REVISTA IBEROAMERICANA DE CONTADURÍA, ECONOMÍA Y ADMINISTRACIÓN

ISSN: 2007 - 9907

Predicción de variables económicas del sector servicios de México con modelos estadísticos clásicos y Bayesianos

Prediction of economic variables of the service sector in Mexico using classical and Bayesian statistical models

Previsão de variáveis econômicas do setor de serviços do México com modelos estatísticos clássicos e Bayesianos

Nery Alejandro Deniz Gálvez
Universidad de Colima, México
nery_deniz@ucol.mx
https://orcid.org/0000-0002-0622-3622

Osval A. Montesinos-López
Universidad de Colima, México
oamontes1@ucol.mx
https://orcid.org/0000-0002-3973-6547

Emeterio Franco-Pérez
Universidad de Colima, México
francoperez@ucol.mx
https://orcid.org/0000-0003-1191-9787

J Jesús García-Martínez
Universidad de Colima, México
jgarciam@ucol.mx
https://orcid.org/0000-0002-9148-7517

Autor para correspondencia: Emeterio Franco Pérez· Escuela de Mercadotecnia Universidad de Colima. Av. Universidad 333, Col. Las Víboras, C.P. 28040, Colima, México. Email: francoperez@ucol.mx Teléfono celular 3121141265

ISSN: 2007 - 9907

Resumen

La Encuesta Mensual de Servicios (EMS) es una actividad que realiza el Instituto Nacional de Estadística y Geografía (Inegi) con el objetivo de generar información estadística básica sobre el sector económico de los servicios en México. Considerando que se aplica mensualmente y su costo es relevante, en esta investigación se proponen modelos estadísticos (Bayesianos y clásicos) para poder predecir los indicadores de las cuatro variables agregadas que se generan a partir de los resultados de la encuesta.

Se estudiaron 42 métodos resultantes de combinar 7 modelos (tres modelos multivariados y cuatro univariados), con 6 métodos de estimación (cuatro bayesianos, uno por mínimos cuadrados y otro por máxima verosimilitud restringida). Los modelos Bayesianos permiten introducir información *a priori* con el objetivo de obtener un ajuste más preciso de los parámetros.

De los siete modelos estadísticos utilizados, el que tuvo mejor capacidad predictiva es el MP1 univariado, seguido por los modelos MP2, MP4 y MP3 multivariados; al final estuvieron los MP5, MP6 y MP7 univariados autoregresivos. De los seis métodos utilizados, el que tuvo mejor capacidad predictiva fue el BayesA, seguido por BayesB, BRR, máxima verosimilitud restringida, BayesC y mínimos cuadrados. En el caso en que predecimos para 3, 6, 12 y 18 meses, los modelos MP1 univariado, MP2, MP3 y MP4 multivariados obtuvieron la mejor capacidad predictiva utilizando los métodos BayesA, BayesB y mínimos cuadrados.

De acuerdo con los resultados obtenidos, es razonable predecir con los modelos propuestos para aquellos indicadores con una correlación de 0.4 o mayor. Con los modelos implementados se encontró que es factible predecir los resultados de la encuesta hasta para tres meses, lo que ayudaría a reducir los costos actuales en forma considerablemente.

Palabras clave: encuesta mensual, mínimos cuadrados, modelos Bayesianos, predicción, regresión lineal.

ISSN: 2007 - 9907

Abstract

The monthly service survey is an activity carried out by the National Institute of Statistics and Geography of Mexico (INEGI), with the aim of generating basic statistical information on the economic sector of services. Since the survey is applied monthly it is very expensive, therefore, in this research statistical models (Bayesian and classical) are proposed to predict the indicators of the four aggregate variables that are generated from the results of the survey.

Forty two methods resulting from combining 7 models (three multivariate models and four univariate models) were studied, with 6 estimation methods (four Bayesians, one by least squares and another by restricted maximum likelihood). Bayesian models allow us to incorporate prior information in order to obtain a more precise parameter estimates.

Model MP1 was the best in terms of prediction accuracy and this is a univariate model, followed by the multivariate models MP2, MP4 and MP3; MP5, while the worst models were MP6 and MP7 which are univariate autoregressive models. Of the six methods used, the one with the best prediction accuracy was BayesA, followed by BayesB, BRR, restricted maximum likelihood, BayesC and least squares. In the case where we predict for 3, 6, 12 and 18 months, the univariate models MP1, MP2, MP3 and MP4 multivariate models obtained the best prediction performaqnce using the BayesA, BayesB and least squares methods.

According to the results obtained, it is reasonable to predict with the proposed models for those indicators with a correlation of 0.4 or greater. With the models implemented it was found that it is feasible to predict the results of the survey for up to three months, which would help reduce current costs considerably.

Keywords: monthly survey, least squares, Bayesian models, prediction, linear regression.

REVISTA IBEROAMERICANA DE CONTADURÍA, ECONOMÍA Y ADMINISTRACIÓN
ISSN: 2007 - 9907

Resumo

A Pesquisa Mensal de Serviços (EMS) é uma atividade realizada pelo Instituto Nacional de Estatística e Geografia (INEGI) a fim de gerar informações estatísticas básicas sobre o setor de serviços económicos no México. Considerando aplicado mensalmente eo custo é relevante, esta pesquisa modelos estatísticos (Bayesian e clássicos) são propostos para prever os indicadores das quatro variáveis acrescentado gerado a partir dos resultados da pesquisa.

Métodos resultantes da combinação de 42 7 modelos (três e quatro modelos multivariados univariadas), 6 métodos de estimação (quatro Bayesiana um outro dos mínimos quadrados de probabilidade máxima restrita) foram estudados. Os modelos Bayesianos permitem inserir informações a priori para obter um ajuste mais preciso dos parâmetros.

Dos sete modelos estatísticos, que tiveram melhor capacidade preditiva é o MP1 univariada, seguida por MP2, MP3 MP4 e modelos multivariados; no final, estavam os MP5, MP6 e MP7 auto-regressivos não reagidos. Dos seis métodos, que tinham melhor capacidade de previsão foi o BAYESA, seguido por BayesB, BRR, restrito de probabilidade máxima e mínimos quadrados BayesC. No caso em que prevemos para 3, 6, 12 e 18 meses, modelos univariados MP1, MP2, MP3 e MP4 multivariada obteve a melhor capacidade preditiva usando BAYESA, BayesB e métodos de mínimos quadrados.

De acordo com os resultados obtidos, é razoável prever com os modelos propostos para aqueles indicadores que tenham uma correlação de 0,4 ou maior. Com os modelos implementados, constatou-se que é possível prever os resultados da pesquisa por até três meses, o que ajudaria a reduzir consideravelmente os custos atuais.

Palavras-chave: pesquisa mensal, mínimos quadrados, modelos bayesianos, predição, regressão linear.

