

Transición de un modelo de regresión lineal múltiple predictivo, a un modelo de regresión no lineal simple explicativo con mejor nivel de predicción: Un enfoque de dinámica de sistemas

Transition from a predictive multiple linear regression model to an explanatory simple nonlinear regression model with higher level of prediction: A systems dynamics approach

Roberto Baeza-Serrato^{1}, José Antonio Vázquez-López²*

¹Dirección de Vinculación y Extensión, Instituto Tecnológico Superior del Sur de Guanajuato. CP. 38980. Guanajuato, México.

²Departamento de Ingeniería Industrial, Instituto Tecnológico de Celaya. CP. 38010. Guanajuato, México.

(Recibido el 14 de febrero de 2013. Aceptado 27 de enero de 2014)

Resumen

Uno de los supuestos principales del análisis de regresión lineal es la existencia de una relación de causalidad entre las variables analizadas, sin que el análisis de regresión lo permita demostrar. Esta investigación demuestra la causalidad entre las variables analizadas a través de la construcción y análisis de la retroalimentación entre las variables en estudio, plasmada en un diagrama causal y validado a través de simulación dinámica. Una de las principales contribuciones de ésta investigación, es la propuesta de utilizar un enfoque de dinámica de sistemas, para desarrollar un método de transición de un modelo de regresión lineal múltiple predictivo a un modelo de regresión no lineal simple explicativo, que incrementa el nivel de predicción del modelo. El error cuadrático medio (ECM) es utilizado como criterio de predicción. La validación se realizó con tres modelos de regresión lineal obtenidos experimentalmente en una empresa del sector textil, mostrando una alternativa para incrementar la fiabilidad en los modelos de predicción.

-----*Palabras clave:* Regresión lineal, dinámica de sistemas, causalidad, modelo predictivo, modelo explicativo, error cuadrático medio

* Autor de correspondencia: teléfono: + 52 + 445 + 4575055, correo electrónico: vinculacion@itsur.edu.mx (R. Baeza)

Abstract

One of the main assumptions of the linear regression analysis is the existence of a causal relationship between the variables analyzed, which the regression analysis does not demonstrate. This paper demonstrates the causality between the variables analyzed through the construction and analysis of the feedback from the variables under study, expressed in a causal diagram and validated through dynamic simulation. The major contribution of this research is the proposal of the use of the system dynamics approach to develop a method of transition from a multiple regression predictive model to a simpler nonlinear regression explanatory model, which increases the level of prediction of the model. The mean square error (MSE) is taken as a criterion for prediction. The validation in the transition model was performed with three linear regression models obtained experimentally in a textile company, showing a method for increasing the reliability of prediction models.

-----**Keywords:** Linear regression, system dynamics, causality, model predictive, explanatory model, mean square error

Introducción

Un proceso industrial con X variables independientes $[x_1, x_2, \dots, x_k]$ que afecten o puedan afectar a una variable dependiente Y , puede ser representado por modelos estadísticos de regresión que permiten explicar y predecir su comportamiento, asumiendo el supuesto de causalidad en las variables analizadas, sin poder demostrarlo; sabiendo esto, entonces, con el modelo adecuado, se pueden realizar estimaciones sobre la variable Y a partir de las variables X . Mientras que la variable Y sea predicha por las variables X , entonces el modelo para las x_k variables de la ecuación (1) representa la relación entre las variables, la cual puede ser simple (si $k=1$) o múltiple (si $k>1$), además de lineal o no lineal de acuerdo con los exponentes de las variables intervinientes en el modelo.

$$Y = \beta_0 + \beta^T X + \varepsilon \quad (1)$$

Donde:

$$\beta = [\beta_1, \beta_2, \dots, \beta_k]^T; X = [x_1, x_2, \dots, x_k]^T$$

De acuerdo con la ecuación (1), Y es la variable dependiente, x_1, x_2, \dots, x_k son las variables

independientes, β_0 es la ordenada al origen del plano de regresión, β_1 es el parámetro que indica el cambio esperado de la respuesta Y por cambio unitario en x_1 , cuando las otras x 's se mantienen constantes. β_k es el parámetro que indica el cambio esperado de la respuesta Y por cambio unitario en x_k , cuando las otras x 's se mantienen constantes.

De acuerdo con [1] las principales premisas que se han hecho hasta ahora al estudiar el análisis de regresión son las siguientes:

1. La relación entre la respuesta Y y los regresores es lineal, al menos en forma aproximada.
2. El término del error ε tiene media cero.
3. El término del error ε tiene varianza σ^2 constante.
4. Los errores no están correlacionados.
5. Los errores tienen distribución normal.

Algunos usos de los modelos de regresión lineal aplicados a diversos procesos industriales, de servicios y de salud pueden verse en [2], quienes propusieron un modelo de regresión lineal múltiple para modelar la eficiencia de un sistema

