

PREDICCIÓN DE LOS PRECIOS PAGADOS AL PRODUCTOR TAMBERO ARGENTINO¹

FORECASTING ARGENTINE DAIRY FARM PRICE

Autores

Vicentin Masaro Jimena; Depetris Giguët Edith

E-mail

jvicentin@fce.unl.edu.ar

Eje Temático

Economía

Palabras claves: Predicción; Precios; Productor Tambero; VECM; ANN

Resumen

Realizar pronósticos sobre precios en un sector productivo permite reducir la incertidumbre de los posibles escenarios a los que se puede enfrentar dicho sector. Los precios son una variable clave en el análisis sectorial, y contar con estimaciones confiables de los mismos no es tarea fácil, mucho menos en países con alta volatilidad como lo es Argentina. Los métodos lineales multivariados han liderado en esta área de predicciones económicas, aunque empiezan a implementarse en trabajos recientes mecanismos no lineales que, mediante un diseño sencillo, han logrado performance predictiva competentes a las primeras en el área económica. Dentro de estos mecanismos no lineales se encuentran las redes neuronales. En el presente trabajo se ajusta un modelo de red neuronal dinámico para realizar predicciones del precio pagado al productor tambero, y se compara con las predicciones de un VECM. Se modela sobre una base de datos mensual desde 2000 a 2012, y los resultados

¹ Trabajo realizado en el marco del PICT 2130 y CAID 0101

obtenidos muestran que pueden obtenerse mejores resultados predictivos respecto a alternativas más tradicionales de series de tiempo.

1. Introducción y Justificación

La predicción de los precios es una de las cuestiones claves en el análisis sectorial. Sin embargo, tanto el nivel de producción como los precios de los sectores agropecuarios son características altamente variables con fuertes dependencias de eventualidades.

El complejo lácteo argentino no escapa a dicha complejidad y problemas de variabilidad a lo largo de su estructura. Es un sector que está integrado por niveles, a saber, la producción primaria por un lado; la industrialización por otro; la comercialización mayorista interna y externa; y por último, el consumo final (Castellano, Issaly, Iturrioz, Mateos, & Terán, 2009; Depetris, Rossini, García Arancibia, & Vicentin Masaro, 2011; Terán, 2008). El flujo de productos e insumos a lo largo de toda la cadena está coordinado por el sistema de precios, que además de ser clave para la distribución de ingresos sectoriales, constituye uno de los elementos críticos de los incentivos en la misma, y en la determinación de su competitividad.

En general, los constantes conflictos entre los agentes a lo largo de la cadena productiva devienen, principalmente, de la falta de previsibilidad y transparencia en la formación de precios del sector. Contar con estimaciones ayudaría en la toma de decisiones tanto de tipo productivas como de inversiones; y dicha necesidad se torna aún más apremiante sobre los agentes primarios, quienes tienen menor fuerza de negociación y capacidad de capturar información más integral del sector. El motivo principal de su escaso poder de negociación radica en lo altamente perecedera que resulta su producción. Pero los beneficios no solo se acotan al sector privado, sino también sería de utilidad en el sector público para la toma de decisiones.

En términos teóricos, no existe una sola herramienta de predicción: las más comunes en problemas económicos multivariados de series de tiempo son los VECM², por tener en cuenta las no estacionariedad de las series (e.g.

² Modelos de Vectores de Corrección por el Error.

Ahumada, & Cornejo, 2014; Basso, 2009; Lütkepohl, 2007). Sin embargo, existe una utilización creciente de las ANN³ con objetivos predictivos, por su flexibilidad para reconocer tanto patrones lineales como no lineales (e.g Huarng, & Yu, 2006; Maciel, & Ballini, 2010). El presente trabajo tiene como objetivo utilizar esta última herramienta en el mercado lácteo argentino para luego, analizar su performance con respecto a resultados de otras metodologías comúnmente usadas, bajo la hipótesis de que podría resultar mejor para realizar predicciones del precio pagado al productor en argentina.