Fecha Recepción: Enero 2018 **Fecha Aceptación:** Marzo 2018

Introduction

The National Institute of Statistics and Geography (INEGI) is responsible for generating basic statistical information of a socioeconomic nature that allows knowing relevant aspects of the country for the implementation of public policies and projects of social interest. The service sector in Mexico shows significant growth: in 2013 it represented 33.6% of gross domestic product (GDP), which makes it a key player for the country's economy. INEGI applies a monthly survey to a representative sample at a national level of all the establishments of the services sector with the objective of measuring and knowing the behavior of the different sub-sectors of this sector. This survey is called the Monthly Service Survey (EMS) and it began to be applied from 1993 to the present (INEGI, 2014a, p.2).

The survey represents a significant expense for the federal government, since, as the name implies, it has a monthly periodicity. For this reason, it is essential to have alternative, low-cost mechanisms that allow the application time to be spaced out. Similar studies have been done in other countries such as the Czech Republic (Bouda, 2014, p.5), Canada (Chernis and Sekkel, 2017, p.2), Sweden (Zhang, 2013, p.12) and Liechtenstein (Brunhart , 2012, p.4), where models have been implemented to predict the behavior of GDP. According to Chernis and Sekkel (2017, p.3), the management of monetary policy and economic policy in general require an assessment of the state of the economy in real time in order to reduce unnecessary costs. In this regard, Kolbachev, Kolbacheva and Salnikova (2015) carried out a study where they analyzed trends in the development of research in the areas of economics and administration, using methods that were developed in the areas of science and engineering. Also, given the phenomenon that macroeconomic indicators are released with significant delays, institutions tend to use prediction models, and even their own judgment to predict the behavior of the economy.

Some studies have analyzed the predictive capacity of the indicators of the surveys applied to study growth in private consumption; the results show that certain indicators have a predictive capacity for private consumption (Dudek, 2008, p. 19).

Surveys applied to businesses and consumers are increasingly popular in the field of macroeconomic forecasting (Lehmann and Weyh, 2014, p.2). Most prediction studies focus on measuring the predictive capacity of survey indicators for economic variables such as

GDP, industrial production and inflation. Lehmann and Weyh (2014, p.5), however, did an analysis of employment and labor market expectations.

There are other works in which dynamic Bayesian models are designed and implemented to describe the economy of countries using Monte Carlo Markov chain methods. For example, Otrok and Whiteman (1998, p.997) proposed an index-based model, however, for the scheme they used artificial data. Dynamic models have also been implemented to predict GDP. Along these lines, Porshakov, Deryugina, Ponomarenko and Sinyakov (2015, p.29) used the dynamic factor model approach for predicting Russia's GDP; One of the key results suggests that models based on few latent factors and that encompass large sets of macroeconomic variables produce quite plausible results, just as it was observed that models with a number of predictors greater than 100 obtained more accurate results. However, there are studies that suggest not including too many predictors in the factor model due to the possible noise contained in many time series. There are even studies where the regression models are compared with the time series models in the prediction of the behavior of the GDP of a whole country. As shown, Stundziene (2013, p.732) carried out an analysis of the suitability of multiple regression models and time series models (Arima, for its acronym in English) for the prediction of the GDP of Lithuania: the The obtained result showed that the multiple regression model was the most appropriate for prediction purposes.

On the other hand, in Australia a study was carried out in which a probit model was implemented to determine the capacity of financial variables, with the goal of predicting future economic events in other countries and it was concluded that these financial variables can reasonably predict the economic activity of the country (Edirisuriya, 2015, page 67). This study, however, is limited to Australian territory only.

For the aforementioned, this article proposes the use of statistical models to predict the indicators of EMS in different periods in order that, instead of applying monthly, such survey is carried out more spaced in time (for example, every three months). The statistical models that were studied are classical and Bayesian regression models under the predictive approach, which allow to predict the behavior of the variables of interest, as long as certain criteria are met. The use of these models is simple and regularly provides a reasonable predictive capacity for the problem under study.

Method

EMS is an activity that INEGI has been carrying out since 1993 with the aim of providing statistical information on the behavior of the services sector. The general objective is to generate statistical information in a timely and permanent manner on the activities of services provided by the non-financial private sector at a national level that allows to know and analyze their monthly and annual behavior. The conceptual design of the EMS is based on the document International Recommendations on Trade, Distribution and Services Statistics of the United Nations (UN). In accordance with the recommendations of the UN, it was decided not to include financial services, since there is enough statistical information on them, just as the public sector was not included (INEGI, 2014a, p.2).

In accordance with what was suggested by the UN and with the information needs that INEGI satisfies, the thematic coverage of the EMS is as follows (INEGI, 2014b, p. 11):

- Personnel employed: Includes all persons who were working in the observation unit under their direct control in the reference month, covering at least one third of the working day of the same or 15 weekly hours either of plant or eventual, receiving a payment regularly, or even without receiving it. In this section information is gathered about the monthly average of the personnel that depends on the observation unit, separated into paid and unpaid. A chapter is also opened where the number of people who worked in the observation unit is requested, but who do not depend on the company name, but are provided by another company name, as well as the staff hired for fees and commissions, and develop substantive activities in the economic unit.
- Hours worked: The hours worked are requested for the variables of paid and unpaid personnel, as well as for personnel not dependent on the company name. The variable of hours worked includes the normal waiting time at work, time not worked due to technical failures and the time of work preparation and cleaning of machinery, equipment and tools. It does not include the time paid staff stopped working because

of strikes, work stoppages, vacations, illness, permits, natural phenomena or any other cause.

- Remuneration: Refers to all payments made by the observation unit during the reference month in favor of paid personnel, both plant and casual. The variables requested in this chapter are the following: a) salaries, the amount of all payments in money entered in the payroll before any deduction, made during the reference month to compensate the normal and extraordinary work of the remunerated personnel; b) employer contributions to social security schemes, the amount of payments made by the observation unit for employer contributions to the Mexican Institute of Social Security (IMSS), to the National Workers' Housing Fund Institute (Infonavit)) and the Retirement Savings System (SAR); c) other social benefits, under this concept are considered all the additional perceptions to the salaries and wages that the unit of observation granted to the worker, be it in money, services or species; d) profits distributed to workers, payments for this concept made in the reference month, and e) payments for compensation or liquidation to staff for separation from the establishment.
- Expenses for consumption of goods and services: This is the amount of expenditures made by the observation unit during the reference month for the provision of services and activities, supply of personnel, fees or commissions, other expenses for consumption of goods and services and expenses not derived from the activities. The variables requested in this section are the following: a) Materials consumed for the provision of the service, which refers to the amount of goods actually consumed in the development of the main, secondary and auxiliary activities of the observation unit. It is important to note that the products and materials for the provision of the service, the goods purchased for resale, as well as the paints and materials used for stowage and packaging and all other consumed goods are valued at acquisition cost; b) acquisition cost, the purchase price of the goods and services acquired plus all the expenses incurred to put them in the observation unit, such as: taxes paid for the acquired goods (indirect taxes, except the value tax) aggregate (VAT), such as imports and the special on production and services), insurance, freight, storage in

