

Pronóstico del Índice de Producción Industrial con una red neuronal artificial

Andrés Velasco M.

Agosto 12 de 2004

Contenido

1. Importancia de un buen pronóstico del IPI
2. Variables clave
3. Metodología usada para el pronóstico
4. Resultados

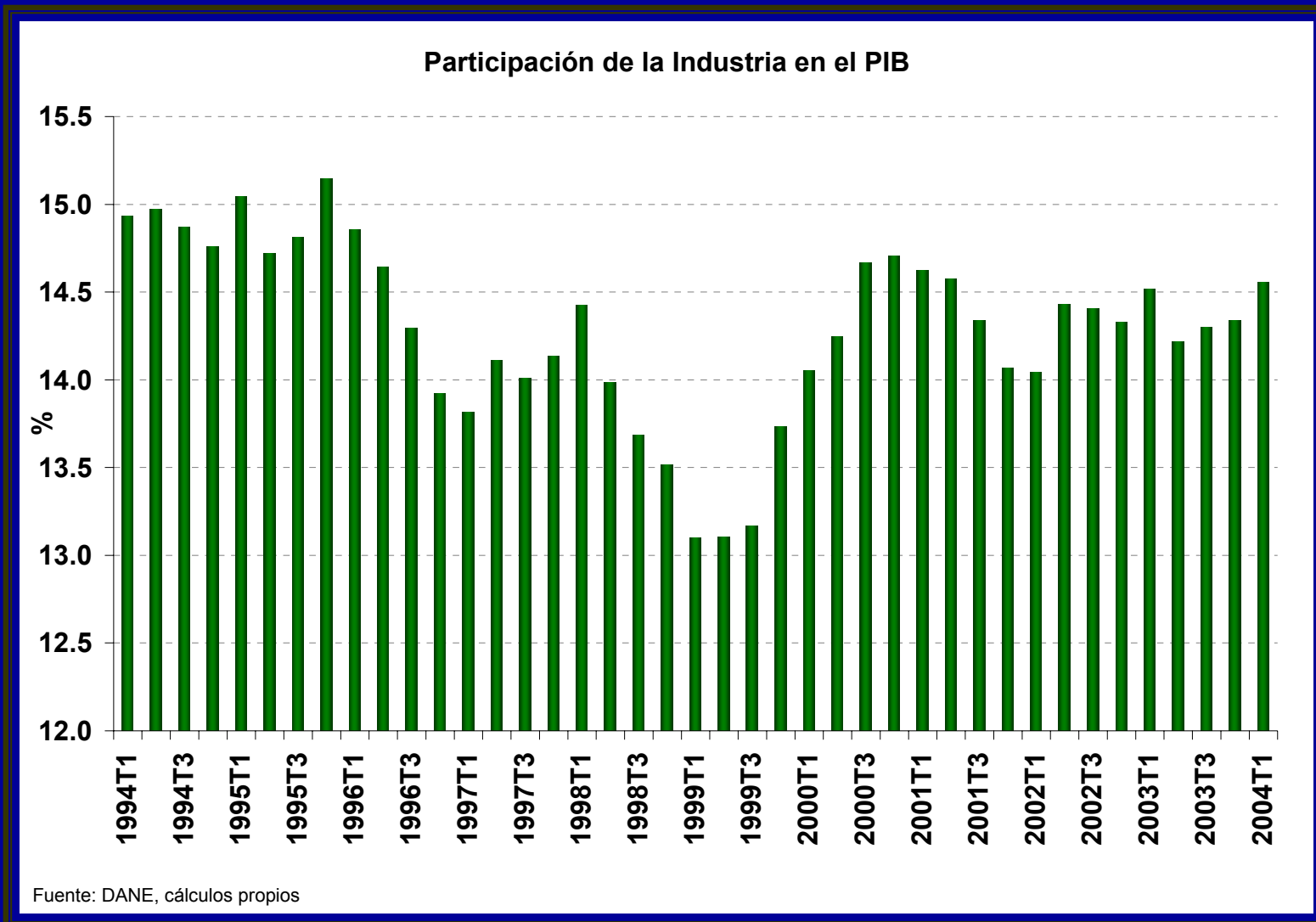
Contenido

- 1. Importancia de un buen pronóstico del IPI**
 - A. En el pronóstico del PIB
 - B. Como insumo para otros modelos
2. Variables clave
3. Metodología usada para el pronóstico
4. Resultados