2. Marco Teórico

En el equilibrio, Heien (1977) afirma que la cantidad ofrecida y demandada en cada nivel del sector se igualan, donde hay factores que afectan tanto a oferta como demanda para cada nivel, es decir,

$$\begin{aligned} Q_{n_t}^o &= f_{on}(p_{n_t}^e, z_t^{on}), \\ Q_{n_t}^d &= f_{dn}(p_{n_t}^e, p_{n+1_t}^e, z_t^{dn}), \\ Q_{n_t}^o &= Q_{n_t}^d, \end{aligned} \quad (1)$$

donde $Q_{n_t}^o$ son las cantidades ofrecidas en el nivel n , en el momento t , y $Q_{n_t}^d$ son las cantidades demandadas correspondientes; $p_{n_t}^e$ son los precios esperados del bien en el nivel n ; $p_{n+1_t}^e$ los precios esperados del producto que del bien se realiza en el nivel siguiente $n+1$; z_t^{on} hace referencia a otras variables que pueden influir sobre las cantidades ofrecidas (o) o demandadas (d) en el n -ésimo nivel.

Sin embargo, en el corto plazo se producirán desequilibrios entre la oferta y demanda, produciendo fluctuaciones en los precios resultantes, mientras que en el mediano/largo plazo deberán equilibrarse. Por ejemplo, puede suceder que la demanda minorista de lácteos se incremente por un aumento en el ingreso de los consumidores, lo que se traslada a la industria que, a su vez, demandará mayor cantidad de leche al sector primario para responder a dicho

³ Por sus siglas en inglés *Artificial Neural Networks*

desequilibrio. Como la producción de leche cruda es inelástica en el corto plazo y no puede responder en forma inmediata, pasará un cierto tiempo hasta que la oferta pueda aumentar su producción y adecuarse a la nueva demanda. Mientras tanto, la industria habrá reducido/agotado los stocks, y ante la escasez relativa de leche, los precios al productor aumentarán. Ese aumento será trasladado al minorista, lo que enfriará la demanda, y ese aletargamiento o disminución, recomenzará un nuevo ciclo en la cadena. En el corto plazo existen efectos de las cantidades sobre los precios a lo largo de la cadena, lo que Heien llama «Noción de la Dependencia del Precio».

3. Metodología

En el presente trabajo se utilizan dos metodologías de estimación aplicables a series de tiempo: por un lado, la de ANN tipo perceptrón multicapa; y por otro, VECM.

Una formulación general de una ANN proalimentada⁴ (o *feedforward*) con cuatro capas (dos ocultas) la posibilidad de incluirse una segunda capa oculta con un total de e nodos:

$$y_u = f_o \left(\sum_{d=1}^e w_d f_{h_2} \left(\sum_{j=1}^s w_j f_{h_1} \left(\sum_{i=1}^n w_{ij} x_i \right) \right) \right) \quad (2)$$

donde n es el total de nodos en la capa de entrada; s la cantidad total de nodos en la capa oculta y, m la cantidad de nodos en la capa salida. Por otro lado, x_i es el i -ésimo input, w corresponde a las ponderaciones, las cuales pueden ser: (a) interconexiones inputs-hidden layer (w_{ij}), (b) hidden layer-output (w_j), o (c) input-output (w_{ij}), f_{h_k} y f_o son las funciones de transferencias (o activación) de las capas ocultas y del output, respectivamente; y por último, y_u es el u -ésimo output.

⁴ El flujo de información recorre a lo largo de la red en una única dirección o sentido

En el presente trabajo se utilizan dos tipos de funciones de activación, a saber, la logística para los nodos ocultos y la lineal para la salida. El diseño se corresponde con un formato *ad-hoc* ya que las variables de entrada surgen de la teoría económica; y se utiliza la variable de interés predictivo como neurona de salida, y la cantidad de capas ocultas y sus correspondientes nodos, se seleccionan a partir de un procedimiento de prueba y error. El procedimiento de aprendizaje es supervisado⁵ y se utiliza un algoritmo de backpropagation para la estimación de los pesos que interconectan los nodos de la red.

Por otro lado, para la modelización por medio de VECM, se sigue la metodología de Johansen (e.g. Dickey, & Fuller, 1979, 1981; Enders, 1995).⁶

Con cada metodología se pronostica el último año disponible (2012) de manera dinámica⁷, es decir, es un proceso iterativo. El primer paso es mediante una base de datos, se estiman los parámetros por primera vez, y a partir de allí, se predice de a un período por vez, reestimando los parámetros del modelo, y volviendo a predecir el período siguiente, el cual será información de entrada para la reestimación del modelo, y predicción del siguiente período.