de colas de un banco, donde los factores utilizados fueron el número de servidores, la cantidad promedio de llegadas y la cantidad promedio de servicios como variables independientes, para el tiempo de espera en la fila, como variable dependiente. En [3] examinaron los factores de riesgo significativos para contraer cáncer oral, a través de un modelo de regresión lineal, donde la condición de salud de los pacientes representó la variable dependiente y La identificación de la información, tales como la edad, el género de los pacientes y la información sobre tabaquismo, consumo de alcohol y tabaco de mascar fueron utilizados como variables independientes. En [4] se propuso una transformación de los parámetros estimados de un modelo de regresión lineal, para mejorar la predictibilidad del modelo, dividiéndolos entre una constante, la cual modifica la pendiente de la ecuación lineal. En [5] propusieron controlar el pH en el proceso de teñido, como variable dependiente a través de identificar la relación entre las variables clave del proceso de pintura, obteniendo mejores resultados para colorear con un mínimo o ningún ajuste. Las variables independientes tomadas en cuenta fueron el tipo de carpeta en onzas por yarda cuadrada, el peso de la carpeta en libras, el peso del colorante en gramos y el peso del colorante como un porcentaje del peso de la carpeta. En [6] propusieron una alternativa de clasificación para un modelo de regresión lineal difuso, utilizando de base el modelo Tanaka analizando la satisfacción del cliente como variable dependiente y nueve factores utilizados como variables independientes, tales como ingresos mensuales, frecuencia de uso, la existencia de un apoyo para la espalda, la accesibilidad para respaldar el ajuste del ángulo, satisfacción por el resto del brazo, la satisfacción por la rigidez del asiento, el uso previsto de la cabecera, peso y hip-poplítea longitud. En [7] desarrollaron un modelo de regresión para la reducción del peso en un automóvil y mejorar el diseño de la resistencia a los choques en un accidente de coche frontal, en base a la cantidad de absorción de energía durante la reacción ante impactos. En [8] evaluaron los cambios dimensionales de los

tejidos lisos de fábricas de lana en el planchado al vapor, después de hacer un acondicionamiento en aire y después de un procedimiento de un ciclo de lavado, señalando que el efecto del planchado de vapor en las dimensiones es temporal e inefectivo en controlar la forma y dimensiones del tejido, así como el cambio de relajación dimensional después del ciclo de lavado depende de la densidad lineal del hilo, la opresión en el tejido y el grado de extensión impartido durante el planchado de vapor.

En cada una de estas investigaciones realizadas en diversos sectores de manufactura, de servicios y de salud, se utilizaron modelos de regresión lineal múltiple asumiendo la causalidad entre las variables analizadas, sin poder demostrarla, vislumbrando un área de oportunidad considerada en ésta investigación para demostrar tal causalidad entre las variables.

Todo análisis de regresión tiene como propósito final, hacer estimaciones precisas de la variable dependiente Y , una vez conocidos los valores de las variables independientes X ó explicar la relación existente entre dichas variables, sin controlar el crecimiento de la variable dependiente Y en una línea recta. Ver figura 1a y 1b.

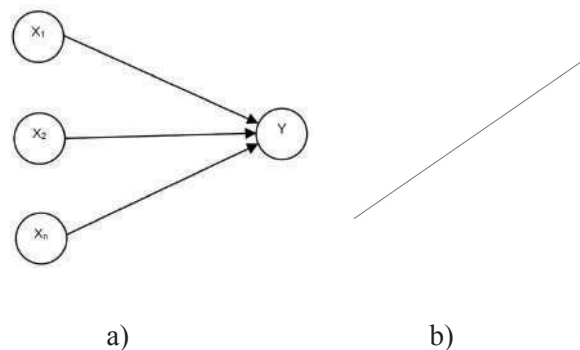


Figura 1 a) Diagrama de la regresión lineal b) Comportamiento del sistema

Esta investigación propone el uso del enfoque de dinámica de sistemas para demostrar la relación causal entre las variables independientes y dependiente a través de bucles de retroalimentación con reforzamiento positivo en X y control del

crecimiento de la variable dependiente a través de bucles de balanceo, validado con simulación dinámica, permitiendo realizar la transición de un modelo de regresión lineal múltiple predictivo a un modelo de regresión no lineal simple explicativo, el cual puede explicar y demostrar la relación entre las variables analizadas con un mejor nivel de predicción. Ver figura 2a y 2b.

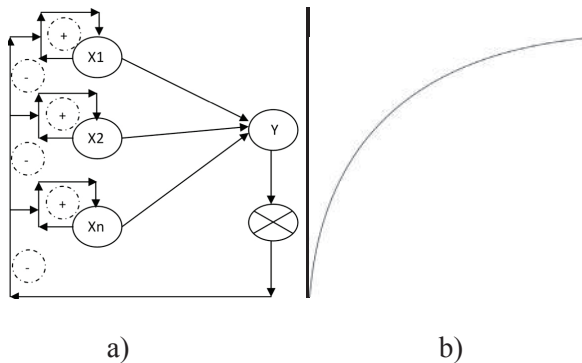


Figura 2 a) Diagrama del sistema de regresión no lineal b) Comportamiento del sistema

Este artículo propone utilizar el enfoque de dinámica de sistemas para demostrar el supuesto de causalidad asumido en el análisis de regresión lineal, mejorando el nivel de fiabilidad del modelo. Esta investigación presenta tres contribuciones principales.

Un diagrama causal que demuestra la relación de causalidad entre la variable dependiente y las variables independientes del modelo de regresión lineal, el cual se explica en la metodología.

Un diagrama de Forrester que utiliza las variables independientes como variables de estado, incrementadas a través de bucles de reforzamiento positivo, hasta alcanzar la meta a través de bucles de balanceo, obteniendo modelos no lineales con un comportamiento dinámico denominado “buscando metas”, el cual se explica en la metodología.

Un método para realizar la transición de un modelo de regresión lineal múltiple predictivo, a un modelo de regresión no lineal simple explicativo, que permite incrementar el criterio

de predicción del modelo, el cual se explica en la metodología.