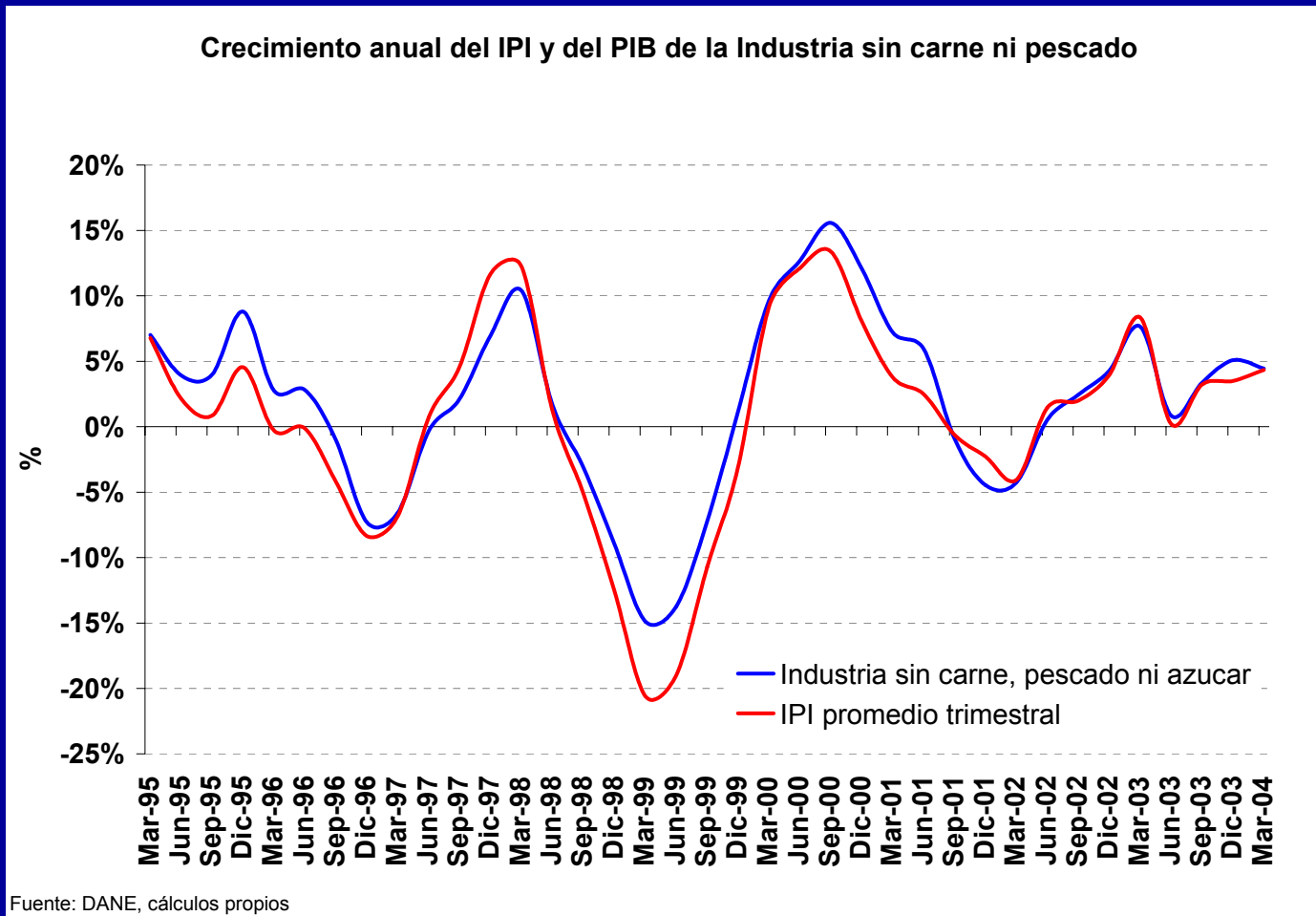
Contenido

- 1. Importancia de un buen pronóstico del IPI**
 - A. En el pronóstico del PIB**
 - B. Como insumo para otros modelos
2. Variables clave
3. Metodología usada para el pronóstico
4. Resultados

La participación de la industria en el PIB es del 14,5% aproximadamente...



... y El IPI explica buena parte de la dinámica del sector industrial.