transit, loading and unloading maneuvers, etc., having to deduct compensation, discounts, rebates and other concessions received; c) fuels and lubricants, amount that the establishment or company paid for the consumption of fuels and lubricants for the development of its activity in the reference month; d) rental of transport equipment, amount of the expenses incurred by the transport and courier company, for the lease (except the financial lease) or rental of transport equipment owned by third parties; e) spare parts, parts and accessories for minor repairs and current maintenance, includes payments to third parties for repair services and current maintenance of fixed assets of the company, as well as for current maintenance of fixed assets of the company, as well as for the consumption of spare parts and parts used in those repairs carried out by the same company in the fixed assets of its property, the parts and spare parts used in the production or in major repairs of the own fixed assets; f) payments for the supply of personnel, the payments made by the establishment or company to another company that provided personnel for the performance of the services or transport and courier activities. Excludes payments for the provision of surveillance services, cleaning, gardening, among others; g) fees or commissions, payments made to personnel who do not receive a base remuneration, but who perform substantive tasks, covering at least one third of the working day; h) other expenses for consumption of goods and services, include expenses such as telephone service, electricity, rents, fees to professionals who do not work exclusively in the economic unit, etc., and i) expenses not derived from activities, this chapter variable Expenses refers to those that are not due to the principal, secondary or auxiliary activity of the observation unit, but to other causes. This concept includes expenses of a financial nature, such as bank interest and commissions, dividends paid to third parties (investors), the payment of taxes and duties, fines and surcharges, loss in exchange rates, and so on.

Table 1 shows in a simplified way the aggregated variables described above, as well as the reagents of the questionnaires used for its construction.

Ricea REVISTA IBEROAMERICANA DE CONTADURÍA, ECONOMÍA Y ADMINISTRACIÓN
ISSN: 2007 - 9907

Tabla 1. Variables agregadas para la Encuesta Mensual de Servicios.

Variable Agregada	**Cuestionario Mensual**	
	Establecimiento de Servicios	**Empresas de Transportes y Mensajería**
Personal ocupado total	$H_{000}A+I_{000}A$	$H_{000}A+I_{000}A$
Remuneraciones totales	$J_{000}A+K_{610}A+K_{620}A$	$J_{000}A+K_{610}A+K_{620}A$
Gastos totales	$K_{200}A+K_{999}A$	$K_{411}A+K_{950}A+K_{999}A$
Ingresos totales	$M_{200}A$ Para el sector 53, $M_{200}A + M_{500}A$	$M_{210}A+M_{220}A+M_{230}A$

Fuente: Elaboración propia.

It is important to mention that, for each of the aggregate variables, 109 indicators are constructed and measured to give a total of 436 indicators, measured monthly in the continuous numerical scale from January 2008 to June 2016.

Proposed models

Table 2 shows the proposed models, which are multiple linear regression under the classical and Bayesian approach. The models implemented consider as a dependent variable the index constructed from the aggregate variable in question for each of the 109 indicators of the Industrial Classification System of North America. The models studied are shown below, along with the corresponding independent variables and dependent variable:

ISSN: 2007 - 9907

Tabla 2. Modelos propuestos (*MP* denota modelo propuesto)

Modelo	**Matriz diseño**	**Tipo de modelo**
MP1	$y_{ij} = \mu + A_i + M_j + AM_{ij} + \varepsilon_{ij}$	Univariado
MP2	$y_{ijk} = \mu + A_i + M_j + V_k + AM_{ij} + AV_{ik} + MV_{jk} + AMV_{ijk} + \varepsilon_{ijk}$	Multivariado
MP3	$y_{ijk} = \mu + AM_{ij} + AV_{ik} + MV_{jk} + AMV_{ijk} + \varepsilon_{ijk}$	Multivariado
MP4	$y_{ijk} = \mu + AV_{ik} + MV_{jk} + AMV_{ijk} + \varepsilon_{ijk}$	Multivariado
MP5	$y_{ij} = \mu + A_i + M_j + AM_{ij} + y_{ij-1} + \varepsilon_{ij}$	Univariado
MP6	$y_{ij} = \mu + A_i + M_j + AM_{ij} + y_{ij-1} + y_{ij-2} + \varepsilon_{ij}$	Univariado
MP7	$y_{ij} = \mu + A_i + M_j + AM_{ij} + y_{ij-1} + y_{ij-2} + y_{ij-3} + \varepsilon_{ij}$	Univariado

Fuente: Elaboración propia.

In the seven proposed models, the subscript i denotes the years and takes values from 1 to 9; the subscript j represents the month, and takes values from 1 to 12; while k represents the aggregate variable in question, and takes the values from 1 to 4 (Montesinos et al., 2017, p.4).

In the M1 model it can be seen that the factors year (Ai), month (Mj) and year-month interaction (AMij) are being considered. In the M2 model, the factors year, month, aggregate variable, year-month interaction, aggregate year-variable interaction (AVik), aggregate month-variable interaction (MVjk) and the aggregate year-month-variable interaction (AMVijk) are considered. In the M3 model, the factors interaction year-month, interaction year-variable aggregate, interaction month-variable aggregate, and interaction year-month-variable added are considered. For the M4 model, the aggregate year-variable interaction, the aggregate month-variable interaction and the aggregate year-month-variable interaction are being considered. In the M5 model, the factors year, month, year-month interaction and the response variable of the previous indicator. For the M6 model, the year, the month, the year-month interaction, the response variables of two previous indicators are considered. While for the M7 model the factors year, month, year-month interaction, and the response variables

Ricea REVISTA IBEROAMERICANA DE CONTADURÍA, ECONOMÍA Y ADMINISTRACIÓN
ISSN: 2007 - 9907

of the last three indicators are taken into account. In all cases the effects are considered as fixed effects, except for the error term (εij), which is assumed with normal distribution, with zero mean and variance σ2 (Pérez and de los Campos, 2014, p.3).

We call univariate models those that do not include as variables independent aggregate variables, and multivariate to those that take them into account. Therefore, the MP1 model is univariate, the MP2, MP3 and MP4 models are multivariate (the four aggregate variables are considered simultaneously) and the MP5, MP6 and MP7 models, given in Table 2, are univariate autoregressive.

Table 3 shows the methods used, which were the result of combining the seven models given in Table 2 with four a priori distributions, and of using two classical methods for linear adjustment:

Tabla 3. Métodos resultantes de combinar los siete modelos de la Tabla 2 con los diferentes métodos de ajuste. (*MC* denota el método de mínimos cuadrados y *MVR* denota el método de máxima verosimilitud restringida)

Método \ Modelo	**MP1**	**MP2**	**MP3**	**MP4**	**MP5**	**MP6**	**MP7**
BRR	M1	M7	M13	M19	M25	M31	M37
BAYES A	M2	M8	M14	M20	M26	M32	M38
BAYES B	M3	M9	M15	M21	M27	M33	M39
BAYES C	M4	M10	M16	M22	M28	M34	M40
MC	M5	M11	M17	M23	M29	M35	M41
MVR	M6	M12	M18	M24	M30	M36	M42

Fuente: Elaboración propia.

