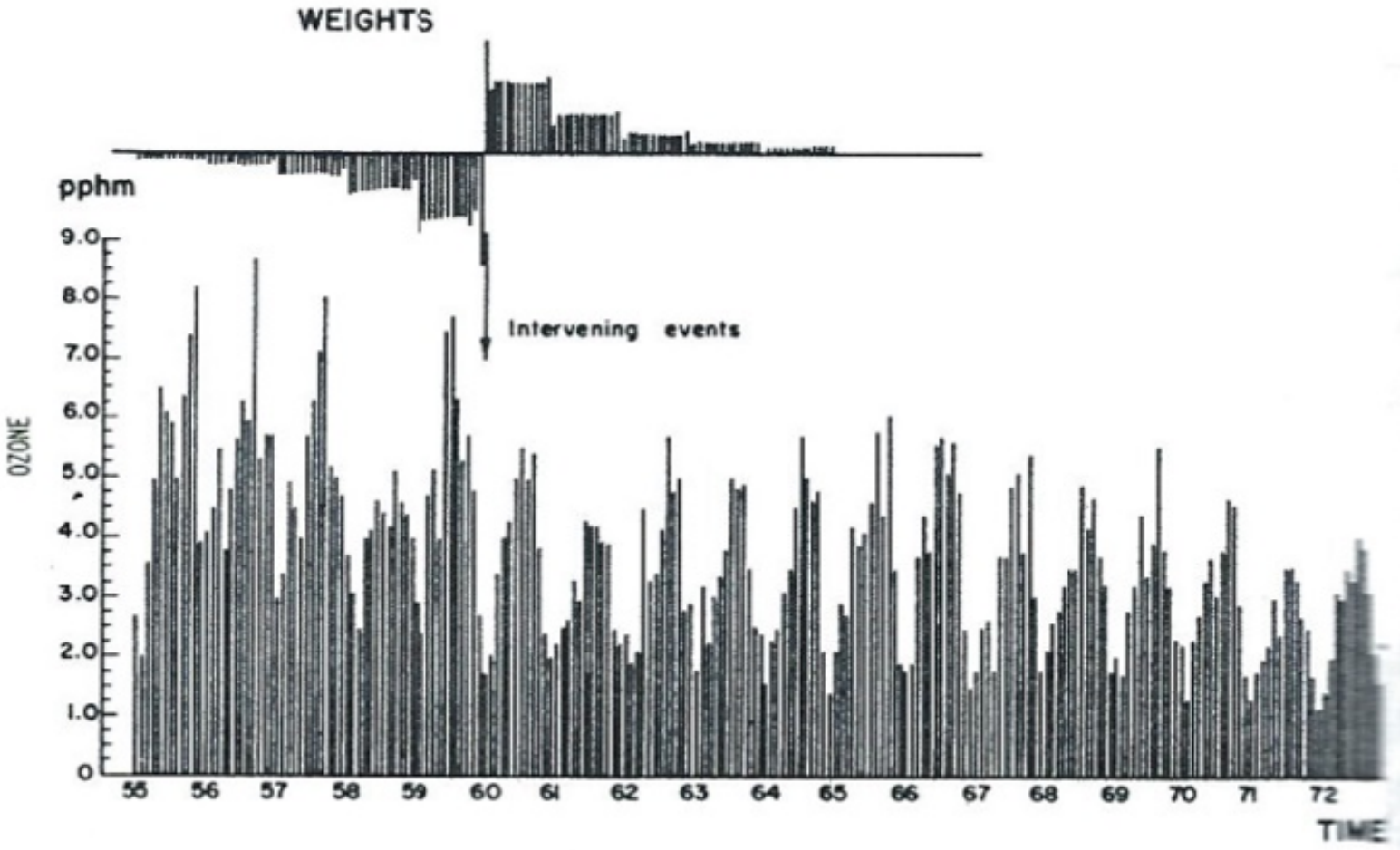
Two "interventions" in the early 1960:

1. Diversion of traffic by opening the Golden State Freeway
2. New law which reduced the allowable portion of reactive hydrocarbons in the gasoline.

Model: A seasonal time series model with intervention effects was fitted to the series.

LA Ozone Data; Estimation of Intervention Effect

A. Monthly Average of Hourly Readings of O_3 (pphm) in Downtown Los Angeles (1955–1972) [a]



[a] With the weight function for estimating the effect of intervening events in 19

Multivariate Extensions

The values of k different variables observed at each time point t

Goal: To describe and model the interrelations between the series

Vector Autoregressive Models, or VAR models include leads and lags and allow for feedback relationships between the series. Advocated by Sims (1980) as an alternative to traditional system-of-equations models.

Vector Autoregressive Moving Average (VARMA) Models discussed by Tiao and Box (1981) and many others.