Metodología

El efecto es raramente proporcional a la causa, lo que pasa localmente en un sistema frecuentemente no aplica en otros estados del sistema. La aplicación de dinámica de sistemas a la regresión lineal propuesta en esta investigación, permite hacer la transición de un modelo de regresión lineal predictivo a un modelo de regresión no lineal explicativo, que permite disminuir el ECM tomado como criterio de predicción de la relación entre las variables independientes X y la variable aleatoria Y , que representa la característica de calidad C , donde el efecto no es proporcional a la causa.

Como primer paso se hace el planteamiento de la hipótesis dinámica para explicar la causalidad entre las variables en estudio, mostrada en el diagrama causal de la figura 3, así como para explicar el comportamiento dinámico del modelo no lineal obtenido. En el diagrama causal se observa que las variables independientes actúan de manera positiva en la variable dependiente, es decir si el valor asignado en las variables independientes aumenta, el valor obtenido en la variable dependiente aumenta, de manera contraria, si los valores en las variables independientes disminuyen, el valor predicho en la variable dependiente disminuye. Para realizar la simulación del modelo, los cambios en cada una de las variables independientes son desarrollados a través de bucles de reforzamiento positivo, utilizando variables de flujo para cada una y variables auxiliares denominadas tasas de cambio. Las tasas de cambio alteran el comportamiento del modelo de regresión.

El valor de la variable de respuesta no puede ser infinito en ningún caso práctico, por lo que en esta investigación el crecimiento de la variable de respuesta está restringido hasta alcanzar el objetivo o la meta buscada. Para lograr esto es utilizado un bucle de balanceo o negativo, es decir si aumenta el valor de la

variable de respuesta, disminuye el valor de la discrepancia obtenida para alcanzar la meta y si la variable de respuesta disminuye, el valor de la discrepancia obtenida aumenta y se realiza la acción correctiva correspondiente. El diagrama causal de la figura 3 muestra la interacción entre los bucles de reforzamiento positivos con los bucles de balanceo, permitiendo encontrar un comportamiento dinámico denominado buscando metas, que sigue una distribución lognormal. La estructura de la retroalimentación presente a través de bucles de reforzamiento positivos y bucles de balanceo, permite identificar un comportamiento no lineal, que difiere del comportamiento lineal asumido de los modelos de regresión lineal.

Los elementos del diagrama causal son:

- i. Variables independientes (Factores significativos) X
- ii. Variables de flujo Fx_i
- iii. Tasas de cambio Ix_i
- iv. Variable de respuesta Y
- v. Discrepancia D
- vi. Meta M

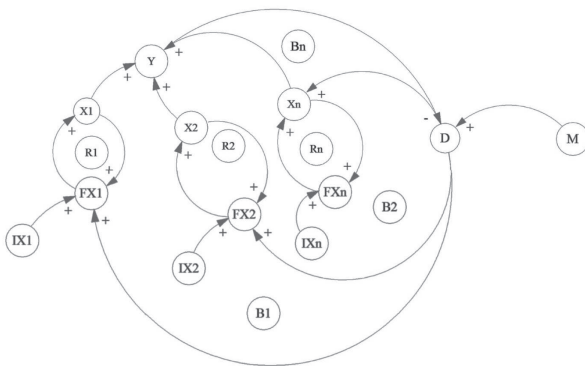


Figura 3 Hipótesis dinámica del modelo de regresión no lineal simple explicativo

Una vez planteada la hipótesis dinámica en un diagrama causal, como segundo paso se hace la construcción del diagrama de Forrester para realizar las simulaciones dinámicas. Cada una

de las variables independientes del modelo de regresión lineal múltiple predictivo, corresponde a una variable de estado o reservorio en un diagrama de Forrester. Ver figura 4. En ellas se acumula el crecimiento a realizar en cada una de las variables, lo cual genera el desequilibrio dinámico del sistema, transformando el modelo de regresión lineal múltiple predictivo a un modelo de regresión no lineal simple explicativo. Los flujos están determinados por la tasa de cambio establecida para cada variable independiente y la acción correctiva a realizar para obtener la meta buscada. Este enfoque posibilita asignar diversos valores en la tasa de cambio, lo cual permite obtener distintos modelos no lineales, calcular el ECM y R^2 para cada uno, hasta obtener el mejor modelo considerando el criterio de predicción y explicar el comportamiento de la causalidad entre las variables en estudio. Por cada valor asignado en la tasa de crecimiento se obtiene un modelo no lineal.

Para la validación del modelo dinámico, se utilizó el software Stella 9.02, realizando la transición de tres modelos de regresión lineales, obtenidos en un diseño experimental en una empresa textil por [9], a modelos de regresión no lineales disminuyendo el ECM tomado como criterio de predicción.

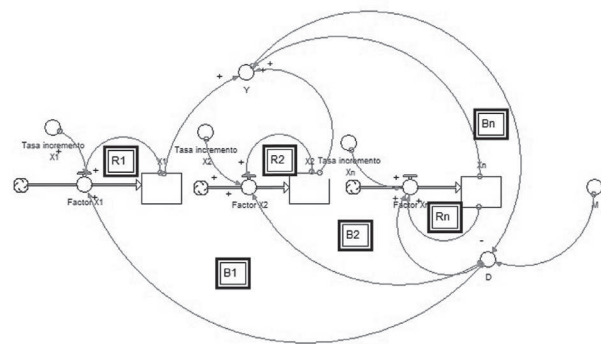


Figura 4 Diagrama de Forrester modelo de regresión no lineal simple explicativo

Los fundamentos básicos de dinámica de sistemas, permiten explicar la causalidad existente en un modelo de regresión lineal múltiple predictivo,

tales como la acumulación y la retroalimentación entre las variables en estudio. Se utilizan los elementos básicos de un sistema dinámico, tales como variables de estado, variables de flujo y variables auxiliares, describiendo la transición de las variables independientes y la variable dependiente, de los modelos estadísticos de regresión lineal múltiples predictivos, para construir el diagrama de Forrester y obtener el modelo de regresión no lineal simple explicativo.

