STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–10.
Published online in International Academic Press (www.IAPress.org)

Estimating and Forecasting Multivariate Autoregressive Time Series Models Using Goal Programming

Dina Ali Bekhit*, Ramadan Hamed, Mohamed Ismail, Mahmoud Rashwan

Faculty of Economics and Political Science, Cairo University, Egypt

Abstract This paper presents a new method for estimation and prediction of vector autoregressive (VAR) models by goal programming (GP). VAR models have been widely used in economics, finance, and engineering, but most of the times they are estimated and forecasted using least squares or maximum likelihood estimation, which are sensitive to non-normality, outlier, and small samples. The proposed GP is a re-formulation of the estimation of VAR models as an optimization problem, where the absolute difference is minimized to enhance the robustness. Validation of the methodology is carried out through a large Monte Carlo simulation study based on 108 scenarios, and the real world application with economic series data for USA and environmental and climate science data for China. Results show that GP has more forecasting accuracy and robustness than comparison methods (ordinary least squares (OLS), Bayesian VAR, M-estimator), especially when contaminated and outlier conditions arise. This study is also a contribution to the literature as it is the first to use GP to solve a multivariate time series problem, which offers a flexible and reliable tool for researchers and practitioners who face problems in data-driven decision making.

Keywords Vector Autoregressive Models; Goal Programming; Multivariate Time Series; Forecasting; Robust Estimation

DOI: 10.19139/soic-2310-5070-4052

1. Introduction

The ability to predict is a central one in the decision-making process of economy, finance, engineering and several other sciences. Policymakers need forecasts of economic activity to make decisions about monetary and fiscal policies, financial institutions need forecasts of risk and volatility to allocate capital wisely, and engineers need forecasts of sensor and process data to optimize the operation of an industrial process. In all these applications, accurate and robust forecasting methods for multivariate time series (MTS) are critical; they enable the modeling of interactions among variables, and offer more information and insight for predicting than univariate methods.

The MTS model has been widely adopted for its flexibility and its ability to describe interdependencies among multiple time series, in particular among the vector autoregressive (VAR) models. The VAR model has been used and has been shown to be effective for macroeconomic forecasting, financial econometrics, energy demand prediction, and even for industrial process control. Although they offer flexibility, the modeling and prediction of VARs are still difficult. Classical approaches, like ordinary least squares (OLS) and maximum likelihood estimation (MLE) work well when the data are normally distributed, but are not always accurate when the data are found to have structural breaks, non-normality, heavy tails or outliers in the data set. Also methods like BVAR, M-estimator are inflexible to add additional constraints and sensitive to prior selection.

To overcome the above drawbacks, the Goal Programming (GP) is introduced as a better alternative. The GP framework is very flexible and enables handling many complex constraints and meeting target aspirations, unlike BVAR or M-estimation approaches, which are restricted to specific data distributions and/or prior sensitivities.

*Correspondence to: Dina Ali Bekhit, Faculty of Economics and Political Science, Cairo University, Egypt. Email: docdina2021@gmail.com

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Several studies have shown that linear programming (LP) and goal programming (GP) give better robustness and accuracy in AR, ARMA and ARIMA models. For example, Kiountouzis [1] proved that LP is a better estimator if the errors have a specific distribution (Laplace or Cauchy). Baker and Fitzpatrick [2] utilized a multistep method to identify the optimal parameters of an exponential smoothing model for forecasting emergency medical service (E.M.S.) demand across four counties in South Carolina. Additionally, Feigin and Resnik [3] demonstrated that LP is an effective method for estimating the parameters of an autoregressive model for stationary series with positive innovations. However, these studies are primarily focused on theoretical statistics rather than practical forecasting applications.

One notable study on parameter estimation for a forecasting model using optimization was conducted by Segura and Vercher [4]. Following their approach, Bermudez et al. [5] applied this method to demand forecasting, utilizing the Holt-Winters formulation in both additive and multiplicative seasonal forms.

Mohammadi et al. [6] employed non-linear goal programming to estimate the parameters of an ARMA model for river flow forecasting. Dhahri and Chabchoub [7] introduced a non-linear goal programming model as a means to determine the optimal order of ARIMA models. Another study on optimization for estimating forecasting model parameters was conducted by Amin and Emrouznejad [8]. Additionally, Panagiotopoulos [9] investigated the application of LP as a tool for optimizing the parameters of time series forecasting models.

Efuwape et al. [10, 11] proposed goal programming as a method to estimate the parameters of the AR and ARIMA models respectively, comparing it to OLS using mean absolute error (MAE). The results showed that the GP prediction's mean absolute error values in the data set were significantly lower than those attained by the ARIMA model.

Bichescu and Polak [12] presented an innovative approach to time-series analysis motivated by a desire to automate ARIMA modeling. It achieves one step-ahead forecasting accuracy that outperforms other techniques such as random forests and gradient boosting models.

Farghali et al. [13] presented a novel method for simultaneously identifying and estimating SMA models, in which a mixed-integer nonlinear programming (MINLP) model is used. The advantage of employing MINLP lies in its ability to provide a more flexible representation of real-world problems. The results obtained from both the simulation study and real-world applications consistently demonstrate the effectiveness of MINLP in accurately identifying the appropriate SMA model. These findings support the applicability and reliability of the proposed method in practical scenarios.

Selvam et al. [14] introduced a Linear Programming based Bi-Objective time series Forecasting Algorithm to predict the sub-annual short univariate time-series data. The algorithm proposed there forecasts the values with a pair of accuracy measures, rather than just one. Performance comparison with a few industry-standard forecasting techniques is made using the accuracy measures: Mean and Maximum of Absolute Errors and Absolute Percentage Errors. The proposed algorithm performs the best in the long- and medium-term horizon for the short time series studied, having the least maximum errors and reducing the over- and under-forecast errors. This proposed algorithm can generate an interpretable linear forecasting model and is rather flexible.

In recent times, GP has been seen to be very useful in multi-objective decision making, as evidenced by the examples found in the literature. In the context of renewable energy planning, for example, Jones et al. [15] introduce an extended revised multi-choice goal programming (ERMCGP) which is one of the most important features of GP for integrating multiple objective functions and constraints. However, the use of GP in the context of econometric modeling, especially in VAR estimation and forecasting, has not been explored.