La producción industrial sin carne, pescado ni azúcar **pesa 94%** en el total de la producción industrial.

Correlación entre las dos series 1994:1-2004:1: de 96%

Adicionalmente el PIB industrial está altamente correlacionado con otros sectores productivos, algunos ejemplos:

- Coeficientes de correlación con:
 - Comercio reparación restaurantes y hoteles (10,8% del PIB): 85%
 - Transporte sin correos ni telecomunicaciones (5,5% del PIB): 78%
 - Electricidad, gas de ciudad y agua (3,2% del PIB): 50%
 - Construcción (5% del PIB): 51%

Un cálculo con un modelo de equilibrio general computable indica que la dinámica de la industria explica cerca de una tercera parte de la dinámica del PIB

Contenido

- 1. Importancia del pronóstico del IPI**
 - A. En el pronóstico del PIB
 - B. Como insumo para otros modelos**
2. Variables clave
3. Metodología usada para el pronóstico
4. Resultados

El IPI es usado en varios modelos que sirven como herramientas de política monetaria:

- Modelo RNA para el IPC de no transables, mensual
- Modelo RNA para el IPC de transables, mensual
- Modelo para IPC total, P-estrella, trimestral (velocidad de M1 estacionaria)
- Modelo P-estrella modificado, trimestral (velocidad de la base monetaria no estacionaria)
- Modelo de índice coincidente de la actividad económica y del indicador líder