Con las predicciones realizadas con cada metodología, se obtienen las diferencias entre las predicciones y los valores reales observados se calculan distintas medidas de error: RSME⁸, MAD⁹ y MPE¹⁰.

Las variables utilizadas son: a) los precios mensuales pagados al productor lácteo (\$/lt), b) los precios mensuales minorista (\$/kg.); c) los precios mensuales mayorista (\$/kg.); d) los precios de exportación (u\$/kg.); e) las cantidades (lts.) producidas mensuales de leche cruda, y f) una indicadora de la estacionalidad alta en los precios al productor. Para los precios minorista, mayorista y de exportación son índices contruídos de modo de simplificar la multiplicidad de bienes. Estos índices se calculan en una canasta de bienes relevantes en cada nivel de la cadena, ponderados por su importancia en dicho

⁵ Proceso que incorpora un tutor externo y/o información global de control, de modo que, con cada conjunto de *inputs*, se suministre a la red el *output* esperado y, permita evaluar el error producido. En economía, este tipo es el más utilizado.

⁶ Aquí no se expone mayor detalle dado que es una metodología muy usada en el ámbito predictivo.

⁷

⁸ raíz del error cuadrático medio

⁹ desviación media absoluta

¹⁰ máximo porcentaje de error

nivel; y la variable relacionada con la estacionalidad ha sido construida de manera dummy indicando las estaciones con altos niveles de precios o por encima del nivel medio, cuando los precios son constantes. Todas tienen una periodicidad mensual que va desde 2000 a 2012; y las mismas (excepto la de estacionalidad) son transformadas en logaritmo, ya que de esta forma se obtienen errores del modelo VECM que cumplen los supuestos del mismo.

El conjunto de datos se subdivide en tres: como se quieren realizar 12 predicciones y evaluar las mismas, se dejan de prueba los últimos 12 períodos disponibles. Esto hace una proporción aproximada de 92% y 8% para el conjunto de entrenamiento-validación y el de prueba, respectivamente.

Las fuentes son varias, pero principalmente es el Ministerio de Agricultura, Ganadería y Pesca de la Nación Argentina, sección Lechería (MiniAgri), de donde se obtuvo información sobre a) precios mayoristas por productos, con periodicidad mensual; b) cantidades destinadas al consumo interno de lácteos, mensuales por productos; así como c) la cantidad de leche cruda producida mensualmente en la Argentina.

En particular, los precios al productor primario lechero se obtienen tanto de la Secretaría de Lechería de la Provincia de Santa Fe y del MiniAgri. La serie de precios del MiniAgri tiene una longitud menor, y recién a partir de noviembre de 2009 comenzaron a publicar información sobre pagos por litro de leche a nivel nacional, por ello se hace un empalme entre las series de estas dos fuentes a modo de obtener una serie de precios con mayor longitud.