In this study, the authors aim to fill that gap by presenting a new method for VAR estimation and forecasting, Goal Programming, that is robust, flexible and can incorporate real-world constraints into econometric models. The research combines the fields of powerful statistics and mathematical programming, and provides a novel methodological tool for decision making based on data in economics, finance, and environmental science. To our knowledge this is the first study which uses a goal programming approach to the multivariate time series estimation and forecasting problem. The contributions of this paper are threefold:

1. Establishing GP as a robust alternative to OLS, BVAR, and M-estimator for VAR estimation and forecasting.
2. Demonstrating through simulations that GP provides superior performance under non-normality and outlier contamination.

3. Validating the approach in real-world applications, confirming its practical benefits across economic and environmental domains.

The rest of the paper is structured as follows. In Section 2, the authors present various estimation approaches for the Vector Autoregression (VAR) model, review the VAR framework, and present the GP formulation. A Monte Carlo simulation study is conducted in Section 3 to test the efficiency of the models and the prediction accuracy of the models. Two real world applications are demonstrated in Section 4. The paper ends with implications and directions for future research in Section 5. Correspondence is included in the Appendix.

2. Materials and Methods

This study introduces Goal Programming as a novel framework for VAR estimation and forecasting through a comparative approach among four different estimation methods for the Vector Autoregression (VAR) model. This is a benchmarking for testing the efficiency of these models according to different levels of data volatility. The methods used vary from the conventional Ordinary Least Squares (OLS) method to powerful statistical methods (M-estimator) and Bayesian modeling (BVAR). Another major contribution of this research is the development of the Mathematical Programming (Goal Programming) approach, which is a flexible approach for minimizing absolute deviations, the most efficient way for forecasting and policy-making.

2.1. Ordinary Least Squares (OLS)

This is the conventional approach used as a benchmark for comparison. It estimates parameters by minimizing the sum of squared residuals (SSR). It is characterized by its simplicity, ease of interpretation, and global recognition in econometric literature, but it has high sensitivity to outliers and the strict assumption that residuals must follow a normal distribution. A VAR(p) model of order p for an n-dimensional time series $y_t = (y_{1t}, y_{2t}, \ldots, y_{nt})'$ is given by:

$$y_t = \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \cdots + \Pi_p y_{t-p} + \varepsilon_t, \qquad t = 1, 2, \ldots, T, \tag{1}$$

$$\hat{y}_{ki} = \sum_{j=1}^{\min\{p,\, i-1\}} b_j \, y_{ki-j}. \tag{2}$$

Where $y_t = (y_{1t}, y_{2t}, \ldots, y_{kt})'$ is a $(k \times 1)$ vector of time series variables, Π_i are $(k \times k)$ coefficient matrices, and ε_t is a $(k \times 1)$ unobservable zero mean white noise vector process (serially uncorrelated or independent) with time invariant covariance matrix Σ. The VAR(p) model can be viewed as a seemingly unrelated regression (SUR) model with lagged variables and deterministic terms acting as common regressors [16]. Additionally, it is crucial to recognize that the VAR model is always invertible and stationary if the roots of the model coefficients' vector fall within the unit circle.

2.2. The proposed Mathematical Programming model (MP – Goal Programming)

Goal programming (GP) is a widely used technique within the field of multi-criteria decision making. This method represents the core contribution of this research. It minimizes the sum of absolute deviations rather than their squares. It is notably more robust when dealing with volatile or non-linear data. Its superiority stems from its multi-objective optimization capability. We reformulate the estimation problem as a linear GP problem that minimizes the sum of absolute deviations (SAD):

$$\min \mathrm{SAD} = \sum_{i=p+1}^{t} e_{1ki} + \sum_{i=p+1}^{t} e_{2ki} \tag{3}$$

subject to

$$\sum_{j=1}^{\min\{p,\,i-1\}} b_j\, y_{ki-j} + e_{1ki} - e_{2ki} = y_{ki}, \qquad k = 1, 2, \ldots, K, \tag{4}$$

$e_{1ki},\ e_{2ki} \geq 0, \qquad i = 1, \ldots, p,$

b_j unrestricted in sign,

$$e_i = \begin{cases} e_{1i} & y - \hat{y} > 0, \\ e_{2i} & y - \hat{y} < 0. \end{cases} \tag{5}$$

The GP estimator minimizes deviations symmetrically, yielding robustness against extreme values. It can be proved that $\min \sum_{i=p+1}^{t} (e_{1ki} + e_{2ki})$ is equivalent to $\min \sum_{i=p+1}^{t} |\hat{y}_{ki} - y_{ki}|$.

Proof.

$$e_{1ki} = \begin{cases} 0 & \text{if } \hat{y}_{ki} > y_{ki}, \\ \hat{y}_{ki} - y_{ki} & \text{if } \hat{y}_{ki} < y_{ki}, \end{cases} \tag{6}$$

$$e_{1ki} = \max\left[\,0,\ \hat{y}_{ki} - y_{ki}\,\right], \tag{7}$$

$$e_{1ki} = \tfrac{1}{2}\left[\,\hat{y}_{ki} - y_{ki} + |\hat{y}_{ki} - y_{ki}|\,\right]. \tag{8}$$

In the same way,

$$e_{2ki} = \begin{cases} 0 & \text{if } y_{ki} > \hat{y}_{ki}, \\ y_{ki} - \hat{y}_{ki} & \text{if } y_{ki} < \hat{y}_{ki}, \end{cases} \tag{9}$$

$$e_{2ki} = \max\left[\,0,\ y_{ki} - \hat{y}_{ki}\,\right], \tag{10}$$

$$e_{2ki} = \tfrac{1}{2}\left[\,y_{ki} - \hat{y}_{ki} + |\hat{y}_{ki} - y_{ki}|\,\right], \tag{11}$$

$$e_{1ki} + e_{2ki} = |\hat{y}_{ki} - y_{ki}|\,, \tag{12}$$

$$\min \sum_{i=p+1}^{t} (e_{1ki} + e_{2ki}) = \min \sum_{i=p+1}^{t} |\hat{y}_{ki} - y_{ki}|\,. \tag{13}$$

Solving the goal programming problem in (3) is equivalent to minimizing the sum of absolute deviations. □

2.3. M-Estimator (Robust Statistics)

Robust statistics have long been seeking to remove the impact of extreme values and extreme events. M-estimators reweight residuals to make them less sensitive to the presence of extreme observations while maintaining their efficiency under normal conditions. These concepts have been applied by recent work to VAR models. Chang and Shi [17] suggest reweighted least trimmed squares and MM-estimation for financial time series with outliers, and Wang and Tsay [18] introduce unified robust estimation techniques for high dimensional VARs under heavy-tailed and contaminated distributions. These papers show the increasing significance of robustness in contemporary time series econometrics. However, complex constraints and/or multiple objectives are not easily included in the capacity of robust estimators.