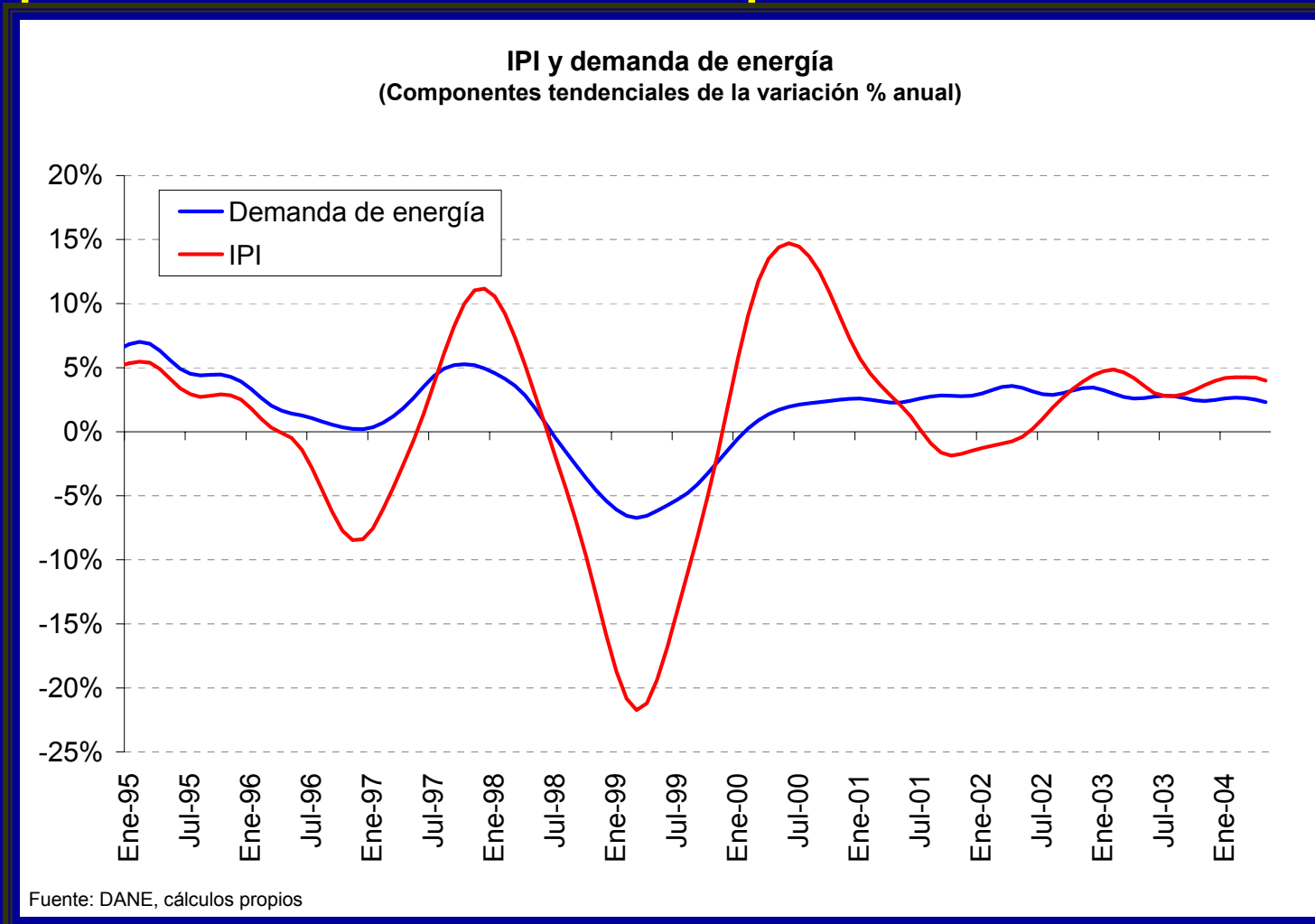
Contenido

1. Importancia de un buen pronóstico del IPI
- 2. Variables clave**
 - A. Demanda de energía
 - B. Número de días hábiles
3. Metodología usada para el pronóstico
4. Resultados

Contenido

1. Importancia de un buen pronóstico del IPI
- 2. Variables clave**
 - A. Demanda de energía**
 - B. Número de días hábiles
3. Metodología usada para el pronóstico
4. Resultados

La relación entre la demanda de energía y la producción industrial parece evidente



Se encontró un coeficiente de correlación contemporánea de 72% entre las dos series. La demanda de energía se publica con un mes de anticipación al IPI.