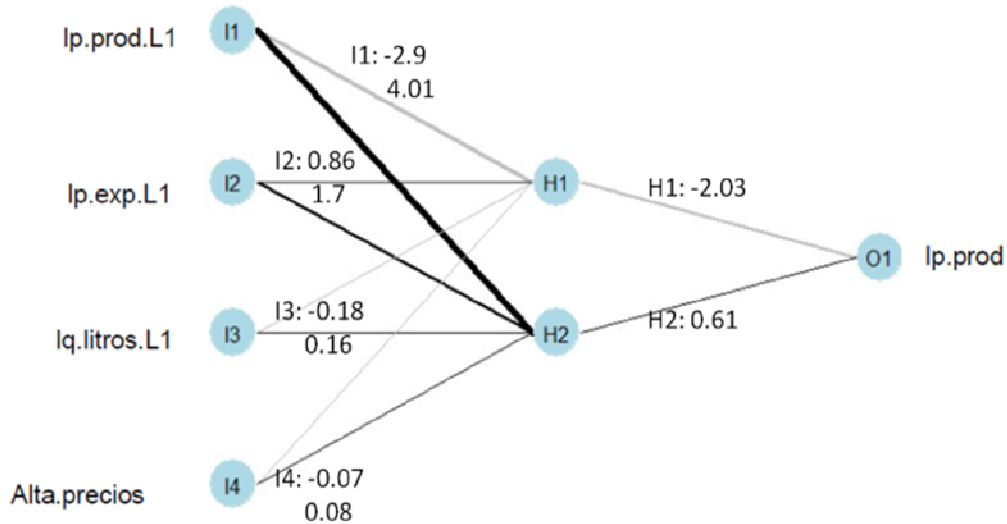
Por otra parte, los precios minoristas mensuales por productos, son obtenidos del Instituto de Estadísticas y Censos de la Provincia de Santa Fe (IPEC). A nivel nacional no se dispone de datos sobre los precios minoristas, por lo tanto, teniendo en cuenta que la provincia mencionada es una de las más importantes en el consumo nacional, se asume que los precios de la misma se comportan como lo hacen a nivel nacional. Claramente, ésta resulta una simplificación importante; sin embargo, ante la escasez de información, se propone como paliativo ante dicha situación.

4. Resultados

Con las variables expuestas en la metodología, se ha estimado tanto el VECM como ANN, cuyos parámetros no son expuestos por que varían en función al step de predicción, dado que se trata de predicciones dinámicas. Sin embargo, el VECM es un sistema multiecuacional, en donde en cada estimación el número de parámetros obtenidos es alto, ya que no sólo se estiman las relaciones de largo plazo (o parámetros de cointegración) sino que también los de corto plazo.

En cuando al modelo ANN, se ha logrado un diseño con tres capas: cuatro nodos de entrada, dos en la oculta y uno sólo en la salida. En todas las estimaciones (*step*), el proceso converge en menos de 100 iteraciones¹¹. En la figura 1 se muestra la arquitectura seleccionada con los coeficientes estimados para las conexiones del primer *step* de predicción del proceso. Comparativamente, la ANN es menos costosa en cantidad de parámetros en cada *step*.

Figura 1. Estructura de la Red para Predecir el precio al productor tambero en el primer *step*.

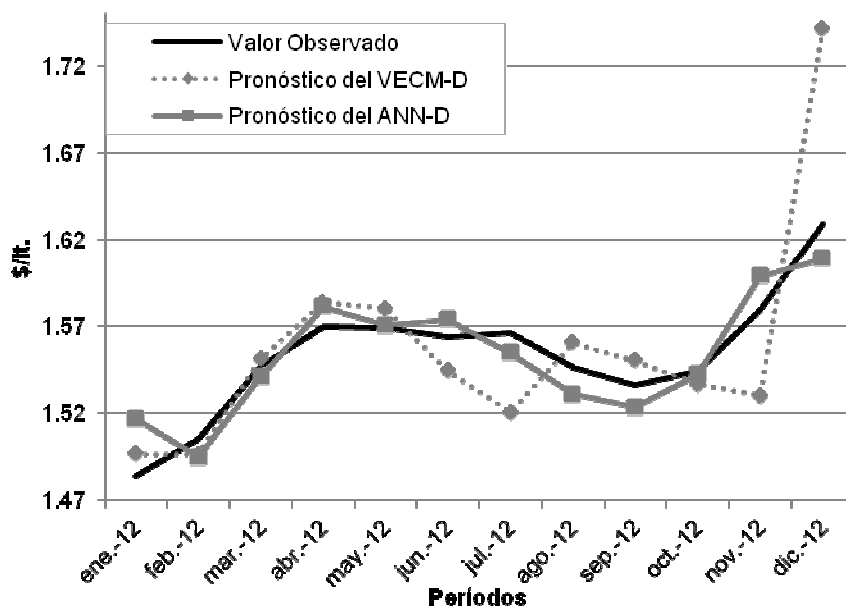


Nota: Los nodos reguladores del error están contemplados en el modelo, aunque no son expuestos.
Fuente: Elaboración propia.