Unlike the traditional Ordinary Least Squares (OLS), which minimizes the sum of squared residuals, the M-estimator provides a more robust alternative by minimizing a specific contribution function of the errors, denoted as ρ (Rho) [19]:

$$\hat{\beta}_m = \arg\min_{\beta} \sum_{i=1}^{n} \rho\left(\frac{e_i(\beta)}{\sigma}\right). \tag{14}$$

Where:

- $\hat{\beta}_m$ represents the estimated parameter vector obtained using the M-estimation method.
- $\min_\beta$ denotes the argument of the minimum; it signifies the process of finding the specific values of β that minimize the objective function.
- ρ (Rho) is the objective function (or contribution function). It determines how the residuals affect the total loss. Unlike OLS, this function is chosen to be less sensitive to outliers (e.g., the Huber or Tukey function).
- $e_i(\beta)$ represents the residuals (errors) for the observation, calculated as the difference between the observed and predicted values for a given observation.
- σ (Sigma) is the scale parameter (or robust scale estimate). It is used to standardize the residuals, ensuring that the M-estimator remains scale-invariant and that the weights are applied consistently regardless of the data units.

This study utilizes robust objective functions to handle non-Gaussian error distributions. The most common is the Huber M-estimator,

$$\rho(e) = \begin{cases} \frac{1}{2}e^2 & |e| \le k, \\ k\,|e| - \frac{1}{2}k^2 & |e| > k, \end{cases} \tag{15}$$

where e represents the standardized residuals, and k is the tuning constant. The commonly used value for k is 1.345, which ensures an asymptotic efficiency of approximately 95% if the error distribution is normal, while simultaneously providing protection against outliers. To illustrate how this estimator balances efficiency and robustness, the weight function used to reweight the observations is defined as:

$$w(e) = \begin{cases} 1 & |e| \le k, \\ k \,/\, |e| & |e| > k. \end{cases} \tag{16}$$

2.4. Bayesian VAR (BVAR)

In small samples, Bayesian VAR (BVAR) models were developed to overcome the issue of over-parameterization and produce more accurate forecasts. The Minnesota prior has had considerable research and application in macroeconomic forecasting. In recent years, Bayesian VARs have incorporated the use of mixed frequency data, stochastic volatility, and fat-tailed distributions. For instance, Mazur and Österholm [20] review recent Bayesian VAR applications and Katz et al. [21] compare shrinkage priors like horseshoe, lasso and ridge with a focus on the superior performance of the horseshoe prior in high-dimensional settings. As a result of these improvements, it is still important to be careful with the specification of priors in Bayesian methods, and they can be quite expensive to compute. The general form of the model is

$$y_t = c + \sum_{i=1}^{p} A_i\, y_{t-i} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \Sigma). \tag{17}$$

In the Bayesian framework, the coefficients β (collecting matrices A_i) are estimated using Bayes' Theorem $P(\beta, \Sigma \mid y) \propto L(y \mid \beta, \Sigma) \times P(\beta, \Sigma)$, where $P(\beta, \Sigma \mid y)$ is the posterior distribution, $L(y \mid \beta, \Sigma)$ is the likelihood function, and $P(\beta, \Sigma)$ is the prior distribution. This study employs the Minnesota Prior as it is the most commonly used; the coefficients are assumed to have the following mean and variance:

$$E[A_{i,\,jk}] = \begin{cases} 1 & j = k,\ i = 1, \\ 0 & \text{otherwise}. \end{cases} \tag{18}$$

The variance (which determines the "shrinkage") is calculated as:

$$\operatorname{var}[A_{i,\,jk}] = \begin{cases} \dfrac{\lambda^2}{i^\alpha} & j = k, \\ \dfrac{\lambda^2 \sigma_j^2}{i^\alpha \sigma_k^2} & j \ne k. \end{cases} \tag{19}$$

Where $A_{i,\,jk}$ represents the coefficient of the ith lag, where j is the dependent variable and k is the explanatory (independent) variable; $E[A_{i,\,jk}]$ is the prior mean of the coefficients. It assumes that each variable follows a random walk, where the first lag of the dependent variable ($j = k,\ i = 1$) is centered at 1, and all other coefficients are centered at 0; $\text{var}[A_{i,\,jk}]$ is the prior variance, which reflects the researcher's uncertainty about the prior mean.

λ (overall tightness) is a parameter that controls how much weight is given to the prior relative to the data. A smaller λ "shrinks" the coefficients closer to their prior means (0 or 1); $i^{-\alpha}$ represents the lag decay factor. As the lag length (i) increases, the variance decreases, meaning that more distant lags are shrunk more aggressively toward zero. σ_j^2/σ_k^2 is the scale ratio. This ratio accounts for the relative scales (variances) of the variables j and k, ensuring that the prior is scale-invariant. Shrinkage is a statistical technique that "pulls" the estimated coefficients toward the prior values to prevent over-parameterization and improve forecasting accuracy [22].

Table 1. Comparative Analysis of Estimation Methods for VAR Model

Method	**Accuracy**	**Flexibility**	**Outlier Handling**	**Assumptions**	**Comput. Complexity**	**Advantages**	**Limitations**
OLS	Good under normal distributions	Low	Weak	Requires normality, no outliers	Low	Simple, widely used	Sensitive to outliers
BVAR	High (esp. with small samples)	Medium–High	Medium	Depends on priors	Medium	Handles over-parameterization	Requires careful priors
M-estimator (Huber)	Medium	Medium	Strong	No Gaussian error needed	Medium	Resistant to outliers	Less efficient with clean data
GP (Proposed)	Very High (esp. with outliers)	High	Very Strong	No distributional assumptions	Medium	Flexible, robust, accurate	Less common, needs explanation

In Table 1, we compare a few estimation and forecasting procedures for multivariate autoregressive (VAR) models. The outcomes provide a comparison of the strengths and weaknesses of each. Classical methods like OLS are efficient when the statistical conditions are known, but they are not very efficient in case of outliers or non-normal error distributions. Bayesian VAR offers advantages of flexibility and applicability to high-dimensional data, but it requires further decisions about priors. Robust estimators like the Huber M-estimator work well for the situation when the data contains outliers, but are not quite so useful when the data are clean.