Recommended Reading:

- Tsay, Ruey S. (2014), *Multivariate Time Series Analysis, With R and Financial Applications*, Wiley
- Reinsel, G. C. (1991) Elements of Multivatiate Time Series Analysis, Springer Verlag

Canonical Analysis of Five Agricultural Series; Box and Tiao (1977), Biometrika Vol. 64

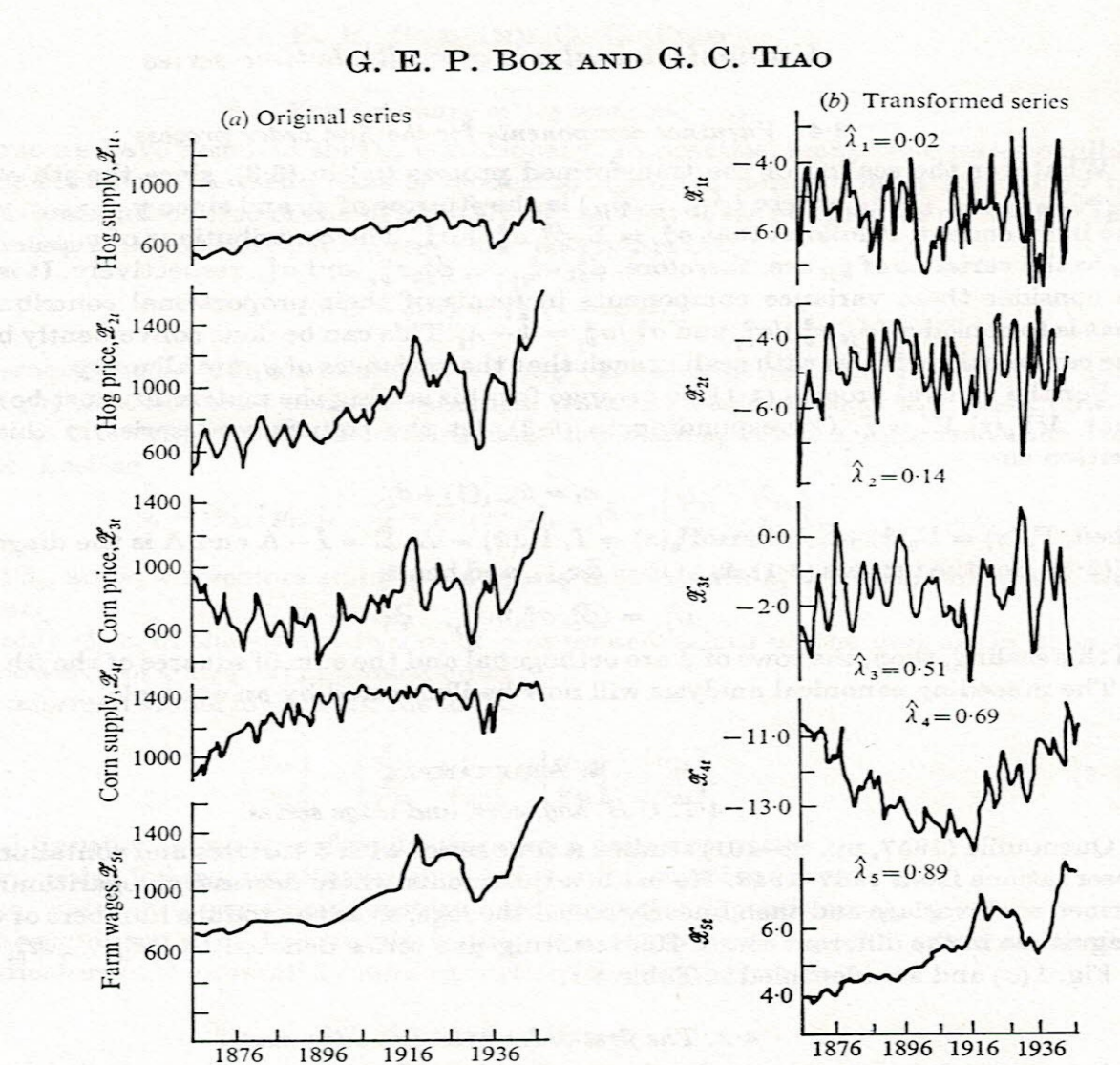


Fig. 1. U.S. hog data. (*a*) Original series, (*b*) transformed series.

Cointergration in Multiple Time Series

Definition: Two non-stationary time series are cointegrated if a linear combination of the two series is stationary

Cointegration suggests stable long-term relationships between non-stationary time series.

Granger and Engle (1987) presented a two-step procedure for estimating the cointegration vector and modeling the dynamic behavior of the different series.

Tools: Unit-root tests are used to determine the existence of cointegration. An Error Correction Model incorporates the cointegration relationships into the model.

Thank You, George!