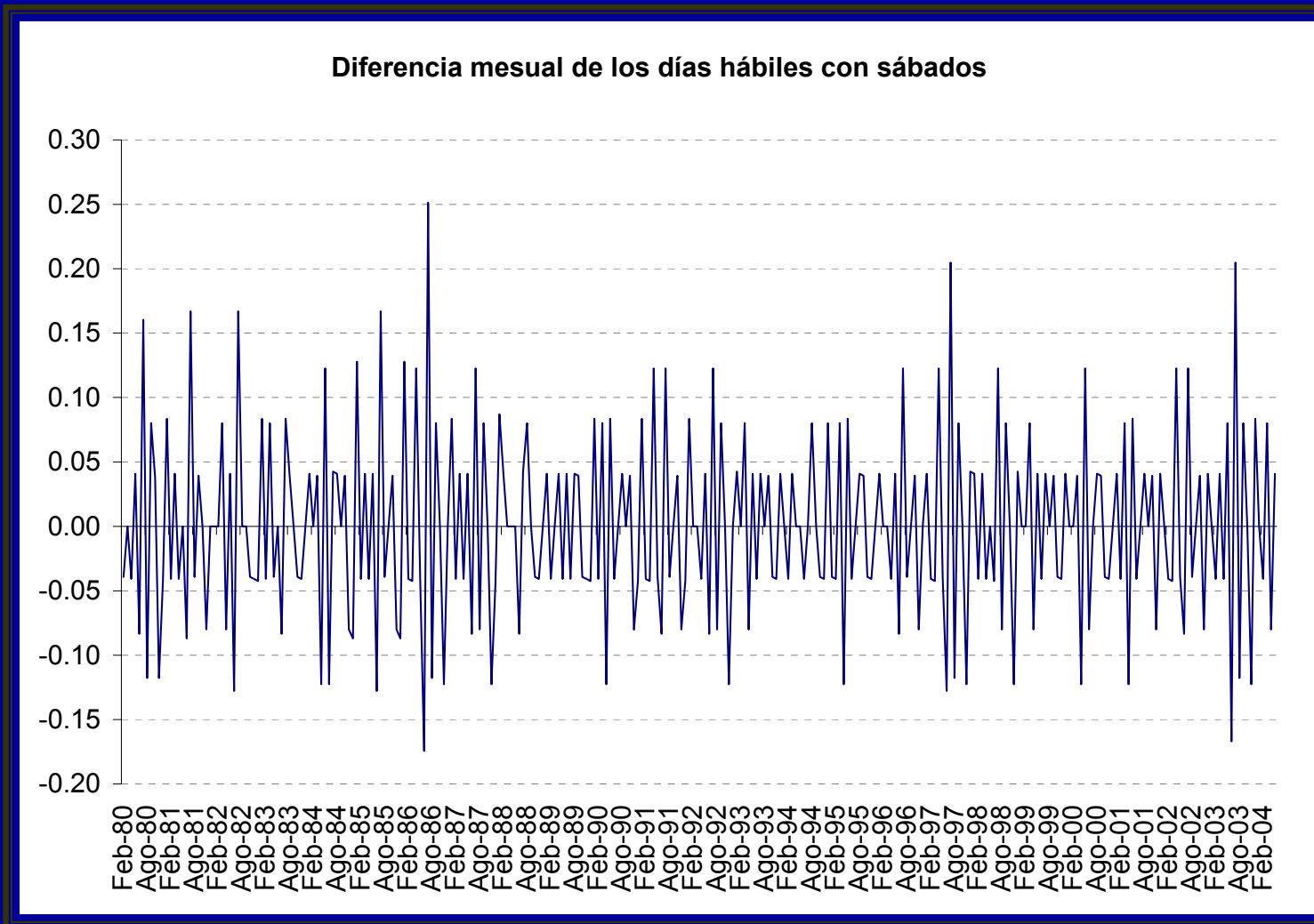
Contenido

1. Importancia de un buen pronóstico del IPI
- 2. Variables clave**
 - A. Demanda de energía
 - B. Número de días hábiles**
3. Metodología usada para el pronóstico
4. Resultados

Se ha encontrado una relación importante entre el número de días hábiles sin sábados y la producción industrial (R^2 -ajustado de 35%)

- Para obtener una variable continua se trabajó con la diferencia mensual del número de días hábiles.
- En un trimestre, un día hábil adicional en una base de 72 días hábiles puede significar 1,4% adicional en el crecimiento de la industria, como sucedió en el primer trimestre de éste año.
- Dos días adicionales en un mes, como es el caso de junio de 2004, pueden significar cerca de 9% de crecimiento adicional en el IPI.

Se trabajó con la diferencia del logaritmo natural mensual del número de días hábiles



Se ha encontrado una relación importante entre el número de días hábiles sin sábados y la producción industrial (R^2 -ajustado de 35%)

- Para obtener una variable continua se trabajó con la diferencia mensual del número de días hábiles.
- En un trimestre, un día hábil adicional en una base de 72 días hábiles puede significar 1,4% adicional en el crecimiento de la industria (p.e.: primer trimestre de 2004).
- Dos días adicionales en un mes, como es el caso de junio de 2004, pueden significar cerca de 9% de crecimiento adicional en el IPI.

Contenido

1. Importancia de un buen pronóstico del IPI
2. Variables clave
3. **Metodología usada para el pronóstico:
Redes neuronales análogas**
 - A. Las redes neuronales
 - B. La red en detalle
 - C. Estimación de la red
4. Resultados

Contenido

1. Importancia de un buen pronóstico del IPI
2. Variables clave
3. **Metodología usada para el pronóstico:
Redes neuronales análogas**
 - A. **Las redes neuronales**
 - B. La red en detalle
 - C. Estimación de la red
4. Resultados

Las redes neuronales

- Es un modelo diseñado para imitar el comportamiento de un cerebro, cuyas neuronas se activan y desactivan ante la presencia de impulsos dependiendo de la intensidad del estímulo. Esto puede asociarse a un **comportamiento no lineal** ante las circunstancias. Kuan y White (1994) y Swanson y White (1995).
- En el contexto de las series de tiempo puede clasificarse como un modelo no lineal entrenado, en la medida en que puede aprender patrones que se repitan dentro de muestra para aplicarlos fuera de muestra

Las redes neuronales

- La gran ventaja del proceso de aprendizaje de la red es también su mayor desventaja a la hora de realizar pronósticos (puede aprender ruidos o movimientos aleatorios).
- Una buena red aísla los choques aleatorios mientras que aprende el componente de la serie que se repite. Al escoger una red es indispensable que ésta sea efectiva en sus pronósticos dentro y fuera de muestra.