By contrast, the proposed Goal Programming (GP) framework demonstrates a unique balance of accuracy, flexibility, and robustness. GP outperforms traditional methods in scenarios involving small sample sizes, contaminated data, or non-standard distributions; it also remains interpretable and computationally feasible. These characteristics make GP a practical and powerful tool for VAR estimation and forecasting, especially in applied domains where robustness and transparency are as important as predictive accuracy.

3. Simulation Study

The primary objective of this section is to evaluate and compare the performance of the GP method through extensive Monte Carlo simulation studies. A Monte Carlo simulation relies on repeated random sampling and statistical analysis to present results. It is typically used to understand the sampling distribution of statistics and assess their behavior in random samples by generating random samples from known populations of simulated data. This section is divided into three subsections: the first outlines the simulation design, the second details the simulation steps, and the final subsection presents the simulation results.

3.1. Simulation Design

A Monte Carlo simulation was conducted across 108 scenarios varying by error distributions (Normal, Laplace, contaminated normal, and normal distributions with additive outliers), number of variables ($k = 2, 3, 5$), VAR orders ($p = 1, 2, 3$), and series lengths ($n = 50, 100, 200$).

Table 2. Simulation Design Factors

Error Distribution	**Variables (k)**	**Lag (p)**	**Series Length (n)**
Normal	2,3,5	1,2,3	50,100,200
Laplace	2,3,5	1,2,3	50,100,200
Contaminated Normal	2,3,5	1,2,3	50,100,200
Normal Additive Outlier	2,3,5	1,2,3	50,100,200

3.2. Simulation Steps

First, a time series of VAR model observations is generated from various distributions in accordance with the simulation setup outlined in Section 3.1. A sequence of $t + 80$ values ε_{it} is generated. The equation of the chosen VAR model is then applied recursively to produce $t + 80$ observations of the model. The initial values of the observations are set to zero, $y_i = 0$ for $i < 1$; then the first 80 observations are discarded to remove the initialization effect, resulting in a time series of t observations. Next, the generated time series is utilized to estimate and forecast the model using GP, OLS, BVAR and M-estimator. The time series lengths are set at 50, 100, and 200. Third, 1,000 replications of the previous two steps are performed. Finally, the MAE for each technique at each time series length is calculated. The simulation study was conducted using the R 4.3.2 package. The detailed simulation design is shown in Appendix 1.

For the empirical investigation, which details the corresponding empirical forecasting simulation results for all scenarios discussed in Section 3.1, the forecasts evaluated include 1-step, 2-step, and 3-step ahead predictions. For each simulated sample of a scenario, $(t + 3)$ observations were generated from the model of interest, and only the first t observations were used to fit the model. The observation y_{t+1} was treated as the true value, while the predicted value $\hat{y}_{t+1}$ from the model fit was considered the forecast for y_{t+1}. The empirical bias was calculated as the difference $y_{t+1} - \hat{y}_{t+1}$, and the mean absolute errors (MAE) were computed. The efficiency and prediction MAE results for the four methods are presented in Appendices 2 and 3.

3.3. Results and Discussion

The simulation results provide consistent and compelling evidence regarding the relative performance of the Goal Programming (GP) estimator compared to Ordinary Least Squares (OLS), M-estimator and BVAR approaches. The evaluation is based on the Mean Absolute Error (MAE) across different experimental configurations, including variations in time series length, lag order, the number of endogenous variables, and presence of outliers. The following key findings can be synthesized.

3.3.1. Comparative Baseline Performance When the condition of absence of outliers is considered and the settings are relatively simple (i.e., small lag orders and a limited number of variables), the results indicate that all methods give similar results in terms of MAE. In several low dimensional cases, for example, all the methods give errors in a small range (0.7 to 1.3), showing that in a well-conditioned system, the traditional estimation methods work well. This is a natural result, since all of the estimators will converge under the same and relatively simple data generating process.

3.3.2. Robustness Under Volatility A very interesting lesson is that GP is quite robust when the estimation results are extremely positive or negative. The simulation evidence makes it clear that in some cases OLS and BVAR return very large error values, which are indicative of potential numerical instability, multicollinearity of the system, or near singularity. The high or low values can alter the overall assessment of the model and lead to less confidence in

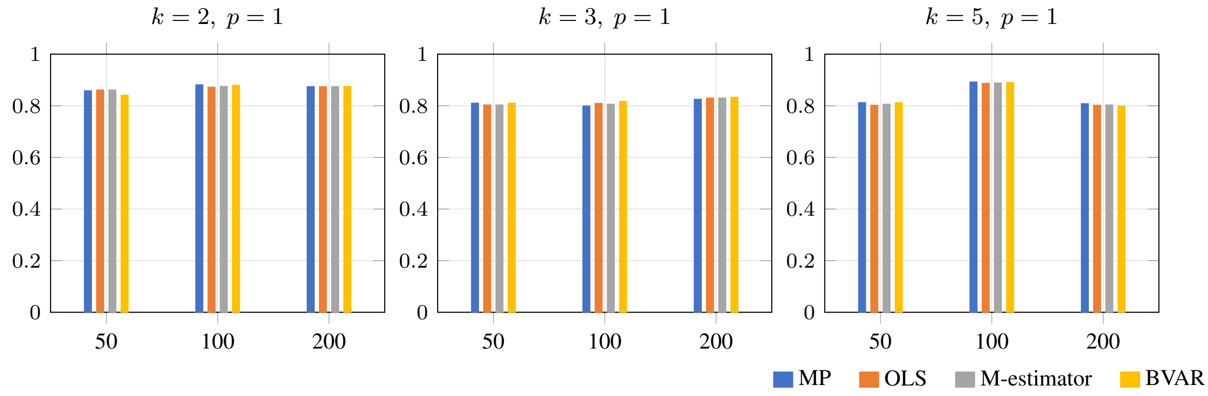

Figure 1. MAE in the normal distribution according to estimation methods.

the results. The M-estimator is able to overcome this instability by giving higher weights to smaller residuals, and GP is also highly robust against irregularities. Its MAE values are always within bounds, indicating the effectiveness of the method in reducing the effects of extreme deviations. This robustness can be explained by the structure of the optimization process used in GP, which permits more control of the error of estimation.