Método de construcción diagrama de Forrester

1. Obtener un modelo de regresión lineal que relacione las variables independientes con la variable dependiente.
2. Definir cada una de las variables independientes del modelo de regresión lineal como variables de estado (**X**).
3. Definir una variable de flujo para cada variable de estado (Fx_i).
4. Definir la tasa de cambio para cada variable de flujo (Ix_i).
5. Definir cada una de las variables dependientes como variables auxiliares de respuesta (**Y**).
6. Definir cada uno de los coeficientes del modelo de regresión como variables auxiliares (**β**).
7. Definir la ordenada al origen del plano de regresión como variable auxiliar (**β_0**).
8. Definir la meta u objetivo a alcanzar como una variable auxiliar (**M**).
9. Definir la discrepancia (**D**) como una variable auxiliar.
10. Agregar la retroalimentación entre las variables de estado y las variables de flujo.
11. Agregar la retroalimentación entre las tasas de cambio y las variables de flujo.
12. Agregar la retroalimentación entre las variables de estado y la variable auxiliar de respuesta.

13. Agregar la retroalimentación entre las variables auxiliares que representan los coeficientes del modelo de regresión lineal y la ordenada al origen con la variable de respuesta.
14. Agregar la retroalimentación entre la variable de respuesta y la discrepancia.
15. Agregar la retroalimentación entre la meta y la discrepancia.
16. Agregar la retroalimentación entre la discrepancia y las variables de flujo.

Las variables de estado, los flujos y la retroalimentación presente entre las variables en estudio, son los fundamentos teóricos de los sistemas dinámicos [10]. Los valores obtenidos en cada variable de estado actúan de manera positiva, en la variable de respuesta del modelo. Los valores de la variable de respuesta son controlados con bucles de balanceo, hasta encontrar la meta buscada e interactúan en los bucles de reforzamiento positivos de cada variable de estado.

Las variables de estado o reservorios y los flujos, están basados en la metáfora hidráulica del flujo de agua que entra y sale del reservorio. La representación matemática de cada reservorio es a través de ecuaciones integrales. Las variables de estado acumulan o integran sus flujos. El flujo neto dentro de la variable de estado es la tasa de cambio de la variable de estado [10]. Ver ecuación (2).

$$\psi(t) = \int_{t_0}^t [\phi_e(s) - \phi_s(s)] ds + \psi(t_0) \quad (2)$$

Donde:

$$\phi_e = \text{Flujo de entrada}$$

$$\phi_s = \text{Flujo de salida}$$

$$\psi = \text{Reservorio}$$

$\phi_e(s)$ representa el valor del flujo de entrada en cualquier tiempo s entre el tiempo inicial t_0 y el tiempo presente t . Equivalentemente la tasa neta

de cambio de cualquier reservorio es su derivada, es decir, el flujo de entrada menos el flujo de salida, definida por la ecuación (3). [10]

$$\frac{d(\psi)}{dt} = \phi_e(t) - \phi_s(t) \quad (3)$$

En general los flujos son funciones de los reservorios y otras variables de estado o parámetros.

Los reservorios acumulan los cambios pasados realizados en cada variable independiente y representan fuentes de demora. Las acciones correctivas o los incrementos en las variables independientes sufren una demora para poder realizarse, ya que los ajustes no pueden ser en forma instantánea y requiere de un tiempo para llevarse a cabo. Las acumulaciones y las demoras crean el desequilibrio dinámico, que permite realizar la transformación del modelo de regresión lineal múltiple predictivo a un modelo de regresión no lineal simple explicativo.

Las unidades utilizadas en los reservorios son las mismas a utilizar en los flujos de entrada y de salida, siendo limitadas a periodos de tiempo determinados.

La comparación de conceptos entre la regresión lineal y el enfoque dinámico se observan en la tabla 1.

Simulación y resultados

Para realizar la validación en la transición de modelos de regresión lineales múltiples predictivos a modelos de regresión no lineales simples explicativos, se utilizaron tres modelos de regresión lineal múltiples predictivos obtenidos experimentalmente por Baeza et al. (2011). Ver ecuaciones (4-6). La simulación se hizo en el Software Stella 9.02 utilizando el diagrama de Forrester de la figura 5.

$$Y_1 = 19,0297 - 0,0859375x_1 + 0,145313x_5 - 0,201563x_1x_5 \quad (4)$$

$$R^2 = 66,5998\%; \text{ECM} = 0,099$$

$$Y_2 = 87,8813 + 1,15625x_2 - 0,71875x_3 + 1,30625x_4 \quad (5)$$

$$R^2 = 77,04\%; \text{ECM} = 2,1336$$

$$Y_3 = 0,0881719 + 0,00120313x_2 + 0,000953125x_4 + 0,000546875x_2x_4 \quad (6)$$

$$R^2 = 75,69\%; \text{ECM} = 1,489E-06$$

Tabla 1 Comparación de conceptos

| <i>Modelo de regresión lineal</i> | <i>Simbología</i> | <i>Dinámica de sistemas</i> | <i>Simbología</i> |
|--|----------------------|--|----------------------------------|
| Factores significativos | X | = Variables de estado | X |
| Variable de respuesta | Y | = Variable auxiliar de respuesta | Y |
| Coefficientes de los factores significativos | β | = Variable auxiliar coeficientes de los factores | β |
| Ordenada al origen del plano de regresión | β₀ | = Ordenada al origen del plano de regresión | β₀ |
| | | Variables de flujo | F_{x_i} |
| | | Tasas de cambio en las variables de flujo | I_{x_i} |
| | | Discrepancias con el valor objetivo | D |
| | | Meta u objetivo | M |

