

A BAYESIAN APPROACH TO ANALYZING
UNCERTAINTY AMONG STOCHASTIC MODELS

Eric F. Wood

September 1974

Research Reports are publications reporting on the work of the author. Any views or conclusions are those of the author, and do not necessarily reflect those of IIASA.

A Bayesian Approach to Analyzing
Uncertainty among Stochastic Models

Eric F. Wood

Abstract

The statistical uncertainty, resulting from the lack of knowledge of which modelling represents a given stochastic process, is analyzed. This analysis of model uncertainty leads to a composite Bayesian distribution. The composite Bayesian distribution is a linear model of the individual Bayesian probability distributions of the individual models, weighted by the posterior probability that a particular model is the true model. The composite Bayesian probability model accounts for all sources of statistical uncertainty-- both parameter uncertainty and model uncertainty. This model is the one that should be used in applied problems of decision analysis, for it best represents the knowledge-- or lack of it--to the decision maker about future events of the process.

Introduction

Applied scientists are often confronted with the problem of choosing one statistical model from many contending models. An example of this selection problem is frequently encountered by hydrologists in flood frequency analysis. The examples and applications in this paper will be addressed to that problem.

Consider the problem of the hydrologist who must make a decision between a number of alternate designs that propose to prevent or decrease the occurrence of future floods. His first task is to make inferences about the underlying

process that generates these events but, in addressing this problem, he is faced with a number of sources of uncertainty. These sources of uncertainty have often been summarized into three categories [1]:

1. Natural uncertainty. This is the uncertainty in the stochastic process -- the occurrence of extreme streamflows, q .
2. Statistical uncertainty. This is associated with the estimation of the parameters of the model of the stochastic process due to limited data.
3. Model uncertainty. This is associated with the uncertainty that a particular probabilistic model of the stochastic process may not be the true model. Most hydrologic processes are so complex that no model yet devised may be the true model, or maybe hydrologic events follow no particular model.

Many models seem to fit the available data very well, but often the models lead to different inferences and decisions. In recent years, considerable progress has been made on the development of statistical procedures for comparing alternative models; examples of this are Gaver and Geisel [3], Smallwood [8] and Leamer [4], who all used Bayesian statistical procedures, and Dumonceaux et al. [2] and Pesaran [5] who applied "classical" statistical procedures of hypothesis testing.

Composite Bayesian Distribution

For a particular model of flood events, parameter uncertainty can be accounted for by considering the Bayesian pdf of flood events, which is

$$\tilde{f}(q) = \int_A f(q|A) \cdot f''(\underline{A}) d\underline{A} \quad (1)$$

where $\tilde{f}(q)$ is the Bayesian pdf for q ,
 $f(q|\underline{A})$ is the "modelled" pdf of q , conditional upon the uncertain parameter set \underline{A} , and
 $f''(\underline{A})$ is the posterior pdf for the parameter set \underline{A} .

Model uncertainty can be considered by defining a composite model of the form

$$\hat{f}(q|\underline{A},\underline{\theta}) = \theta_1 \cdot f_1(q|\underline{A}_1) + \dots + \theta_n \cdot f_n(q|\underline{A}_n) \quad (2)$$

where $A = \bigcup_{i=1}^n A_i$.

The composite model, $\hat{f}(q|\underline{A},\underline{\theta})$, is conditioned upon a set of unknown model parameters \underline{A} and an unknown composite model parameter set $\underline{\theta}$.

$f_1(q|\underline{A}_1), \dots$, and $f_n(q|\underline{A}_n)$ is the set of probabilistic models that make up the composite model. These models are conditioned upon a general unknown parameter set \underline{A} .

θ_1, \dots , and θ_n are parameters that take on a value of either 0 or 1; their value is uncertain. If $\theta_1 = 1$, then model $f_1(q|\underline{A}_1)$ is the true model. The constraint

$$\sum_{i=1}^n \theta_i = 1 \quad (3)$$

is imposed, which implies that one and only one model is the true model.