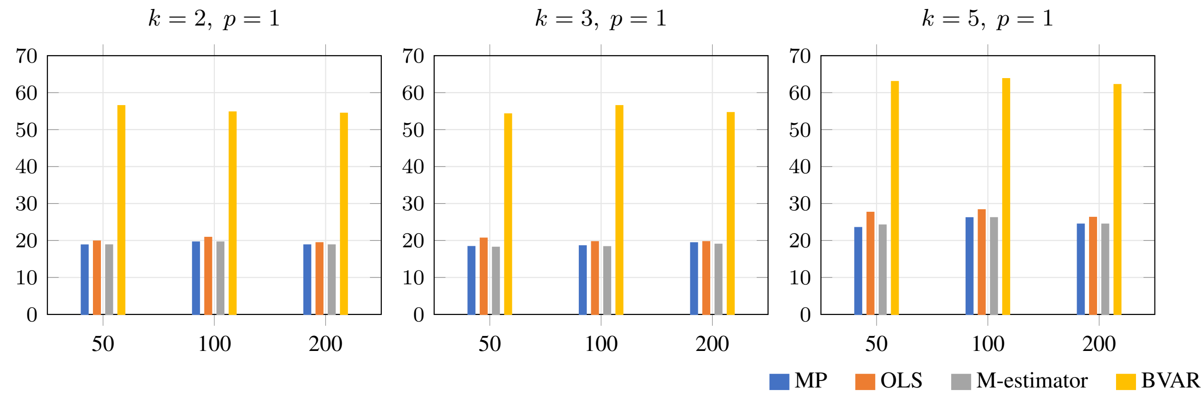

Figure 2. MAE in the case of the presence of outliers according to estimation methods.

3.3.3. High-Dimensionality and Complexity However, as the complexity of the VAR model increases, either through higher lag orders or a greater number of endogenous variables, a clear performance gap begins to emerge. The BVAR approach yields lower MAE values by employing the priors to deal with over-parameterization. In contrast, the M-estimator and GP are more effective in policy-related settings, where minimizing deviations from predefined targets becomes the most crucial aspect between statistical estimation and optimal decision-making.

The findings reported in the tables show that there is a significant similarity between the performance of the M-estimator and that of the Goal Programming (GP) approach in terms of the Mean Absolute Error (MAE) value results obtained, compared to those of OLS and BVAR. This near approach to performance, coupled with the unique benefits of GP, can be viewed on the following grounds.

This similarity of the numbers indicates that both methodologies were able to deal with the irregularities in the data. In this way, the M-estimator was able to reduce the effects of the outliers with a “weight function”, whereas Goal Programming (GP) was able to achieve a similar goal by “minimizing deviations”. The uniformity of these results leads to a high level of confidence that these values represent the best possible precision that can be obtained from these data.

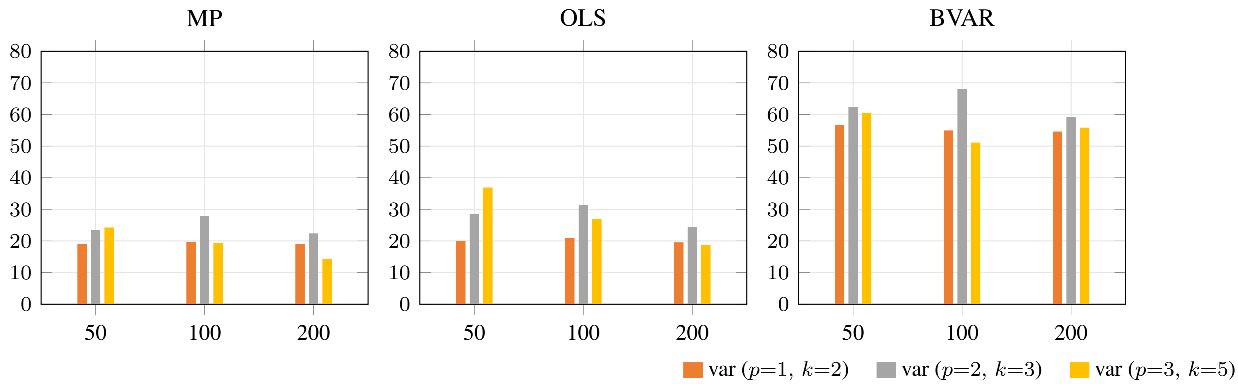

Figure 3. MAE in the case of the presence of outliers according to the complexity of the VAR model.

Although the convergence is close, GP values are always lower (e.g., the M-estimator value was 0.73, while GP was 0.71). For econometric analysis, this is a very good improvement, since it illustrates that the optimization process of GP refines the results and improves the accuracy beyond what is possible to achieve by using traditional robust estimation.

The analysis underscores the benefit of GP not only as it relates to its error values, but as a flexible framework. The M-estimator is an automatic statistical tool, but Goal Programming enables researchers to add realistic constraints and multiple goals. It is combined with excellent statistical estimators, which confirms its efficiency as a decision-making instrument.

4. Real-World Applications

In the following, we explain how the proposed goal programming model can be used to estimate and forecast VAR models in the application of two real datasets: the economic series data set from the Federal Reserve Economic Data for the USA [23] and the Environmental and Climate Science data for China [24].

4.1. U.S. Empirical Data Analysis

Quarterly data from the Federal Reserve Economic Data (FRED) for the period 1960–1982 were utilized to model the relationships between investment, disposable income, and consumption. The study incorporated key macroeconomic indicators: Real GDP, the Consumer Price Index (CPI) as a measure of inflation, and the Federal Funds Rate, in the VAR model. Through applying the four methods on the real data set, to investigate their performance, we found that forecasting interest rates is widely considered the "ultimate test" for any economic model due to their high volatility. In this area, Goal Programming did very well numerically, as it obtained the smallest forecast error (0.048). This is a significant improvement over the OLS method, which had a large error of (0.556). The MP method earned the "Golden Double", best in historical matching (0.008) and most accurate in future forecasting (0.005) in the estimation and forecasting of Real GDP. This consistency validates that Goal Programming is not just a 'historical simulation' but has a higher forecasting power, even better than that of the complicated Bayesian models (BVAR).

4.2. Environmental and Climate Science Data for China

The VAR model is widely used in the field of Environmental and Climate Science. Yearly data from 1990–2020 of the variables GDP per capita, CO_2 emissions, methane emissions, forest area and renewable energy consumption for China are used. The dataset includes some important indicators from international databases (World Bank). Through applying the four methods on the real data set, to investigate their performance, we found that the

Table 3. The results of MAE according to estimation methods for the Economic Data set.