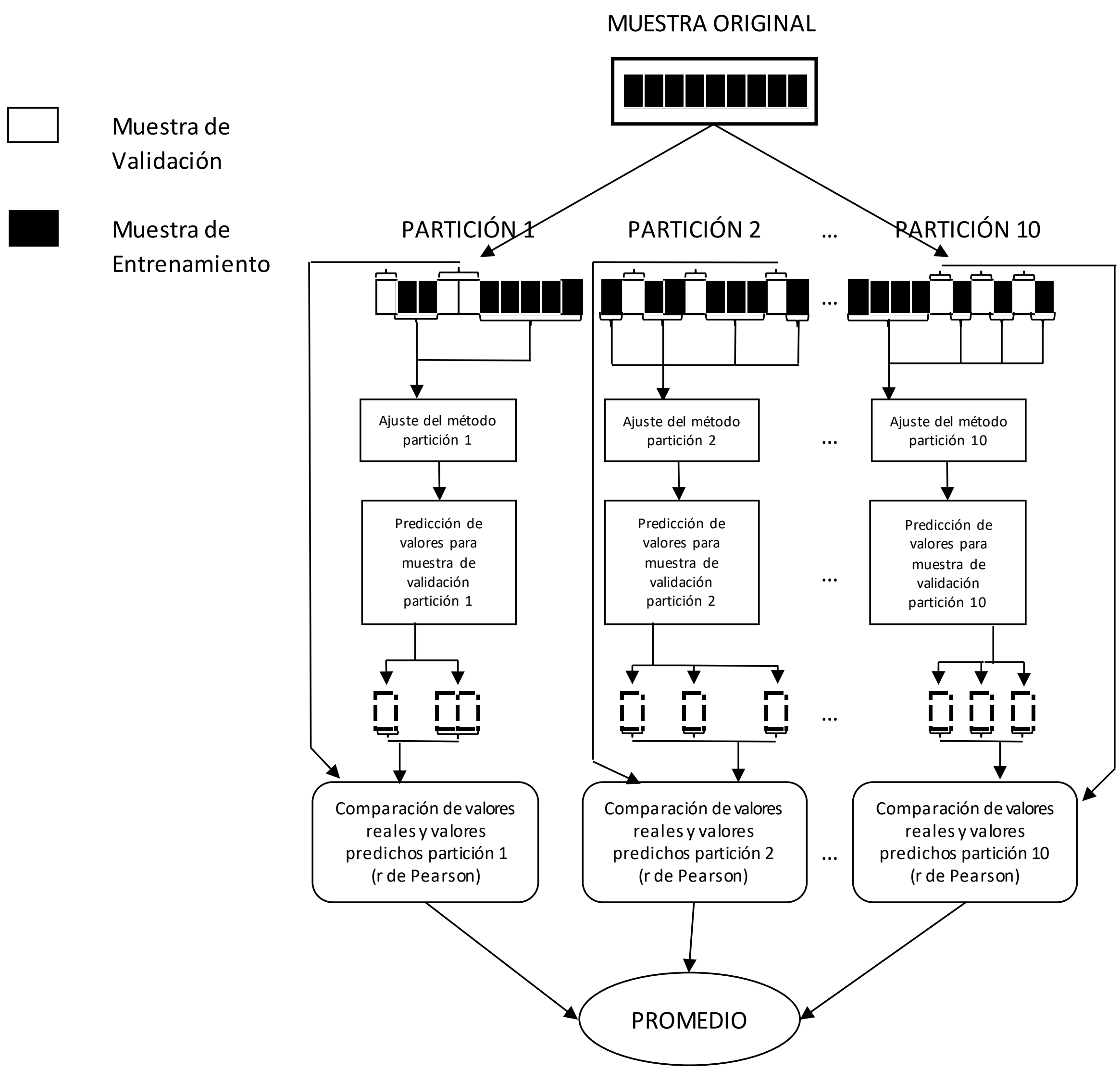
Predictiveness assessment

For the evaluation of the predictive capacity, cross-validation was used, which basically consists of dividing the original sample into two parts. One is called a training sample, and the other is called a validation sample. With the training sample, each of the methods resulting from Table 3 was adjusted, and the predictive capacity was evaluated with the validation sample. For the implementation of the cross-validation, 10 random partitions were made, where 70% of the parts were called a training sample, and the rest was for the validation sample. In each of the partitions, the full number of indicators was maintained (Montesinos et al., 2017, page 8).

The predictive capacity was evaluated (see Figure 1) using the Pearson correlation coefficient (r) between the predicted values and the observed values calculated with the validation information. The average r correlation of the 10 random partitions implemented is reported. It is important to mention that the same number of partitions was used for the training and test sets in all the statistical models that were implemented (Montesinos et al., 2017, page 8) to make fair comparisons.

REVISTA IBEROAMERICANA DE CONTADURÍA, ECONOMÍA Y ADMINISTRACIÓN
ISSN: 2007 - 9907

Figura 1. Validación cruzada



Fuente: Elaboración propia.

Results

The results of this research are presented in the following five sections: first the results corresponding to the comparison between the proposed models, methods and aggregate variables for cross validation are shown; and then the results of the comparison between the models, methods and variables added for 3, 6, 12 and 18 months later are analyzed.

Comparison between the proposed models for cross validation

It is important to remember that the indicators studied that make up each of the aggregate variables are 109; each of them was predicted using the models in Table 3. However, considering the number of variables that we intend to predict, in order to summarize the information, in Table 4 we classified each of the 109 indicators in 6 categories according to their level of predictive capacity observed for each indicator using the Pearson correlation. C1 if the correlation is less than zero, C2 if the correlation observed is between 0 and 0.2, C3 if the correlation observed is between 0.2 and 0.4, C4 if the correlation is between 0.4 and 0.6, C5 if the correlation is between 0.6 and 0.8 and C6 if the observed correlation was between 0.8 and 1. Figure 2, Figure 3, Figure 4, Figure 5 and Figure 6 show the percentage of indicators that had a correlation greater than or equal to 0.4, that is, those indicators which are in classes C4 to C6.