For notational simplicity, consider the case where $n = 2$. The likelihood function for a set of observations Q is just:

$$\begin{aligned} L(\underline{A}, \underline{\theta} \mid \underline{Q}) &= \theta_1 \prod_{i=1}^n f_1(q_i \mid \underline{A}_1) + \theta_2 \prod_{i=1}^n f_2(q_i \mid \underline{A}_2) \\ &= \theta_1 \cdot L_1(\underline{A}_1 \mid \underline{Q}) + \theta_2 \cdot L_2(\underline{A}_2 \mid \underline{Q}) \end{aligned} \quad (4)$$

There are no cross products of the models, due to the limitation imposed on the values that θ_i can take on; and the constraint on $\underline{\theta} \cdot L_i(\underline{A}_i \mid \underline{Q})$ is just the likelihood function of model i , conditional upon the observations, \underline{Q} .

Define now a composite prior distribution on the parameters \underline{A} and $\underline{\theta}$. The prior will be of the form

$$\begin{aligned} f'(\underline{A}, \underline{\theta}) &= \theta_1 f'_1(\underline{A}_1 \mid \theta_1 = 1) \cdot p'(\theta_1 = 1) \\ &+ \theta_2 f'_2(\underline{A}_2 \mid \theta_2 = 1) \cdot p'(\theta_2 = 1) \end{aligned} \quad (5)$$

$f'_i(\underline{A}_i \mid \theta_i = 1)$ is the prior distribution on the parameter set \underline{A} , conditional upon $\theta_i = 1$. $p'(\theta_i = 1)$ is the prior probability that model i is the "true" model.

Bayes' rule can be written as

$$f''(b|\text{data}) = \frac{1}{K} L(b|\text{data}) \cdot f'(b) \quad . \quad (6)$$

$f''(b|\text{data})$ is the posterior distribution of the b , conditional upon the data; $L(b|\text{data})$ is the likelihood function for b ; $f'(b)$ is the prior distribution of b ; and K is a normalizing constant.

The normalizing constant K is often called, in the econometrics literature, the marginal density of the observations or the marginal likelihood [12] and can be found by

$$K = \int_b L(b|\text{data}, \text{model}) \cdot f'(b|\text{model}) db. \quad (7)$$

K_i , the marginal likelihood function for model i , can be thought of as the probability of observing the data, given model i .

The posterior density function for $\underline{A}, \underline{\theta}$ is calculated from Bayes' rule; it is

$$\begin{aligned} f''(\underline{A}, \underline{\theta}) &\propto \{\theta_1 \cdot L_1(\underline{A}|\underline{Q}) + \theta_2 \cdot L_2(\underline{A}|\underline{Q})\} \cdot \{\theta_1 \cdot f'_1(\underline{A}_1|\theta_1=1) \\ &\quad \cdot p'(\theta_1=1) + \theta_2 \cdot f'_2(\underline{A}_2|\theta_2=1) \cdot p'(\theta_2=1)\} \\ &\propto \theta_1 K_1 f''(\underline{A}_1|\theta_1=1) \cdot p'(\theta_1=1) \\ &\quad + \theta_2 K_2 f''(\underline{A}_2|\theta_2=1) \cdot p'(\theta_2=1) \\ &= \theta_1 \frac{K_1}{K^*} p'(\theta_1=1) \cdot f''(\underline{A}_1|\theta_1=1) \\ &\quad + \theta_2 \frac{K_2}{K^*} p'(\theta_2=1) \cdot f''(\underline{A}_2|\theta_2=1) \end{aligned} \quad (8)$$

where K^* is a normalizing constant equal to

$$K^* = K_1 \cdot p'(\theta_1 = 1) + K_2 \cdot p'(\theta_2 = 1) \quad . \quad (9)$$

The posterior model probabilities, $p''(\theta_i)$ are

$$p''(\theta_1=1) = \frac{K_1}{K^*} p'(\theta_1=1) \quad (10)$$

$$p''(\theta_2=1) = \frac{K_2}{K^*} p'(\theta_2=1) \quad . \quad (11)$$

These posterior probabilities for θ_i are the same as those found by Leamer [4], Gaver and Geisel [3], and Smallwood [8], even though their approaches to the problem were different.