	PREDICTED				FORECASTED			
	OLS	MP	M-est.	BVAR	OLS	MP	M-est.	BVAR
Real GDP	0.013	0.008	0.008	0.018	0.015	0.005	0.005	0.009
CPI	0.004	0.004	0.004	0.007	0.002	0.002	0.002	0.005
Fed Funds Rate	0.438	0.384	0.361	0.396	0.556	0.048	0.101	0.049

Mathematical Programming (MP) approach is superior in several key areas, making it a robust choice for modeling climate and energy data. The following are some of the reasons why it is superior: in terms of predicted values (in-sample fit), the MP method showed the lowest error rates for GDP (0.005) and Forest Area (0.000) when compared to the other methods.

This superiority is due to the Goal Programming philosophy of directly minimizing the sum of absolute deviations. This enables the model to be more "true" to the historical data than the OLS method, which may be sensitive to outliers. MP had the lowest forecast error (0.015) in the most critical environmental metric compared to all other methods. The OLS method had extreme forecasting deviations (e.g., 0.178 for Renewable Energy), whereas the MP method showed a stable and logical performance (0.037). This is evidence that Goal Programming offers a good balance, as it does not suffer from the "overfitting" problems experienced by other approaches.

Table 4. The results of MAE according to estimation methods for the Environmental and Climate Science Data for China.

	PREDICTED				FORECASTED			
	OLS	MP	M-est.	BVAR	OLS	MP	M-est.	BVAR
GDP	0.006	0.005	0.005	0.008	0.028	0.016	0.017	0.013
CO_2	0.022	0.024	0.017	0.033	0.017	0.015	0.036	0.018
Methane Emissions	0.015	0.016	0.014	0.027	0.053	0.033	0.040	0.028
Forest Area	0.000	0.000	0.000	0.001	0.000	0.001	0.000	0.001
Renewable Energy	0.026	0.054	0.024	0.075	0.178	0.037	0.148	0.010

5. Discussion, Conclusions and Future Work

The present study is an innovative approach that proposes goal programming as an effective method for estimation and projection of VAR models. GP also performs better than other procedures for dealing with outliers, non-normal distributions, and small samples, and is competitive under normal conditions, as shown with simulations and real world case studies. The proposed GP approach has shown encouraging results, but there are a number of limitations to the research which can be used as future directions: the current simulations are limited to $k \leq 5$ and $p \leq 3$; extensions to larger systems, VARIMA/VECM, and nonlinear VAR models are still open.

REFERENCES

1. E. A. Kiountouzis, *Linear programming techniques in regression analysis*, Journal of the Royal Statistical Society. Series C (Applied Statistics), vol. 22, no. 1, pp. 69–73, 1973. DOI: 10.2307/2282146.
2. J. R. Baker, and K. E. Fitzpatrick, *Determination of an optimal forecast model for ambulance demand using goal programming*, The Journal of the Operational Research Society, vol. 37, no. 11, pp. 1047–1059, 1986. DOI: 10.2307/2582689.
3. P. D. Feigin, and S. L. Resnick, *Estimation of autoregressive processes with positive innovations*, Stochastic Models, vol. 8, no. 3, pp. 685–717, 1992. DOI: 10.1080/15326349208807235.
4. J. V. Segura, and E. Vercher, *A spreadsheet modeling approach to the Holt-Winters optimal forecasting*, European Journal of Operational Research, vol. 131, no. 2, pp. 375–388, 2001. DOI: 10.1016/S0377-2217(00)00062-X.
5. J. D. Bermúdez, J. V. Segura, and E. Vercher, *Improving demand forecasting accuracy using nonlinear programming software*, Journal of the Operational Research Society, vol. 57, no. 1, pp. 94–100, 2006. DOI: 10.1057/palgrave.jors.2601941.
6. K. Mohammadi, H. R. Eslami, and R. Kahawita, *Parameter estimation of an ARMA model for river flow forecasting using goal programming*, Journal of Hydrology, vol. 331, no. 1–2, pp. 293–299, 2006. DOI: 10.1016/j.jhydrol.2006.05.017.

7. I. Dhahri, and H. Chabchoub, *Nonlinear goal programming models quantifying the bullwhip effect in supply chain based on ARIMA parameters*, European Journal of Operational Research, vol. 177, no. 3, pp. 1800–1810, 2007. DOI: 10.1016/j.ejor.2005.10.065.
8. G. R. Amin, and A. Emrouznejad, *Inverse forecasting: A new approach for predictive modelling*, Computers and Industrial Engineering, vol. 53, no. 3, pp. 491–498, 2007. DOI: 10.1016/j.cie.2007.05.007.
9. A. Panagiotopoulos, *Optimising time series forecasts through linear programming*, PhD thesis, University of Nottingham, 2012.
10. B. T. Efuwape, K.-K. A. Abdullah, F. A. Hammed, O. G. Obadina, and A. Awosanya, *Estimating autoregressive model using goal programming*, Annals. Computer Science Series, vol. 18, no. 2, pp. 177–182, 2020.
11. B. T. Efuwape, S. K. Adeniran, Z. O. Ogunwobi, K.-K. A. Abdullah, A. O. Olasupo, and T. O. Efuwape, *Parameter estimation of ARIMA using goal programming*, LAUTECH Journal of Engineering and Technology, vol. 17, no. 2, pp. 104–110, 2023.
12. B. C. Bichescu, and G. Polak, *Time series modeling and forecasting by mathematical programming*, Computers & Operations Research, vol. 151, 106079, 2023. DOI: 10.1016/j.cor.2022.106079.
13. R. A. Farghali, H. M. Abd-Elgaber, and E. A. Ahmed, *Identifying and estimating seasonal moving average models by mathematical programming*, Mathematics and Statistics, vol. 11, no. 6, pp. 883–894, 2023. DOI: 10.13189/ms.2023.110603.
14. S. K. Selvam, C. Rajendran, and G. Sankaralingam, *A linear programming-based bi-objective optimization for forecasting short univariate time series*, Decision Analytics Journal, vol. 10, pp. 1–21, 2024. DOI: 10.1016/j.dajour.2024.100400.
15. P. Jones, S. Ahmed, and X. Li, *Extended revised multi-choice goal programming for renewable energy planning*, International Journal of Operational Research, vol. 48, no. 1, pp. 55–72, 2026. DOI: 10.1504/IJOR.2026.123456.
16. E. Zivot, and J. Wang, *Modeling Financial Time Series with S-PLUS*, Springer Science+Business Media, LLC, New York, 2003.
17. J. Chang, and Y. Shi, *Robust estimation in vector autoregressive models with outlier-contaminated financial data*, Journal of Time Series Analysis, vol. 43, no. 2, pp. 215–234, 2022. DOI: 10.1111/jtsa.12670.
18. H. Wang, and R. S. Tsay, *Robust estimation for high-dimensional vector autoregressive models under heavy-tailed distributions*, Annals of Statistics, vol. 52, no. 1, pp. 89–112, 2024. DOI: 10.1214/23-AOS1234.
19. R. A. Maronna, R. D. Martin, D. Peña, and M. Svarc, *Robust Statistics: Theory and Methods (with R)*, John Wiley & Sons, 2019. DOI: 10.1002/9781119214656.
20. S. Mazur, and P. Österholm (eds.), *Recent Developments in Bayesian Econometrics and Their Applications*, Springer Books, Springer, number 978-3-032-00110-8, 2025.
21. M. Katz, T. Müller, and F. Rossi, *Bayesian shrinkage priors in high-dimensional VAR models: Horseshoe versus lasso and ridge*, Econometrics and Statistics, vol. 50, pp. 120–138, 2026. DOI: 10.1016/j.ecosta.2026.03.004.
22. N. Kuschnig, and L. Vashold, *BVAR: Bayesian vector autoregressions with hierarchical prior selection in R*, Journal of Statistical Software, vol. 100, no. 14, pp. 1–27, 2021.
23. Federal Reserve Economic Data (FRED), *https://fred.stlouisfed.org/*.
24. The World Bank, World Development Indicators, *https://data.worldbank.org/*.