Tabla 4. Categorías de agrupación

Categoría	**Rango de Correlación de Pearson** (r)
C1	$r < 0.0$
C2	$0.0 \leq r < 0.2$
C3	$0.2 \leq r < 0.4$
C4	$0.4 \leq r < 0.6$
C5	$0.6 \leq r < 0.8$
C6	$0.8 \leq r < 1.0$

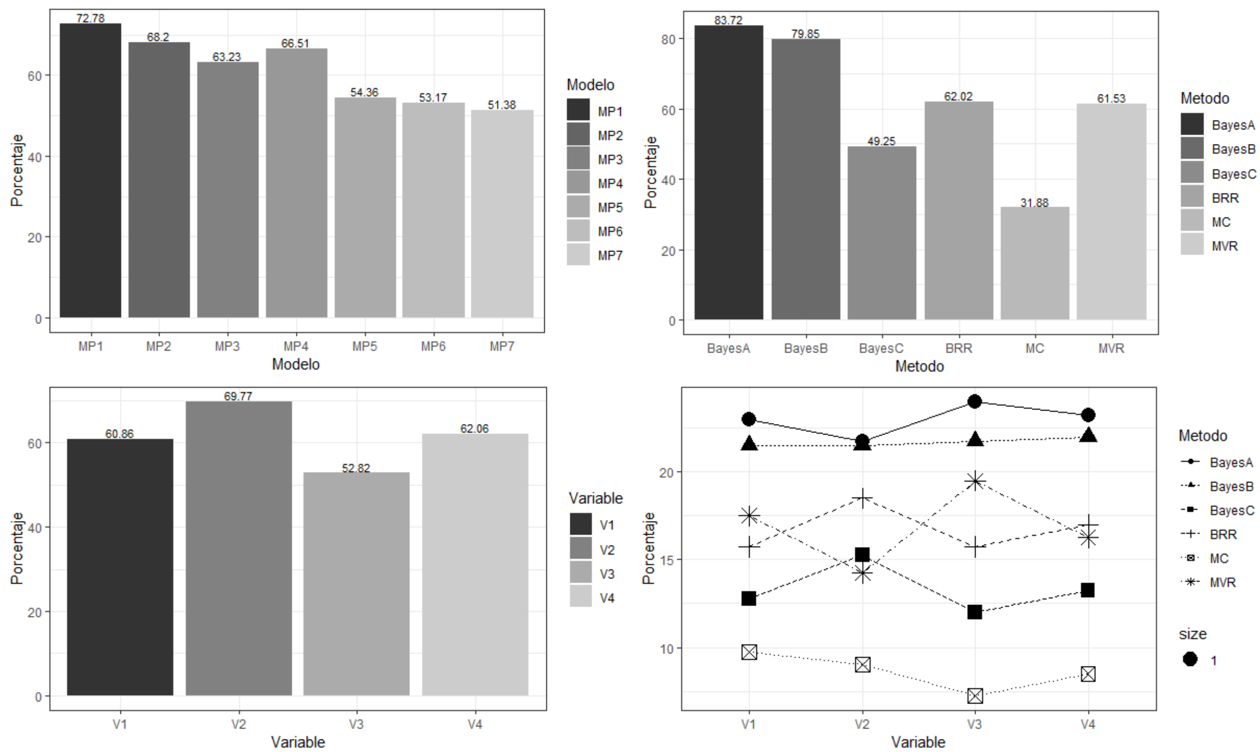
Fuente: Elaboración propia.

Figure 2 shows the results of cross validation for statistical models, methods and aggregate variables. It was observed that the MP1 model, belonging to the univariate type, gave more precise predictions, followed by the multivariate MP2, MP4 and MP3 models, MP5, MP6 and MP7 univariate autoregressive. The best regression method is BayesA, followed by BayesB, BRR, maximum restricted likelihood, BayesC and least squares. It was observed that, of the four aggregate variables, the one that is best predicted is that of total

ISSN: 2007 - 9907

expenses (V2), followed by that of total employed personnel (V4), total income (V1) and total remuneration (V3).

Figura 2. Modelos, métodos y variables agregadas para la validación cruzada.



Fuente: Elaboración propia.

Comparison between the proposed models for 3, 6, 12 and 18 months forward

Economic forecasts are by nature short-term, since in the short term the predictive capacity is better. However, this proposal to evaluate the predictive capacity for 3, 6, 12 and 18 months was to visualize how the predictive capacity was decreasing as there was greater distance in time.

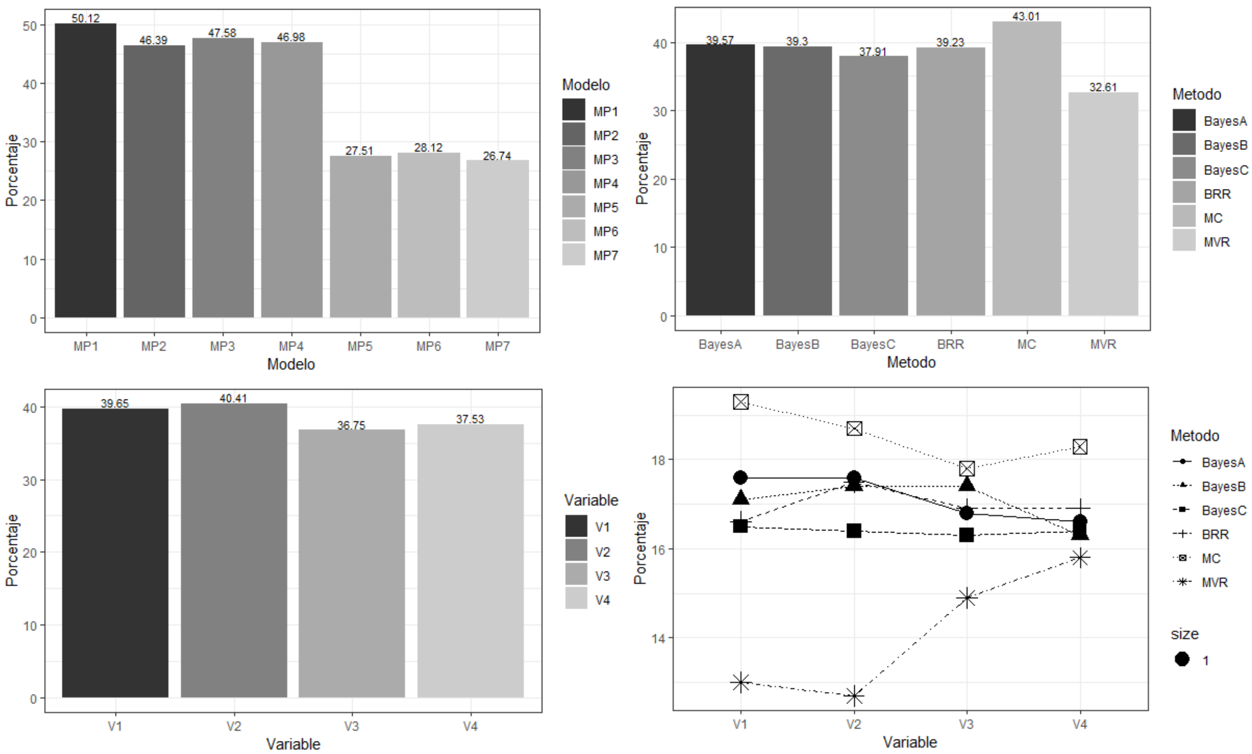
Figure 3 shows the results for the predicted values at three months, where the validation sample is made up of the values corresponding to the last three months, and the training sample is made up of the values of the remaining 99 months.

The MP1 model belonging to the univariate type yielded more accurate predictions, followed by the multivariate MP3, MP4 and MP2, MP6, MP5 and MP7 univariate autoregressive models. The method with the best predictions is MC, followed by BayesA,

ISSN: 2007 - 9907

BayesB, BRR, BayesC and maximum restricted likelihood, respectively. The best aggregate variable in terms of prediction is total expenses (V2), followed by total income (V1), total employed personnel (V4) and total remuneration (V3).