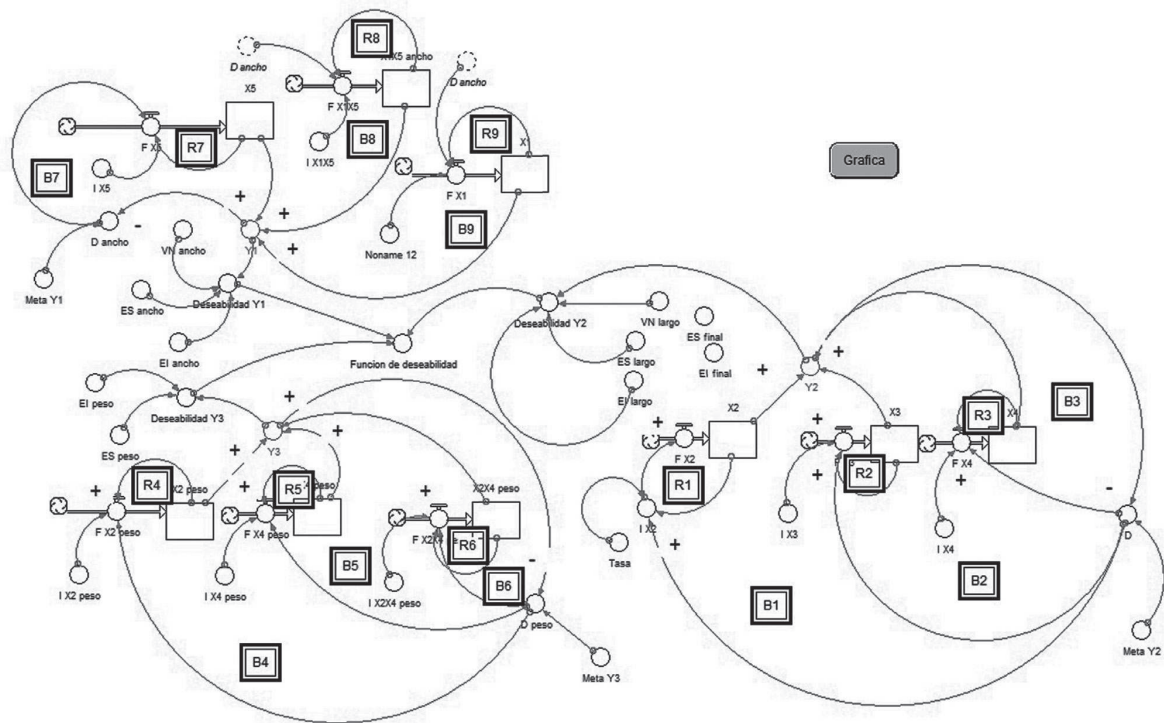


Figura 5 Diagrama de Forrester para la transición de los modelos de regresión lineal múltiples predictivos obtenidos experimentalmente por [9]

En el diagrama de Forrester de la figura 5 para cada modelo de regresión en estudio, se observan tres bucles de reforzamiento positivos y tres bucles de balanceo, en total nueve bucles positivos y nueve bucles negativos. Las variables de estado representan cada una de las variables independientes de las ecuaciones (4-6).

A continuación se muestra el desarrollo para la transición del modelo de regresión lineal múltiple predictivo Y_2 de la ecuación 5. De la misma manera se desarrolla la transición de los otros dos modelos. Las variables de estado están formuladas a través de las ecuaciones (7-9).

$$X_2(t) = \int_{t_0}^t [Factor X_2(s) - X_2(s)] ds + X_2(t_0) \quad (7)$$

$$X_3(t) = \int_{t_0}^t [Factor X_3(s) - X_3(s)] ds + X_3(t_0) \quad (8)$$

$$X_4(t) = \int_{t_0}^t [Factor X_4(s) - X_4(s)] ds + X_4(t_0) \quad (9)$$

Donde los flujos de los Factores X_2 , X_3 y X_4 se calculan a través de las ecuaciones (10-12):

$$FX_2 = \frac{d(X_2)}{dt} = IX_2 \left(\frac{DY_2}{X_2} \right) \quad (10)$$

$$FX_3 = \frac{d(X_3)}{dt} = IX_3 \left(\frac{DY_2}{X_3} \right) \quad (11)$$

$$FX_4 = \frac{d(X_4)}{dt} = IX_4 \left(\frac{DY_2}{X_4} \right) \quad (12)$$

Donde:

IX_2 = Tasa de cambio para X_2

IX_3 = Tasa de cambio para X_3

$$IX_4 = \text{Tasa de cambio para } X_4$$

$$Y = \prod_{m=1}^N x_m \quad (14)$$

$$DY_2 = \text{Discrepancia para } Y_2$$

Donde:

La discrepancia es calculada a través de la ecuación (13).