The composite Bayesian distribution of extreme flood events, q , can also be found by applying first principles:

$$\begin{aligned} \tilde{f}(q) &= \int_{\underline{A}, \underline{\theta}} \hat{f}(q|\underline{A}, \underline{\theta}) \cdot f''(\underline{A}, \underline{\theta}) \, d\underline{A}d\underline{\theta} \\ &= \int_{\underline{A}, \underline{\theta}} \{ \theta_1 \cdot f_1(q|\underline{A}_1) + \theta_2 \cdot f(q|\underline{A}_2) \} \\ &\quad \cdot \{ p''(\theta_1=1) \cdot f''(\underline{A}_1|\theta_1) + p''(\theta_2=1) \cdot \\ &\quad \cdot f''(\underline{A}_2|\theta_2=1) \} \cdot d\underline{A}d\underline{\theta} \\ &= p''(\theta_1=1) \cdot \tilde{f}(q) + p''(\theta_2=1) \cdot \tilde{f}_2(q) \quad . \quad (12) \end{aligned}$$

The composite Bayesian distribution is simply the Bayesian distributions of the models weighted by the posterior probability that a particular model is the true model. This result is extremely convenient.

Analytical Derivation of the Marginal Density Function

The marginal density function of a set of observations is calculated from Equation (7), and represents the probability of observing that set of data. The marginal density function depends upon the probability model for the stochastic process, the prior probability density function over the parameters of the model and the set of observed data. Consider the marginal likelihood function for the following cases:

1. Normal Process

Let the random variable q be distributed with Normal mean μ and precision h . The probability density for q is

$$f(q|\mu,h) = \frac{1}{\sqrt{2\pi}} h^{\frac{1}{2}} \exp \left\{ -\frac{h}{2} (q-\mu)^2 \right\} . \quad (13)$$

Then, given n independent observations of q , \underline{Q} , the likelihood function for μ and h is

$$\begin{aligned} L(\mu,h|\underline{Q}) &= \prod_{i=1}^n f_N(q_i|\mu,h) \\ &= (2\pi)^{-n/2} h^{n/2} \cdot \exp \left\{ -\frac{h}{2} \sum (q_i - \mu)^2 \right\} . \end{aligned} \quad (14)$$

Define the following

$$m = \frac{1}{n} \sum q_i \quad (15)$$

$$v = \frac{1}{n-1} \sum (q_i - m)^2 \quad (16)$$

$$v = n-1 ,$$

then

$$L(\mu, h | \underline{Q}) = (2\pi)^{-n/2} \exp\{-\frac{1}{2} h v v - \frac{1}{2} h n (m - \mu)^2\} \cdot h^{n/2} \quad (17)$$

Assume the prior on (μ, h) is a natural conjugate prior¹ of the form

$$f'(\mu, h) = (2\pi)^{-1/2} \exp\{-\frac{1}{2} h n' (\mu - m')^2 - \frac{1}{2} h v' v'\} \\ \cdot n'^{1/2} h^{1/2} \cdot h^{1/2} v'^{-1} \frac{(\frac{1}{2} v' v')^{1/2} v'}{\Gamma(\frac{1}{2} v')} \quad (18)$$

Then, the marginal likelihood function for the Normal model,

$$K_N = \int_{\mu} \int_h L(\mu, h | \underline{Q}) \cdot f'(\mu, h) d\mu dh$$

is from Equation (14) and (18)

$$K_N = n'^{1/2} (2\pi)^{-v/2} \frac{(\frac{1}{2} v' v')^{1/2} v'}{\Gamma(\frac{1}{2} v')} \\ \int_{\mu} \int_h (2\pi)^{-1/2} h^{1/2} \exp\{-\frac{1}{2} h n'' \cdot (\mu - m'')^2\} \\ \cdot h^{1/2} v''^{-1} \exp\{-\frac{1}{2} h v'' v''\} d\mu dh \quad , \quad (19)$$

where

$$m'' = \frac{n' m' + n m}{n' + n}$$

$$n'' = n' + n$$

¹For the Normal process, the natural conjugate over the mean and precision is Normal-Gamma (Raiffa and Schlaifer, [6]).

$$v'' = \frac{1}{v''} (v'v' + nm'^2 + vv + nm^2 - n''m''^2)$$

$$v'' = v' + v + 1 = n'' - 1$$

The integral is equal to

$$\frac{1}{n''^{1/2}} \cdot \frac{\Gamma(1/2 v'')}{(1/2 v''v'')^{1/2v''}} \quad (20)$$