Contenido

1. Importancia de un buen pronóstico del IPI
2. Variables clave
3. **Metodología usada para el pronóstico:
Redes neuronales análogas**
 - A. Las redes neuronales
 - B. La red en detalle**
 - C. Estimación de la red
4. Resultados

La red en detalle

- La red se compone de dos partes, una lineal con intercepto (modelo AR-X) y otra no lineal.

$$y_t = \phi x_t + \sum_{j=1}^q B_j G(\gamma_j z_t)$$

$$G(x) = \frac{1}{1 + e^{-x}}$$

$$z_t \subseteq x_t$$

La red en detalle

- ϕ es el vector de parámetros estimados con OLS del modelo AR-X, y 'x' contiene el conjunto de información disponible para la parte lineal: rezagos de la variable endógena y de las variables explicativas adicionales
- Se trabaja, generalmente, con las variables normalizadas. (Franses, et. al (1999))

La red en detalle

- La parte no lineal de la red contiene funciones de activación que se denotan con ' q '. Estas son funciones logísticas $G(x)$ que usan para modelar el paso de una variable de un estado a otro.
- Durante el proceso de selección de la red se debe elegir el número óptimo de ' q ', entre más haya mejor será el ajuste del modelo dentro de muestra.

La red en detalle

- La red se compone de dos partes, una lineal con intercepto (modelo AR-X) y otra no lineal.

$$y_t = \phi x_t + \sum_{j=1}^q B_j G(\gamma_j z_t)$$

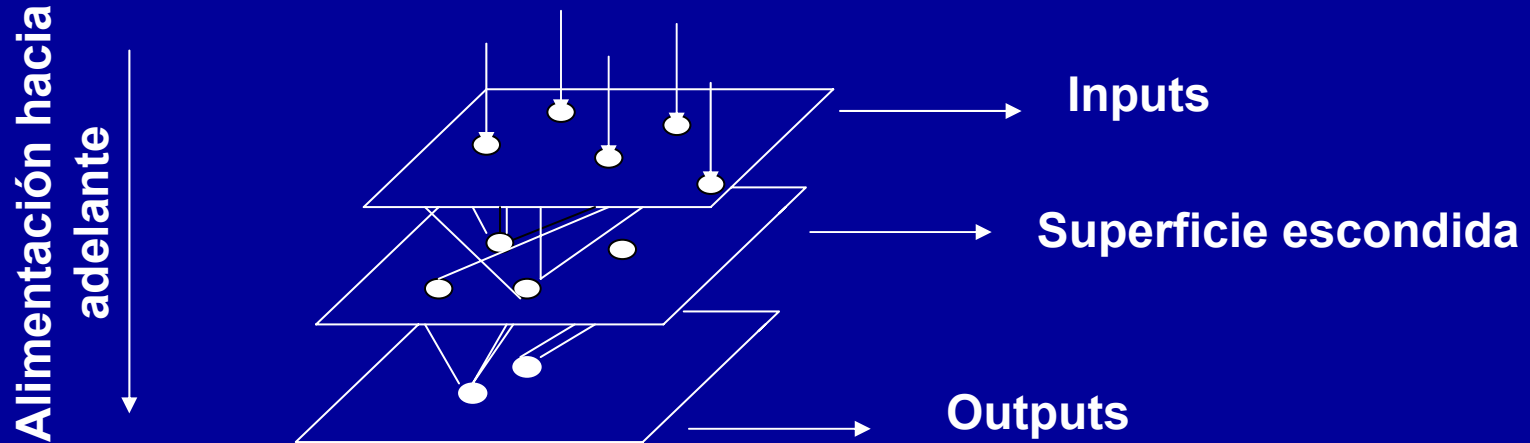
$$G(x) = \frac{1}{1 + e^{-x}}$$

$$z_t \subseteq x_t$$

‘Nodos de decisión’

El proceso de aprendizaje → se basa en los cambios de las ponderaciones de sus conectores. El sistema adquiere nuevos conocimientos ajustando el conjunto de ponderaciones.

Figura 3: Estructura topológica



La red en detalle

- Los γ son los parámetros de las funciones de activación, que provienen inicialmente de una distribución uniforme aleatoria entre $[-1, 1]$. Estos determinan la respuesta cuando se activa cada una de las funciones. Dependiendo de su valor pueden atenuar o amplificar la señal original de la serie.
- Estos γ se escogen mediante un método indirecto de selección a partir de simulaciones (con criterios de calidad de los pronósticos)

La red en detalle

- La red se compone de dos partes, una lineal con intercepto (modelo AR-X) y otra no lineal.

$$y_t = \phi x_t + \sum_{j=1}^q B_j G(\gamma_j z_t)$$

$$G(x) = \frac{1}{1 + e^{-x}}$$

$$z_t \subseteq x_t$$

La red en detalle

- En general $\gamma(i,j)$ corresponde al *i*-ésimo parámetro de la función de activación *j*-ésima. En este caso *i* depende del número de parámetros Z , que es el conjunto de información de la parte no lineal del modelo. *j*, por su parte, es el número de '*q*' del modelo.
- En este trabajo se asumió que Z es un subconjunto de x (No. de elementos de la parte lineal). El número de elementos de Z , se denota por '*p*'.