$x_1 < x_2 < \dots < x_m$, entonces Y es una variable distribuida log-normalmente. Ver ecuación (15).

$$D = Y - M \quad (13)$$

$$Y \sim \text{Log-N}(\mu, \sum_m \sigma_m^2) \quad (15)$$

Donde:

Los distintos modelos no lineales obtenidos en esta investigación para Y_2 se observan en la figura 6, los cuales siguen una distribución lognormal. Cada uno de ellos corresponde a un valor dado para la tasa de cambio (IX_i) y la meta (M) a alcanzar.

A cada una de las variables de estado para cada variable auxiliar de respuesta en estudio, se les asigna un mismo valor para la tasa de cambio, por lo que el modelo de regresión lineal múltiple predictivo, en el que cada variable de estado tienen sus propios valores, hace la transición a un modelo de regresión simple no lineal explicativo, que para cada variable de estado corresponden los mismos valores. Es decir $x_2 = x_3 = x_4$

Esta propuesta basada en un enfoque de dinámica de sistemas, utiliza variables independientes log-normalmente distribuidas con el mismo parámetro μ y permitiendo que σ varíe. Ver ecuación (14).

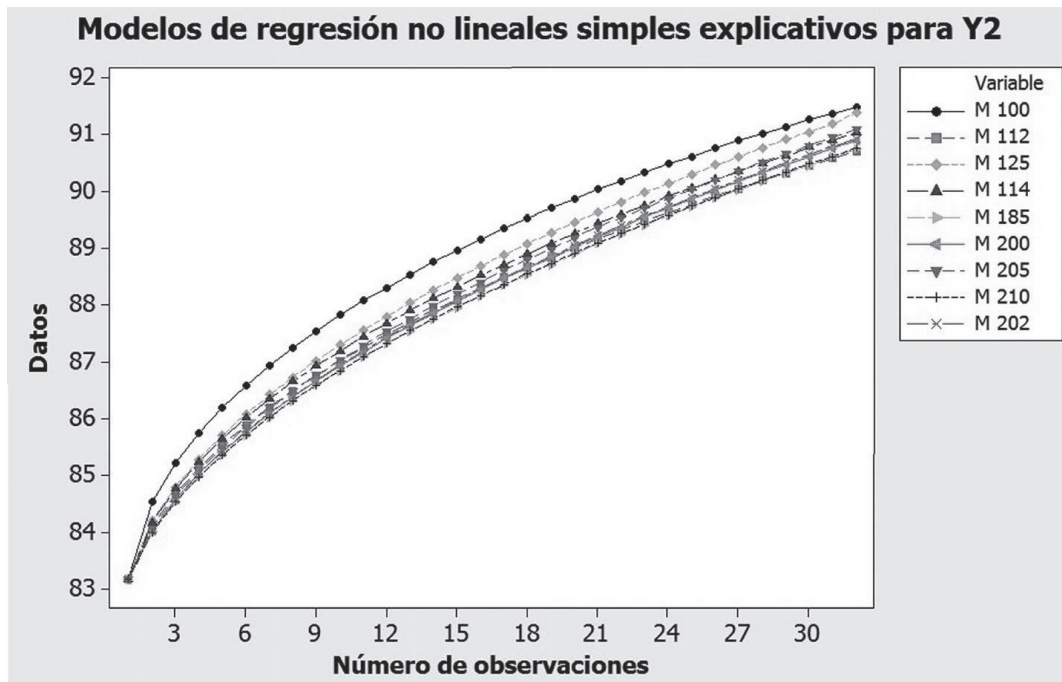


Figura 6 Comparativo de los modelos de regresión no lineales simples explicativos obtenidos para el modelo de regresión lineal múltiple predictivo Y_2

Un criterio que en general se utiliza para ejemplificar lo adecuado de un modelo de regresión ajustado es el coeficiente de determinación múltiple, el cual calcula el porcentaje de varianza justificado por las variables independientes. A medida que su valor es mayor, el ajuste de la recta a los datos es mejor, puesto que la explicación es mayor. Ver ecuaciones (16 y 17).

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (16)$$

$$R^2 = 1 - \frac{SSE}{SST} = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (17)$$

Una de las técnicas más utilizadas para la regresión lineal, es la técnica de estimación de parámetros a través de la suma de mínimos cuadrados, donde la variabilidad total es originada por la variabilidad explicada por la regresión SSR, más

la variabilidad originada por los residuales SSE. Ver ecuación (18). Esta regresión es basada en una línea recta de valores promedios.

$$SST = SSR + SSE \quad (18)$$

El coeficiente de determinación R^2 solo indica qué proporción de la variación total de la respuesta Y se explica mediante el modelo ajustado. [11]

En [12] mencionan y muestran ejemplos numéricos, validando que el criterio de explicación establecido R^2 , puede dar una idea equivocada sobre las capacidades predictivas del modelo de regresión en cuestión. En esta propuesta es calculado y analizado el indicador del ECM como criterio de predicción. Ver ecuación (19) y Tabla 2.