Thus

$$K_N = \left(\frac{n'}{n''}\right)^{1/2} \cdot (2\pi)^{-v/2} \cdot \frac{\Gamma(1/2 v'')}{\Gamma(1/2 v')}$$

$$\cdot \frac{(1/2 v'v')^{1/2v'}}{(1/2 v'v')^{1/2v''}} \quad (21)$$

2. Log-Normal Process

Let $x_i = \ln q_i$ be distributed Normal with mean μ and precision h . Then q_i is distributed Log-Normal by definition. The probability density function for q is

$$f(q|\mu, h) = \frac{1}{q} (2\pi)^{-1/2} h^{1/2} \exp\{-\frac{1}{2}h(x - \mu)^2\} \quad (22)$$

The likelihood function for μ and h , given n independent observations of q is

$$L(\mu, h|Q) = \frac{1}{\prod_{i=1}^n q_i} (2\pi)^{-n/2} \cdot h^{n/2}$$

$$\cdot \exp\{-\frac{1}{2}hn\sum(x_i - \mu)^2\} \quad (23)$$

Assume a Normal-Gamma prior for μ and h of the same form as Equation (18). The marginal likelihood, K_{LN} , is just the integration of μ and h over the product of the likelihood and the prior probability density function.

$$K_{LN} = \frac{1}{n} \prod_{i=1}^n q_i \int_{\mu, h} (2\pi)^{-n/2} h^{n/2} \cdot \exp \left\{ -\frac{1}{2} h n \sum (x_i - \mu)^2 \right\} \cdot f'(\mu, h) d\mu dh. \quad (24)$$

The integral is of the same form as the marginal likelihood for the Normal model. Then, from Equation (21), K_{LN} is:

$$K_{LN} = \frac{1}{n} \prod_{i=1}^n q_i \cdot \left(\frac{n'}{n''} \right)^{1/2} (2\pi)^{-v/2} \frac{(1/2 v'')}{(1/2 v')} \cdot \frac{(1/2 v' v')^{1/2} v'}{(1/2 v'' v'')^{1/2} v''} \quad (25)$$

3. Exceedance Model

Another model of common use in water resources, especially in the analysis of extreme events, is the Exceedance model. (Shane and Lynn, [7]; Wood, [10]; Todorovic and Zelenhasic, [9].) The Exceedance model considers only those extreme events, let's say flood discharges, greater than a specified base level. Such discharges are called exceedance discharges and the probability density function of exceedance discharges is assumed to be of an Exponential type. Furthermore, the arrival of exceedance events is assumed to be a Poisson process. Such a model is of a general form since the upper tails of many distributions

can be approximated by an exponential form.

The second part of this model concerns flood discharges less than the base level. Usually such discharges are of little interest in analyzing extreme events, and the distribution of such events may be quite complex. Here, it will be assumed that the events will follow a uniform distribution. The use of the uniform b density function implies that the posterior probability for the Exceedance model will be underestimated or conservative.

The probability density function for the Exceedance model is

$$\begin{aligned}
 f(q|v,\alpha) &= v \alpha \exp \{-\alpha(q - q_b)\} \quad \text{for } q \geq q_b \\
 &= \frac{1-v}{q_b} \quad , \quad \text{for } 0 \leq q \leq q_b \quad , \quad (26)
 \end{aligned}$$

where v is the arrival rate of floods, α is the event magnitude parameter and q_b is the base level.

Given a sample of n independent discharges, \underline{Q} , of which m are discharges less than q_b and $n-m$ are discharges greater than or equal to q_b , then the likelihood function for v and α can be shown to be,

$$\begin{aligned}
 L(v,\alpha|\underline{Q}) &= \frac{(1-v)^m}{q_b^m} \cdot v^{n-m} \alpha^{n-m} \exp \{-\alpha \sum_{i=1}^{n-m} (q_i - q_b)\} \\
 &\cdot \exp\{-v \sum t_i\} \quad . \quad (27)
 \end{aligned}$$

The marginal likelihood function, K_E , is defined as

$$K_E = \int_{\nu, \alpha} L(\nu, \alpha | \underline{Q}) \cdot f'(\nu) \cdot f'(\alpha) d\nu d\alpha \quad . \quad (28)$$

The conjugate prior density function for ν and α are of the form

$$f'(\nu) = \exp(-s' \cdot \nu) \cdot \nu^{u'-1} \frac{s', u'}{\Gamma(u')} \quad (29)$$

$$f'(\alpha) = \exp(-\ell' - \alpha) \cdot \alpha^{\nu'-1} \frac{\ell', \nu'}{\Gamma(\nu')} \quad . \quad (30)$$