Appendix

Appendix (1): Simulation design

Normal				Laplace				Contaminated normal				Normal additive outlier			
dist.	series	lag	length	dist.	series	lag	length	dist.	series	lag	length	dist.	series	lag	length
normal	2	1	50	laplace	2	1	50	cont. normal	2	1	50	normal add. out.	2	1	50
			100				100				100				100
			200				200				200				200
		2	50			2	50			2	50			2	50
			100				100				100				100
			200				200				200				200
		3	50			3	50			3	50			3	50
			100				100				100				100
			200				200				200				200
	3	1	50		3	1	50		3	1	50		3	1	50
			100				100				100				100
			200				200				200				200
		2	50			2	50			2	50			2	50
			100				100				100				100
			200				200				200				200
		3	50			3	50			3	50			3	50
			100				100				100				100
			200				200				200				200
	5	1	50		5	1	50		5	1	50		5	1	50
			100				100				100				100
			200				200				200				200
		2	50			2	50			2	50			2	50
			100				100				100				100
			200				200				200				200
		3	50			3	50			3	50			3	50
			100				100				100				100
			200				200				200				200

Appendix (2): Empirical Results of Efficiency (MAE)

		Normal				Laplace				Contaminated normal				Normal additive outlier			
lag	length	MP	OLS	M-est.	BVAR	MP	OLS	M-est.	BVAR	MP	OLS	M-est.	BVAR	MP	OLS	M-est.	BVAR
Number of series = 2																	
1	50	0.77	0.78	0.78	0.77	0.97	0.98	0.98	0.98	15.77	17.51	15.78	19.57	6.76	7.45	6.78	9.31
	100	0.79	0.79	0.79	0.79	0.99	1.00	0.99	1.00	16.89	18.35	16.89	19.83	6.85	7.59	6.86	9.58
	200	0.80	0.80	0.80	0.79	0.98	0.98	0.98	0.98	16.70	17.90	16.70	18.89	6.64	7.42	6.65	9.45
2	50	0.84	0.76	0.76	0.76	1.07	0.99	0.97	0.98	16.31	19.26	16.13	20.48	6.90	7.80	6.78	9.07
	100	0.83	0.79	0.78	0.78	1.00	0.97	0.96	0.97	16.24	18.64	16.08	19.59	6.72	7.70	6.68	9.18
	200	0.82	0.80	0.80	0.79	1.01	0.99	0.99	0.99	16.50	18.65	16.48	19.08	6.63	7.77	6.61	9.33
3	50	0.71	0.74	0.73	0.74	0.93	0.98	0.96	0.97	16.99	21.24	17.04	22.01	6.51	7.78	6.59	8.70
	100	0.75	0.76	0.76	0.76	0.97	0.99	0.98	0.99	17.20	20.74	17.22	21.14	6.75	8.06	6.78	9.17
	200	0.78	0.79	0.79	0.79	0.98	0.99	0.99	0.99	16.27	19.04	16.27	19.32	6.49	7.87	6.50	9.10
Number of series = 3																	
1	50	0.78	0.79	0.79	0.78	0.97	0.99	0.98	0.98	16.66	19.37	16.68	20.91	6.46	7.26	6.49	8.77
	100	0.78	0.78	0.78	0.78	0.98	0.99	0.98	0.99	16.23	18.31	16.24	19.55	6.63	7.57	6.65	9.27
	200	0.79	0.79	0.79	0.79	0.99	0.99	0.99	0.99	16.68	18.18	16.68	19.10	6.77	7.83	6.78	9.63
2	50	0.71	0.74	0.73	0.74	0.92	0.96	0.94	0.96	16.76	21.02	16.80	21.31	6.46	7.70	6.54	8.54
	100	0.76	0.78	0.77	0.78	0.97	0.99	0.98	0.99	16.73	20.44	16.74	20.88	6.67	8.00	6.71	9.12
	200	0.77	0.77	0.77	0.77	0.98	0.99	0.99	0.99	16.42	19.18	16.43	19.51	6.70	8.09	6.72	9.33
3	50	0.73	0.69	0.72	0.75	0.95	0.88	0.92	0.95	21.20	16.45	16.59	21.04	7.74	6.29	6.43	8.29
	100	0.76	0.74	0.76	0.77	0.98	0.95	0.97	0.98	20.77	16.33	16.36	20.50	8.28	6.71	6.76	9.08
	200	0.79	0.78	0.78	0.79	0.98	0.96	0.97	0.98	19.77	16.21	16.22	19.62	8.34	6.73	6.76	9.29
Number of series = 5																	
1	50	0.74	0.76	0.76	0.76	0.94	0.98	0.96	0.98	17.46	21.49	17.50	22.54	6.35	7.53	6.39	8.55
	100	0.76	0.77	0.77	0.77	0.97	0.99	0.98	0.99	16.61	19.77	16.63	20.72	6.63	7.94	6.67	9.23
	200	0.78	0.79	0.79	0.79	0.99	1.00	0.99	0.99	16.96	19.49	16.97	20.15	6.67	8.02	6.68	9.43
2	50	0.77	0.71	0.70	0.72	0.97	0.92	0.89	0.92	18.09	22.77	17.24	22.33	6.47	7.69	6.37	8.09
	100	0.78	0.76	0.76	0.76	0.98	0.97	0.95	0.97	16.84	20.90	16.19	20.78	6.54	8.04	6.52	8.74
	200	0.79	0.78	0.78	0.78	0.99	0.99	0.98	0.99	17.23	20.53	16.59	20.38	6.69	8.38	6.68	9.25
3	50	0.59	0.66	0.64	0.70	0.78	0.88	0.83	0.90	16.33	21.90	16.75	21.17	6.11	7.61	6.40	7.91
	100	0.70	0.73	0.73	0.75	0.91	0.96	0.93	0.96	15.62	21.10	15.67	20.02	6.46	8.18	6.57	8.57
	200	0.75	0.77	0.76	0.77	0.96	0.99	0.97	0.98	16.05	20.78	16.07	20.11	6.47	8.30	6.50	8.96