$$ECM = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n} \quad (19)$$

Tabla 2 Comparativo de las diversas tasas de cambio y los coeficientes de determinación para cada modelo no lineal Y_2

| Simulación | Tasa de cambio | Meta | SSR | R^2 | ECM |
|------------|----------------|------|-------------|-------------|------------|
| 1 | 0,1 | 100 | 105,941696 | 0,523340085 | 2,85563781 |
| 2 | 0,04 | 112 | 128,672345 | 0,63562694 | 0,47282037 |
| 3 | 0,03 | 125 | 158,5491433 | 0,783214969 | 0,56686989 |
| 4 | 0,04 | 114 | 142,5285725 | 0,704075148 | 0,50435758 |
| 6 | 0,01 | 185 | 133,1312143 | 0,657653253 | 0,39175676 |
| 7 | 0,09 | 200 | 139,2617837 | 0,687937578 | 0,37187497 |
| 8 | 0,09 | 205 | 147,4507501 | 0,728390153 | 0,38653556 |
| 9 | 0,008 | 210 | 134,287158 | 0,66336349 | 0,38406458 |
| 10 | 0,089 | 202 | 140,420987 | 0,69366391 | 0,37114188 |

En la tabla 2 se observa que así como se va disminuyendo el indicador de explicación, se va disminuyendo el indicador de predicción, lo cual indica que el modelo de regresión no lineal simple explicativo obtenido, es un mejor modelo de regresión no lineal simple predictivo.

El modelo de regresión obtenido experimentalmente para Y_2 obtuvo un valor del $ECM = 2,1336$ y un valor de $R^2 = 0,7704$. Ver Ecuación 5. En esta propuesta de transición se obtiene una disminución en el criterio de predicción, hasta un valor del $ECM = 0,3711$,

mejorando la predicción de los valores en estudio. La tasa de cambio para cada variable de estado es de 0.089 en la simulación 10. Ver tabla 2.

Las estimaciones del modelo más fiable de regresión no lineal simple explicativo, obtenido

en la simulación 10 con una tasa de cambio de 0.089, se observa en el comparativo de la figura 7, con las observaciones reales y las estimaciones realizadas por el modelo de regresión lineal múltiple predictivo obtenido experimentalmente.

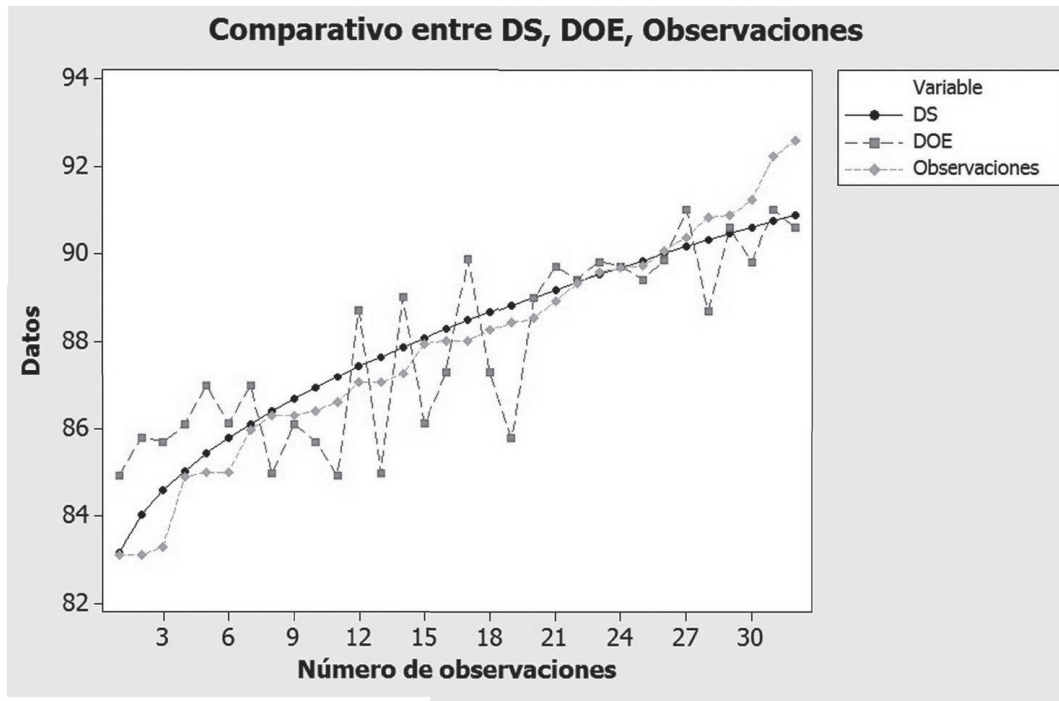


Figura 7 Comparativo entre el mejor modelo de regresión no lineal simple explicativo y las observaciones reales Y_2

En la figura 7 se observa el comparativo del modelo de regresión lineal múltiple predictivo, el modelo de regresión no lineal simple explicativo y las observaciones reales. El comportamiento del modelo de regresión no lineal simple explicativo es del tipo buscando metas, tiene un crecimiento exponencial hasta alcanzar el objetivo o meta. La tabla 3 muestra los valores de X y Y del mejor modelo de regresión no lineal simple explicativo obtenido.