Therefore, from Equation (28) applying Equations (27) and (29) K_E is simply

$$K_E = q_b^{-m} \cdot \frac{s', u'}{\Gamma(u')} \int_{\nu} \exp\{-(s'' + m)\nu\} \cdot \nu^{(u'+n-m)-1} d\nu \cdot \frac{\ell', \nu'}{\Gamma(\nu')} \int_{\alpha} \exp\{-\ell' + \Sigma(q_1 - q_b)\} \cdot \alpha^{(\nu'+n-m)-1} d\alpha \quad . \quad (31)$$

The integral over ν equals

$$\frac{\Gamma(u'')}{(s'' + m)^{u''}} \quad , \quad (32)$$

where

$$u'' = u' + n - m$$

$$s'' = s' + T \quad (\text{or } s'' = s' + \Sigma t_i)$$

and the integral over α equals

$$\frac{\Gamma(v'')}{(\ell'')^{v''}} \quad , \quad (33)$$

where

$$\begin{aligned} v'' &= v' + n - m \\ \ell'' &= \ell' + \Sigma(q_i - q_b) \quad . \end{aligned}$$

Thus, K_E equals

$$K_E = q_b^{-m} \cdot \frac{s'^{u'}}{(s''+m)^{u''}} \cdot \frac{\Gamma(u'')}{\Gamma(u')} \cdot \frac{\ell'^{v'}}{\ell''^{v''}} \cdot \frac{\Gamma(v'')}{\Gamma(v')} \quad . \quad (34)$$

Some computer experiments were carried out with samples generated from known distributions. As an example, a sample growing from 10 to 200 was generated from a Log-Normal distribution with $\mu_{\ell_{ny}} = 7.85$ and $\sigma_{\ell_{ny}} = 0.95$ and the marginal likelihoods were numerically evaluated for the Log-Normal and the Exceedance models assuming diffuse prior distributions on the probability model parameters. Table 1 shows the values of the marginal likelihoods jointly with the posterior model probabilities estimated according to Equations (10) and (11) on the assumption of diffuse prior model probabilities ($p'(\theta_1 = 1) = p'(\theta_2 = 1) = 0.5$). Extensive experiments are presently being performed to evaluate the worth of data on the problem of model selection as well as the influence of prior assessments, and the results will be forthcoming.

An Application to the Blackstone River, U.S.A.

The Blackstone River, at Woonsocket, Rhode Island, has been analyzed by Wood and Rodriguez [11] for prior information for the Bayesian probability density function of its flood discharges (for four different probability models), and for a decision problem concerning local flood protection. Model uncertainty was not considered in the previous paper even though competing models were considered. This section calculates the posterior model probabilities. The parameters for the marginal likelihood functions are summarized in Table 2. The values of the marginal likelihoods are

$$K_N = 7.46 \times 10^{-191} , \quad K_{LN} = 4.76 \times 10^{-160} ,$$
$$K_E = 1.14 \times 10^{-156}$$

for the Normal, Log-Normal, and Exceedance models, respectively.

Assuming uniform prior probabilities on the three models, the posterior probabilities for the models are

$$p''(\theta_N = 1) = 0.0$$
$$p''(\theta_{LN} = 1) = .00418$$
$$p''(\theta_E = 1) = .99582 .$$

The composite Bayesian distribution of flood discharges is, from Equation (12)

$$\tilde{f}(q) = .99582 \tilde{f}_E(q) + .00418 \tilde{f}_{LN}(q) , \quad (35)$$

where $\tilde{f}_E(q)$ is the Bayesian density function for the Exceedance model, and $\tilde{f}_{LN}(q)$ is the Bayesian density function for the Log-Normal model.

The composite Bayesian distribution of Equation (36) is the probability model which should be used in making inferences about future flood discharges. The composite Bayesian model rationally accounts for both parameter and model uncertainty. It is interesting to note that the form of composite Bayesian model is not fixed, but is dynamic and changes as more data becomes available.

Conclusions

This paper considers the problem of model uncertainty within a Bayesian analysis. When there is a set of competing probability models for flood discharges, Bayesian analysis leads to a composite Bayesian model. The composite Bayesian model is a linear model consisting of the Bayesian distribution of the individual models, weighted by the posterior model probability that the individual model is the true model. The posterior model probabilities are calculated from the marginal likelihood function of the observed data and the

prior model probability.