¹¹

Cuando se realiza comparaciones de las predicciones dinámicas de los precios pagados al productor tambero realizadas por ambos modelos (figura 2), puede observarse que, en promedio, las predicciones del VECM tienen errores superiores que las correspondientes a la de ANN (RMSE: 0.024 vs. 0.009; MAD: 0.016 vs. 0.008, y MPE (%): 3.585 vs. 1.894, respectivamente), donde estas última muestra medidas menores a la mitad de la primera.

Figura 2. Valores Observados y Predicciones Realizadas por los Modelos Dinámicos para el Período 2012.



Fuente: Elaboración propia.

5. Conclusiones

La importancia de contar con predicciones de precios en un sector productiva abarca tanto el ámbito privado como público, permite la planificación estratégica productiva y de inversiones. Sin embargo, la predicción en los sectores agrícolas ganaderos resulta complicada por el alto grado de

variabilidad de estos sectores, por su dependencia climática como por los vaivenes políticos. Sin embargo, el sector lácteo argentino no ha sido objetivo de modelización alguna que permitiese contar con algún tipo de predicción sectorial. Es por este que el presente trabajo tiene como objetivo analizar herramientas multivariadas de predicción y la evaluación de las mismas, donde una de ellas ha sido poco utilizada en el ámbito económico, pero que sin embargo está tomando fuerza cuando los objetivos son meramente predictivos.

Mediante la modelización de redes neuronales y VECM, se han realizado predicciones dinámicas del precio pagado al productor, y se ha observado que, en este caso, los resultados de las redes neuronales son más parecidos y siguen similar evolución a los datos reales. Muestran medidas de performance muy inferiores a las del VECM, concluyendo que dicha metodología es más adecuada para realizar las predicciones de precios de este eslabón de la cadena.

6. Bibliografía

- Ahumada, H., & Cornejo, M. (2014). Explaining commodity prices by a cointegrated time series-cross section model. *Empirical Economics*, 48(4), 1667–1690.
- Basso, M. A. (2009). *Variaciones en los precios y tipo de cambio. La paridad de poder de compra Argentina* (Tesis de Maestría). Universidad Nacional de Córdoba, Córdoba, Argentina.
- Castellano, A. I., Iturrioz, L. C., Mateos, G. M., Teran, M., & Castellano, J. C. A. (2009). Análisis de la cadena de la leche en Argentina (Informe Área Estratégica de Economía y Sociología del INTA), 136 pág. Buenos Aires,
- Depetris Guiguet, E., Rossini, G., García Arancibia, R., & Vicentin Masaro, J. (2011). *Competitividad del complejo lácteo santafesino. Período 2000-2009*. UNL Ediciones, 386 págs, ISBN 978-987-657-602-4.
- Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, 49(4), 1057–1072.
- Enders, W. (1995). *Applied econometrics times series*. New York, United States: Wiley.
- Engle, R. F., & Granger, C. W. J. (1987). Cointegration and error correction: representation, estimation and testing. *Econometrica*, 55(2), 251–276.

- Granger, C. W. J. (1981). Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16(1), 121–130.
- Heien, D. (1977). Price determination processes for agricultural sector models. *American Journal of Agriculture Economics*, 59(1), 126–132.
- Huang, K., & Yu, T. H. K. (2006). The application of neural networks to forecast fuzzy time series. *Physica A: Statistical Mechanics and its Applications*, 363(2), 481–491.
- Lütkepohl, H. (2004a). Univariate time series analysis. In H. Lütkepohl y M. Kräzig. (Ed.), *Applied Time Series Econometrics*, Cambridge, United Kingdom: Cambridge University Press.
- Maciel, L. S., & Ballini, R. (2008). Design a neural networks for time series financial forecasting: accuracy and robustness analysis. *Anales do 9º Encontro Brasileiro de Finanças*, Sao Paulo, Brazil.
- Terán, J. C. (2008). Caracterización de la cadena agroalimentaria de la leche en la provincia de Santa Fe (Informe Área Estratégica de Economía y Sociología del INTA), Rafaela.
- Vicentin Masaro, J., Depetris Guiguet, E., García Arancibia, R. & Rossini, G. (2013). Retrasos en la transmisión de precios de exportación entre los principales productos lácteos santafesinos: Leche en polvo entera y quesos. *Ciencias Económicas de la UNL*, 2(10), pp. 11-21.