Figura 3. Modelos, métodos y variables agregadas para predecir tres meses hacia adelante.

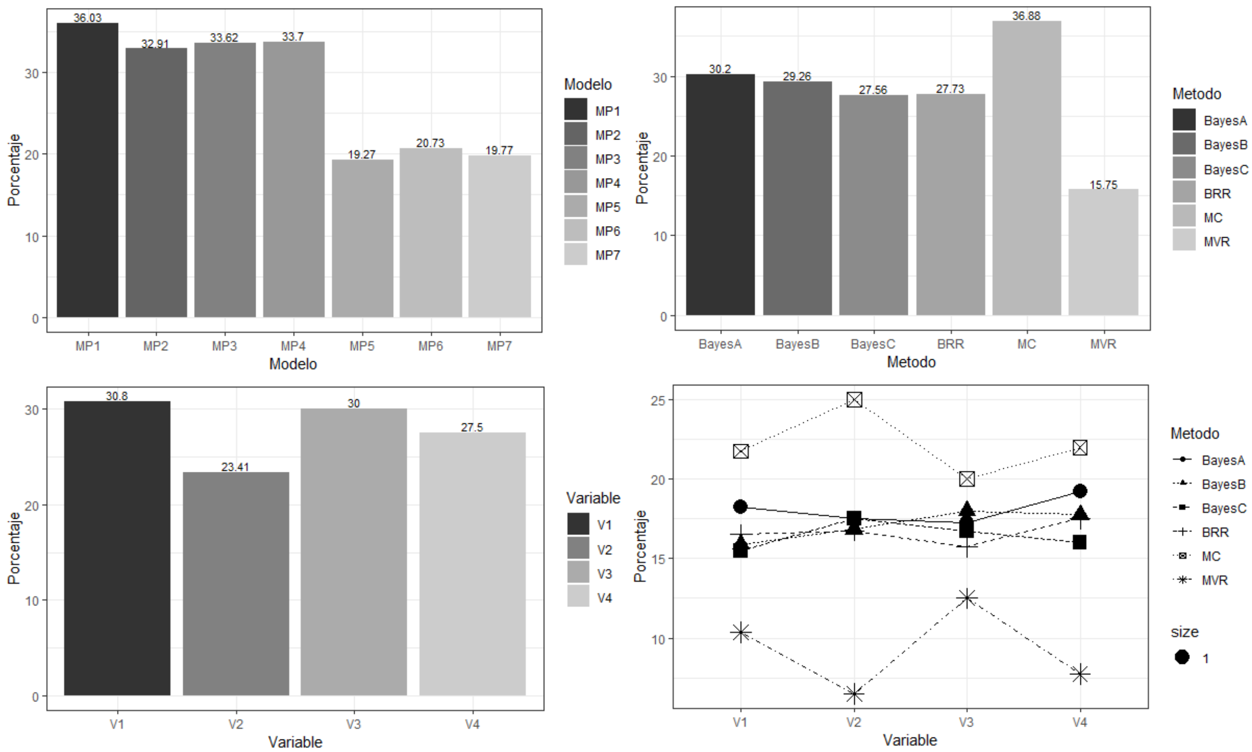


Fuente: Elaboración propia.

Figure 4 shows the results to predict six months later. In this case, the validation sample is made up of the observations of the last six months, while the training sample is made up of the remaining 96 months. It was observed that the MP1 model, belonging to the univariate type, gave more precise predictions, followed by the multivariate MP4, MP3 and MP2 models and the univariate MP6, MP7 and MP5. The method with the best results is the MC, followed by BayesA, BayesB, BRR, BayesC and MVR. It was observed that, of the four aggregate variables, the one that is best predicted is that of total income (V1), followed by total remuneration (V3), total employed personnel (V4) and total expenses (V2).

ISSN: 2007 - 9907

Figura 4. Modelos, métodos y variables agregadas para predecir seis meses hacia adelante.



Fuente: Elaboración propia.

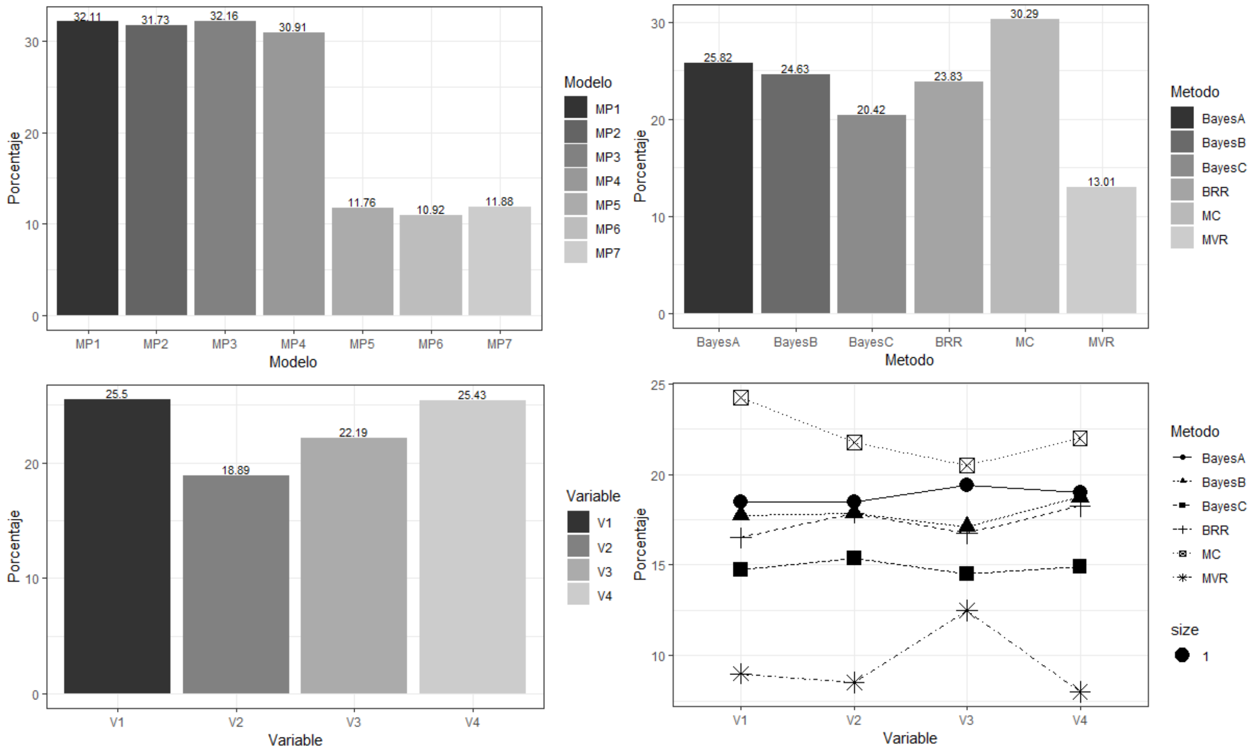
In Figure 5, on the other hand, the results are shown for the case in which it was predicted at 12 months. Here it was observed that the multivariate MP3 model yielded more precise predictions, followed by the univariate MP1, MP2 and MP4 multivariate models and the univariate autoregressive MP7, MP5 and MP6. The method that predicts better is the MC, followed by BayesA, BayesB, BRR, BayesC and MVR. The aggregate variable that gave the best predictions is total income (V1), followed by total employed personnel (V4), total remunerations (V3) and total expenses (V2).