The posterior model probabilities are found by calculating the marginal likelihood function for each competing model. The marginal likelihood function was derived analytically for three commonly used models -- a Normal process, a Log-Normal process, and an Exceedance model. The results have been applied to "real-world" data and favourable results obtained.

Table 1: Marginal Likelihoods and Posterior Model Probabilities
for Samples Generated from Log-Normal Process with
 $\mu_{\ln y} = 7.8$ and $\sigma_{\ln y} = 0.95$

<u>Log-Normal Model</u>		<u>Sample Size</u>	<u>Exceedance Model</u>	
Marginal Likelihood	Posterior Model Probability		Marginal Likelihood	Posterior Model Probability
9.83594×10^{-43}	0.525	10	8.90135×10^{-43}	0.475
10.3726×10^{-83}	0.671	20	5.09288×10^{-83}	0.329
14.4061×10^{-165}	0.678	40	6.85213×10^{-165}	0.322
15.3628×10^{-245}	0.997	60	13.6377×10^{-248}	0.003

Table 2: Marginal Likelihood Parameters for Normal, Log-Normal, and Exceedance Models for the Blackstone River, U.S.A.

Normal Model

$n' = 7$ years	$n'' = 44$ years
$v' = 36$ years	$v'' = 43$ years
$v' = 9.22 \times 10^6$ cfs ²	$v'' = 24.7 \times 10^6$ cfs ²

Log-Normal Model

$n' = 4$ years	$n'' = 41$ years
$v' = 36$ years	$v'' = 40$ years
$v' = .22 \log$ cfs ²	$v'' = .689 \log$ cfs ²

Exceedance Model

$u' = 6$ events	$u'' = 11$ events
$v' = 3$ events	$v'' = 8$ events
$S' = 50$ years	$S'' = 87$ ($S''+m=119$) years
$l' = 10850$ cfs	$l'' = 49468$ cfs
$m = 32$ events	$n = 5$ events
$q_b = 8500$ cfs	

References

- [1] Benjamin, Jack R. and C. Allin Cornell. Probability, Statistics and Decision for Civil Engineers, McGraw-Hill, New York, 1970.
- [2] Dumonceaux, Robert, Charles E. Artle, and Gerald Haas. "Likelihood Ratio Test for Discrimination Between Two Models with Unknown Location and Scale Parameters," Technometrics, Vol. 15, No. 1 (February 1973).
- [3] Gaver, K. and M. Geisel. "Discriminating Among Alternative Models: Bayesian and Non-Bayesian Methods," to appear in Frontiers of Econometrics, Paul Zaremkka, Editor, Academic Press, forthcoming, 1972.
- [4] Leamer, E. "Probabilities of Linear Hypothesis," Harvard University, unpublished working paper, 1973.
- [5] Pesaran, M.H. "On the General Problem of Model Selection," The Review of Economic Studies, Vol. XLI (2), No. 126 (April 1974).
- [6] Raiffa, Howard and Robert Schlaifer. Applied Statistical Decision Theory, MIT Press, Cambridge, Mass., 1961, p. 356.
- [7] Shane, R.M. and W.R. Lynn. "Mathematical Model for Flood Risk Evaluation," Journal of the Hydraulics Division, American Society of Civil Engineers, Vol. 90, No. HY-6 (June 1964).
- [8] Smallwood, R. "A Decision Analysis of Model Selection," IEEE Transactions on Systems Science and Cybernetics, Vol. SSC-4, No. 3, September 1968.
- [9] Todorovic, P. and E. Zelenhasic. "A Stochastic Model for Flood Analysis," Water Resources Research, Vol. 6, No. 6 (December 1970).
- [10] Wood, Eric F. "Flood Control Design with Limited Data - A Comparison of the Classical and Bayesian Approaches," Symposium on the Design Water Resources Projects with Inadequate Data, Madrid, June 1973.
- [11] Wood, Eric F. and I. Rodriguez-Iturbe. "Bayesian Inference and Decision Making for Extreme Hydrologic Events," submitted for publication, Water Resources Research, 1974.
- [12] Zellner, A. An Introduction to Bayesian Inference in Econometrics, Wiley, New York, 1971.