Tabla 3 Valores del mejor modelo no lineal

| X | Y | X | Y | X | Y | X | Y |
|------|-------|------|-------|------|-------|------|-------|
| 1,00 | 83,17 | 3,16 | 86,67 | 4,28 | 88,48 | 5,14 | 89,88 |
| 1,53 | 84,03 | 3,33 | 86,94 | 4,40 | 88,67 | 5,24 | 90,03 |
| 1,87 | 84,58 | 3,48 | 87,19 | 4,51 | 88,86 | 5,33 | 90,19 |
| 2,15 | 85,04 | 3,63 | 87,42 | 4,62 | 89,04 | 5,43 | 90,34 |
| 2,39 | 85,43 | 3,77 | 87,65 | 4,73 | 89,21 | 5,52 | 90,49 |
| 2,61 | 85,78 | 3,90 | 87,87 | 4,84 | 89,38 | 5,61 | 90,63 |
| 2,81 | 86,10 | 4,03 | 88,08 | 4,94 | 89,55 | 5,70 | 90,77 |
| 2,99 | 86,40 | 4,16 | 88,28 | 5,04 | 89,72 | 5,78 | 90,91 |

La ecuación (20) representa el modelo de regresión no lineal simple explicativo obtenido en ésta investigación, el cual disminuye el ECM a un valor de 0.3711, siendo un mejor modelo de predicción.

$$Y_2 = -0,000221 x^2 + 1,62x + 81,55 \quad (20)$$

La ecuación 20 es usada para cualquier predicción usando regresión inversa, dado un valor Y estimado, al obtener el valor X correspondiente. Ver Tabla 4.

Tabla 4 Predicciones de valores

| Valor estimado de Y | Valor obtenido de X |
|---------------------|---------------------|
| 88 | 3,98364639 |
| 87,5 | 3,67468162 |
| 86 | 2,74794371 |
| 89 | 4,60165415 |
| 89,3 | 4,78707683 |

Conclusiones

El enfoque propuesto de dinámica de sistemas utilizado en esta investigación, demuestra la relación causal existente entre las variables independientes y la variable dependiente analizadas. Las variables independientes actúan de manera positiva en la variable dependiente, a través de bucles de reforzamiento positivos. El crecimiento de la variable de respuesta está restringido hasta alcanzar un objetivo o meta, a través de bucles de balanceo, ya que en general, las ecuaciones de regresión sólo son válidas dentro del rango de las variables regresoras contenidas en los datos observados. La presencia de bucles de reforzamiento positivos y de bucles de balanceo permite realizar la transición de un modelo de regresión lineal predictivo a un modelo de regresión no lineal explicativo. La simulación dinámica permite realizar un análisis de sensibilidad en el modelo dinámico, modificando los parámetros de las tasas de cambio y las metas a alcanzar, buscando disminuir el indicador ECM tomado como criterio de predicción. Para

cada valor establecido en la tasa de incremento y meta a alcanzar, se obtiene un modelo no lineal. Se hizo el comparativo con los modelos obtenidos experimentalmente, donde se visualiza la siguiente percepción: el modelo de regresión no lineal simple explicativo es mejor modelo predictivo, que el modelo de regresión lineal múltiple predictivo. Como investigaciones futuras está contemplado analizar el comportamiento de los residuales obtenidos y utilizar el enfoque de dinámica de sistemas, para analizar el comportamiento de la función de deseabilidad para variables de respuesta múltiples, como una herramienta en la toma de decisiones en los sectores productivos, de servicios o de salud.

Referencias

1. P. Montgomery. *Introducción al Análisis de Regresión Lineal*. 1ª ed. Ed. Ediciones CECSA. México DF., México. 2002. pp. 565.
2. E. Imgbemena, O. Mgbemena, C. Chinwuko. "A Regression Analysis Approach to Queuing System modelling: a Case of Banks". *Journal of Applied Sciences Research*. Vol. 7. 2011. pp. 200-212.
3. S. Arulchinnappan, K. Karunakaran, G. Rajendran. "Deduction of Oral Cancer Using Fuzzy Linear Regression". *Journal of computer science*. Vol. 7. 2011. pp. 1141-1145.
4. O. Okereke. "Effect of Transformation on the Parameter Estimates of a Simple Linear Regression Model: A Case Study of Division of Variables by Constants". *Asian Journal of Mathematics and Statistics*. Vol. 4. 2011. pp. 128-134.
5. K. Chung, P. Stacey. "Applying Regression Analysis to Improve Dyeing Process Quality: a Case Study". *Int J Adv Manuf Technol*. Vol. 49. 2010. pp. 357-368.
6. G. Sekelli, G. Köksal, I. Batmaz, Ö. Bayrak. "Classification Models Based on Tanaka's Fuzzy Linear Regression Approach: The Case of Customer Satisfaction". *Journal of Intelligent & Fuzzy Systems*. Vol. 21. 2010. pp. 341-355.
7. G. Bae, H. Huh. "Regression Model for Light Weight and Crashworthiness Enhancement Design of Automotive Parts in Frontal Car Crash". *International Journal of Modern Physics B*. Vol. 22. 2008. pp. 5584-5589.

8. Q. Chen, K. Au, C. Yuen, K. Yeung. "Relaxation Shrinkage Characteristics of Steam-ironed Plain Knitted Wool Fabrics". *Textile Research Journal*. Vol. 72. pp. 463-467.
9. R. Baeza, J. Yáñez, J. A. Vázquez. "Sistema Integral de Control Dinámico *SICD*: Un Enfoque Híbrido para el Cumplimiento de las Especificaciones de calidad". *DYNA*. Vol. 8. 2012. pp. 234-240.
10. J. Sterman. *Business Dynamics*. 1st ed. Ed. McGraw-Hill. Indianapolis, U.S.A. 2000. pp. 135-262.
11. H. Gutiérrez, R. de La Vara. *Análisis y Diseño de Experimentos*. 2^a ed. Ed. Mc Graw Hill. México DF., México. 2008. pp. 338-373.
12. E. Martínez. "Errores Frecuentes en la Interpretación del Coeficiente de Determinación Lineal". *Anuario jurídico y económico escurialense*. Vol. 38. 2005. pp. 315-332.