Figure 6, finally, shows the prediction at 18 months forward. The MP2 model, belonging to the multivariate type, obtained better predictions, followed by the univariate MP1, multivariate MP4 and MP3 models and the univariate MP5, MP6 and MP7. The method with the best predictions is the MC, followed by BayesA, BayesB, BRR, BayesC and MVR. It was observed that, of the four aggregate variables, the one with the best predictions

ISSN: 2007 - 9907

is the total income (V1), followed by the total employed personnel (V4), total remunerations (V3) and total expenses (V2).

Figura 5. Modelos, métodos y variables agregadas para predecir 12 meses hacía adelante.

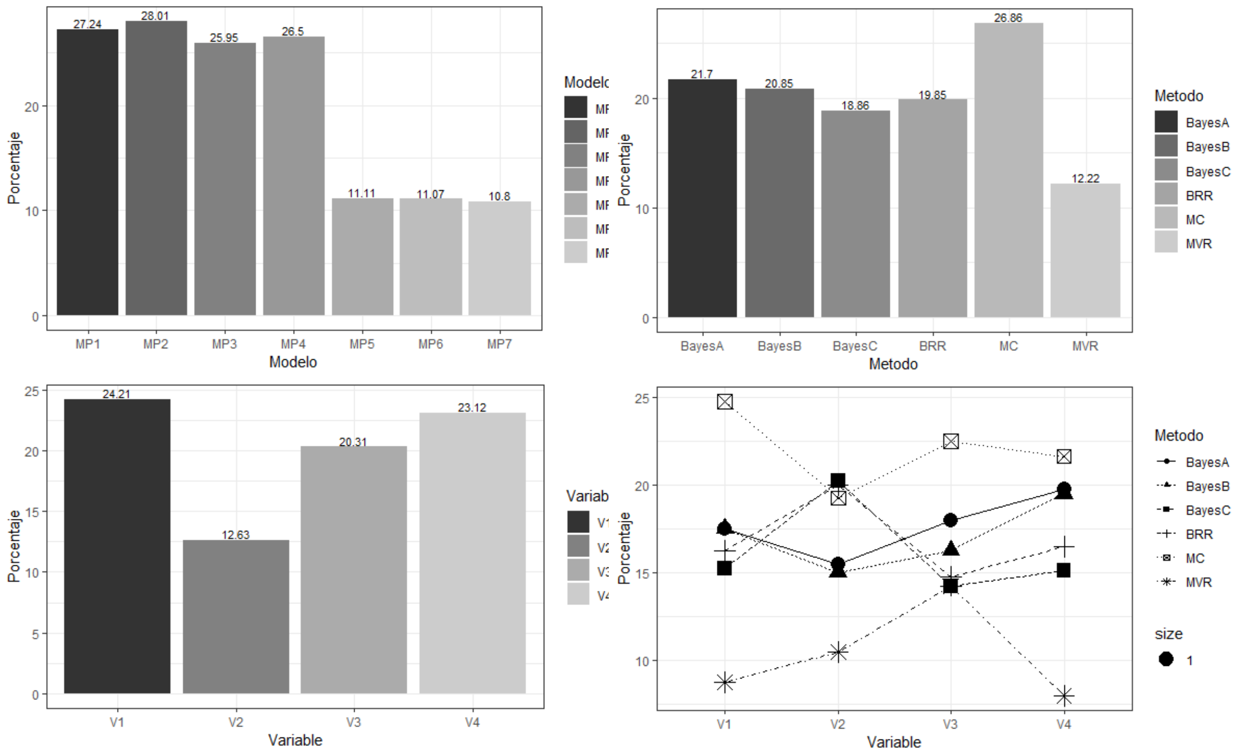


Fuente: Elaboración propia

 DOI: 10.23913/ricea.v7i14.120

Ricea REVISTA IBEROAMERICANA DE CONTADURÍA, ECONOMÍA Y ADMINISTRACIÓN
ISSN: 2007 - 9907

Figura 6. Modelos, métodos y variables agregadas para predecir 18 meses hacía adelante.



Fuente: Elaboración propia

Discussion

According to the results obtained, it can be observed that the univariate M1 model was found to have better predictive capacity than the others, followed by the multivariate M2, M4 and M3 models; at the end, the autoregressive M5, M6 and M7 models were found in that order. Of the six Bayesian methods evaluated, the one that produces the best predictions is BayesA, which has a scaled t-student density function as the marginal distribution of the effects. For computational convenience, this density is implemented as an infinite mixture of normal densities and t-student escalation (Montesinos et al., 2017, p.27). In second place was the BayesB, which is similar to BayesA, since it has a marginal distribution of the t-student effects escalated, however, it introduces an additional parameter π that represents the a priori proportion of the non-zero effects (Montesinos et al., 2017, p.28). Then there was the BRR, whose regression coefficients are assigned normal distributions with zero mean and unknown variance; then comes the maximum likelihood method, which is one of the most common methods in statistics, and for multiple linear regression assumes that the response variable has a normal distribution.

The maximum likelihood method consists of maximizing the probabilistic model taking into account all the data. First the likelihood function is formed, which is the product of all observations, and seeks to maximize it. That is, the values of the parameters that maximize the likelihood function are searched. Then there is the BayesC method, which is similar to the BRR method, with the difference that an additional parameter π is assigned, which represents the a priori proportion of the non-zero effects (Montesinos et al., 2017, p.27). Finally, the classical regression method, which corresponds to the linear least squares adjustment, which seeks to minimize the error calculated with the predicted values and the observed values. This method does not assume an a priori distribution of the data to obtain the value of the parameters; In addition, it works very well when there are normal or almost normal distributions. The added variable V2 was the one that obtained the best predictions after applying the cross-validation technique, followed by the added variables V4, V1 and V3. Said variables correspond to total expenses, total employed personnel, total income and total remunerations, respectively.

For the case in which we predict the future, it could be observed that the univariate MP1 model was the best of all for both three and six months, while the multivariate MP3 and MP2 model were superior for the case in which we predicted 12 and 18 months to future. For the case of regression methods, least squares was the best for 3, 6, 12 and 18 months, followed by BayesA and BayesB. The aggregate variable total expenditures (V2) predicted better at three months, while total income (V1) predicted better at 6, 12 and 18 months.