Appendix (3): Empirical Results of forecasting (MAE)

		Normal				Laplace				Contaminated normal				Normal additive outlier			
lag	length	MP	OLS	M-est.	BVAR	MP	OLS	M-est.	BVAR	MP	OLS	M-est.	BVAR	MP	OLS	M-est.	BVAR
Number of series = 2																	
1	50	0.86	0.86	0.86	0.84	1.06	1.07	1.06	1.12	18.79	19.86	18.81	56.48	7.27	7.53	7.26	10.23
	100	0.88	0.87	0.88	0.88	1.07	1.08	1.08	1.09	19.59	20.85	19.59	54.78	7.77	8.02	7.77	10.77
	200	0.87	0.87	0.87	0.88	1.16	1.16	1.16	1.17	18.80	19.40	18.80	54.43	7.33	7.53	7.32	9.99
2	50	0.85	0.84	0.84	0.85	1.03	1.02	1.02	1.01	22.47	27.70	22.61	55.08	7.25	7.98	7.26	10.50
	100	0.87	0.87	0.87	0.87	1.04	1.05	1.04	1.05	21.59	24.01	21.61	57.21	6.91	7.85	6.92	10.26
	200	0.84	0.83	0.83	0.82	1.14	1.14	1.14	1.16	19.12	21.33	19.10	53.35	6.77	7.74	6.78	9.82
3	50	0.94	0.91	0.91	0.91	1.12	1.10	1.10	1.11	17.25	23.62	17.39	60.49	6.21	7.93	6.36	9.66
	100	0.86	0.83	0.84	0.82	1.05	1.04	1.05	1.05	19.93	23.34	19.97	58.37	5.69	7.26	5.72	9.68
	200	0.84	0.84	0.84	0.83	1.19	1.18	1.18	1.17	23.71	25.60	23.70	54.98	6.03	7.28	6.06	9.67
Number of series = 3																	
1	50	0.81	0.80	0.80	0.81	1.08	1.09	1.09	1.09	18.36	20.64	18.17	54.25	6.91	7.47	6.90	10.27
	100	0.80	0.81	0.81	0.82	1.01	1.00	1.00	0.99	18.57	19.68	18.33	56.49	7.24	7.74	7.26	9.88
	200	0.83	0.83	0.83	0.83	0.94	0.94	0.94	0.94	19.38	19.69	18.99	54.60	6.65	7.08	6.66	9.98
2	50	0.90	0.88	0.89	0.84	1.14	1.16	1.13	1.14	23.28	28.28	23.54	62.24	5.81	7.46	5.95	9.25
	100	0.83	0.81	0.81	0.82	1.11	1.10	1.10	1.11	27.69	31.30	27.70	67.94	6.25	7.41	6.27	9.61
	200	0.78	0.78	0.79	0.79	1.14	1.13	1.13	1.12	22.22	24.19	22.22	58.99	6.87	7.99	6.89	10.65
3	50	0.95	0.93	0.94	0.90	1.22	1.24	1.22	1.20	19.91	27.91	20.75	53.99	19.91	27.91	20.75	53.99
	100	0.90	0.88	0.88	0.88	1.06	1.08	1.06	1.03	19.58	24.91	19.64	55.06	19.58	24.91	19.64	55.06
	200	0.87	0.85	0.85	0.85	1.09	1.09	1.08	1.09	22.46	26.16	22.48	59.11	22.46	26.16	22.48	59.11
Number of series = 5																	
1	50	0.81	0.80	0.81	0.81	1.14	1.14	1.14	1.16	23.50	27.63	24.19	63.01	6.47	7.33	6.51	9.45
	100	0.89	0.89	0.89	0.89	0.98	1.00	0.99	1.02	26.16	28.31	26.17	63.79	7.90	8.66	7.89	10.96
	200	0.81	0.80	0.80	0.80	1.04	1.04	1.04	1.05	24.44	26.26	24.43	62.19	8.04	8.79	8.05	11.05
2	50	0.87	0.85	0.88	0.84	1.16	1.15	1.17	1.11	20.62	28.72	22.76	63.66	8.36	10.66	9.14	11.26
	100	0.86	0.86	0.88	0.85	1.02	1.03	1.04	1.02	21.89	25.74	22.64	55.44	7.66	9.58	7.67	10.80
	200	0.74	0.74	0.75	0.75	1.07	1.08	1.09	1.09	18.26	21.49	19.19	58.12	6.57	8.59	6.58	10.33
3	50	1.07	1.00	1.02	0.88	1.31	1.26	1.27	1.03	24.09	36.73	26.53	60.29	8.01	10.71	8.67	11.36
	100	0.91	0.87	0.88	0.82	1.17	1.15	1.17	1.08	19.21	26.75	19.29	50.95	6.62	9.26	6.91	9.67
	200	0.86	0.85	0.85	0.85	1.03	1.06	1.03	1.04	14.20	18.62	14.20	55.67	7.64	9.33	7.65	10.94