REVISTA IBEROAMERICANA
DE CONTADURÍA, ECONOMÍA
Y ADMINISTRACIÓN

ISSN: 2007 - 9907

Conclusions

According to the results obtained, it is reasonable to predict with the proposed models for those indicators with a correlation of 0.4 or greater (Salkind, 2004, p.81). Likewise, in the results obtained by the different statistical models studied, it could be observed that the univariate MP1 and multivariate MP2 models were found to have better predictive capacity, using the BayesA and BayesB methods. In the case where we predict for 3, 6, 12 and 18 months, the univariate MP1, MP2, MP3 and MP4 multivariate models obtained the best predictive capacity using the BayesA, BayesB and MC methods. Therefore, with the results obtained, it is feasible to predict the percentage of the following indicators: 50.12% at 3 months, 36.06% at 6 months, 32.16% at 12 months and 28.01% at 18 months of the EMS, which may represent significant savings on the part of the INEGI in carrying out the application of the survey more spaced over time. Finally, it can be concluded that the use of statistical models is very useful for prediction purposes, and serve as alternative mechanisms to reduce costs in any area of knowledge when applied correctly.

ISSN: 2007 - 9907

References

Bouda, M. (2014). Econometric Models and their ability to predict GDP growth of the Czech Republic. *AAK- Acta academica karviniensia*, *14*(3), 5-14.

Brunhart, A. (2012). Univariate and Multivariate Models to Predict Liechtenstein's GDP. *Universitat Wien*, 1-17.

Chernis, T. and Sekkel, R. (2017). A Dynamic Factor Model for Nowcasting Canadian GDP Growth. *Springer*, *53*(1), 217-234.

Dudek, S. (2010). Consumer Survey Data and short-term forecasting of households consumption expenditures in Poland. *Munich Personal RePEc Archive*, 1-29.

Edirisuriya, P. (2015). The predictive power of financial variables: new evidence from Australia. Australasian Accounting, *Business and Finance Journal*, *9*(1), 57-70.

Instituto Nacional de Estadística y Geografía (INEGI, 2014a). *Síntesis metodológica de la Encuesta Mensual de Servicios*. Aguascalientes, México: Inegi.

Instituto Nacional de Estadística y Geografía (INEGI, 2014b). *Encuesta mensual de servicios: metodología para el cálculo de los índices*. Aguascalientes, México: Inegi.

Kolbachev, E., Kolbacheva, T. and Salnikova Y. (2015). Application of Natural Science and Engineering Methods as a Trend in the Development of Economic and Management Research and Education. *Procedia Social and Behavioral Sciences*, *214*, 1000-1007.

Lehmann, R. and Weyh, A. (2014). Forecasting employment in Europe: Are survey results helpful? *Ifo Institute-Leibniz Institute for Economic Research at the University of Munich*, 1-33.

Montesinos, A., Montesinos, O., Cuevas, J., Mata, W., Burgueño, J., Mondal, S. and Crossa, J. (2017). Genomic Bayesian functional regression models with interactions for predicting wheat grain yield using hyper-spectral image data. *Plant Methods*, *13*(62).

Otrok, C. and H. Whiteman, C. (1996). Bayesian leading indicators: measuring and predicting economic conditions in Iowa. *International Economic Review*, *39*, 997-1014.

Pérez, P. and de los Campos, G. (2014). Genome-wide Regression & Prediction with the BGLR statistical. *GENETICS*, *198*(1), 483-495

Porshakov, A., Deryugina, E., Ponomarenko, A. and Sinyakov, A. (2015). Nowcasting and short-term forecasting of Russian GDP with a dynamic factor model. *BOFIT Discussion Papers*, *19*, 1-41.

Salkind, N. J. (2013). *Statistics for people who (think they) hate statistics: Excel 2010*. Washington, United States: Sage.

Stundziene, A. (2013). Prediction of Lithuanian GDP: are regression models or time series models better? *Economics and Management*, *18*(4), 721-735.

Zhang, H. (2013). *Modeling and forecasting regional GDP in Sweden* (master thesis in Microdata Analysis). Falun, Sweden: Hogskolan Dalarna.

Ricea REVISTA IBEROAMERICANA DE CONTADURÍA, ECONOMÍA Y ADMINISTRACIÓN
ISSN: 2007 - 9907

Rol de Contribución	Autor (es)
Conceptualización	Dr. Osval Antonio Montesinos López, Dr. Emeterio Franco Pérez, Maestro. Nery Alejandro Deniz Gálvez y Dr. J Jesús García-Martínez
Metodología	Dr. Osval Antonio Montesinos López, Dr. Emeterio Franco Pérez, Maestro. Nery Alejandro Deniz Gálvez y Dr. J Jesús García-Martínez
Software	Dr. Osval Antonio Montesinos López , Maestro Nery Alejandro Deniz Gálvez, Dr. Emeterio Franco Pérez y J Jesús García-Martínez
Validación	Dr. Osval Antonio Montesinos López, Maestro. Nery Alejandro Deniz Gálvez, Dr. Emeterio Franco Pérez, y J Jesús García-Martínez
Análisis Formal	Maestro. Nery Alejandro Deniz Gálvez, Dr. Osval Antonio Montesinos López, Dr. Emeterio Franco Pérez y Dr. J Jesús García-Martínez
Investigación	Maestro. Nery Alejandro Deniz Gálvez, Dr. Osval Antonio Montesinos López , Dr. Emeterio Franco Pérez y Dr. J Jesús García-Martínez
Recursos	Dr. Osval Antonio Montesinos López, Dr. Emeterio Franco Pérez, Maestro. Nery Alejandro Deniz Gálvez y Dr. J Jesús García-Martínez
Curación de datos	Maestro. Nery Alejandro Deniz Gálvez, Dr. Osval Antonio Montesinos López , Dr. Emeterio Franco Pérez y Dr. J Jesús García-Martínez
Escritura - Preparación del borrador original	Maestro. Nery Alejandro Deniz Gálvez, Dr. Osval Antonio Montesinos López , Dr. Emeterio Franco Pérez y Dr. J Jesús García-Martínez
Escritura - Revisión y edición	Dr. Osval Antonio Montesinos López Desarrollo , Dr. Emeterio Franco Pérez Maestro. Nery Alejandro Deniz Gálvez y J Jesús García-Martínez
Visualización	Dr. Osval Antonio Montesinos López, Dr. Emeterio Franco Pérez, Maestro. Nery Alejandro Deniz Gálvez y Dr. J Jesús García-Martínez
Supervisión	Dr. Osval Antonio Montesinos López y Dr. Emeterio Franco Pérez
Administración de Proyectos	Dr. Osval Antonio Montesinos López, Dr. Emeterio Franco Pérez, Maestro. Nery Alejandro Deniz Gálvez y Dr. J Jesús García-Martínez

Ricea
REVISTA IBEROAMERICANA
DE CONTADURÍA, ECONOMÍA
Y ADMINISTRACIÓN
ISSN: 2007 - 9907

Adquisición de fondos	Dr. Osval Antonio Montesinos López, Dr. Emeterio Franco Pérez, Maestro. Nery Alejandro Deniz Gálvez y Dr. J Jesús García-Martínez