La red en detalle

- De tal forma que cuando en la literatura se habla de que la arquitectura de una red neuronal es (p,q) , se está haciendo alusión al número de variables explicativas de la parte no lineal, y al número de funciones de activación.

Contenido

1. Importancia de un buen pronóstico del IPI
2. Variables clave
3. **Metodología usada para el pronóstico:
Redes neuronales análogas**
 - A. Las redes neuronales
 - B. La red en detalle
 - C. Estimación de la red**
4. Resultados

Estimación de la red

- Inicialmente se estima la parte lineal de la red (AR-X):
 - *Forward* (modelos de una variable, con dos variables..., se escoge el mejor en cada uno de los casos según la prueba F, se dejan de incluir variables cuando éstas no resulten significativas al 5%).
 - *Backward* (de modelos grandes hacia más pequeños)
 - *Stepwise*, se escogen los mejores modelos AR-X bajo los criterio de R².

Estimación de la red

- Después se procede a estimar la parte no lineal. Como los γ vienen de una selección aleatoria, se debe considerar un número grande de estimaciones aleatorias de los parámetros para cada una de las arquitecturas (p,q) . En este trabajo se realizaron 50 selecciones diferentes, se ha aprendido que no hay beneficios adicionales de realizar más simulaciones en la medida en que *computacionalmente* es muy exigente.

Estimación de la red

- Una vez estimados los γ , éstos se multiplican por el vector Z , de tal forma que se obtienen j vectores con estos productos. Con OLS, se estiman los parámetros B que ponderan la participación de cada uno de las funciones logísticas dentro de la parte no lineal.
- Con el conjunto de información inicial (parámetros de la parte lineal: ϕ , y de la no lineal: γ y B), se suman la parte lineal con la no lineal y se produce un pronóstico preliminar.
- Se hace la estimación no lineal según una función objetivo (suma de los errores al cuadrado más un término conocido como *weight decay*, que se encarga de penalizar parámetros muy grandes).

Estimación de la red

- Con todo lo anterior se obtiene un valor de la función objetivo optimizando cada uno de los 50 juegos iniciales, para cada una de las arquitecturas.
- Después se escoge para cada una de las arquitecturas el vector de parámetros que produzca el menor valor de la función objetivo.
- Con la estimación de los parámetros para cada arquitectura, se escoge la mejor red según criterios de evaluación de pronóstico: AIC, BIC, R², RMSE, RMSPE, MAE, MAPE, SR y la U-THEIL
- Una vez se escoge la mejor red se procede a hacer 'desconexiones' en el componente no lineal, de acuerdo una vez más a las medidas de evaluación.
- Todo lo anterior se hace dos veces: dentro y fuera de muestra

Contenido

1. Importancia de un buen pronóstico del IPI
2. Variables clave
3. Metodología usada para el pronóstico
4. **Resultados**
 - A. Selección de la arquitectura
 - B. El pronóstico fuera de muestra

Contenido

1. Importancia de un buen pronóstico del IPI
2. Variables clave
3. Metodología usada para el pronóstico
4. **Resultados**
 - A. **Selección de la arquitectura**
 - B. El pronóstico fuera de muestra

Selección de la arquitectura

- Bajo los criterios usados, el mejor modelo AR-X para proyectar el IPI contiene los rezagos 1, 8, 10, 12, 13, 14, 15, 16, 17 de la variable endógena; los rezagos 0, 1, 9, 11, 12, 14, 18 de la demanda de energía y los días hábiles de cada período (contemporáneos con la variable endógena)

Selección de la arquitectura

- Las mejores arquitecturas dentro de muestra fueron:
 1. (16,3)
 2. (15,2)
 3. (15,4)
 4. (16,2)
- Las mejores arquitecturas fuera de muestra fueron:
 1. (3,2)
 2. (4,4)
 3. (15,4)
 4. (3,4)

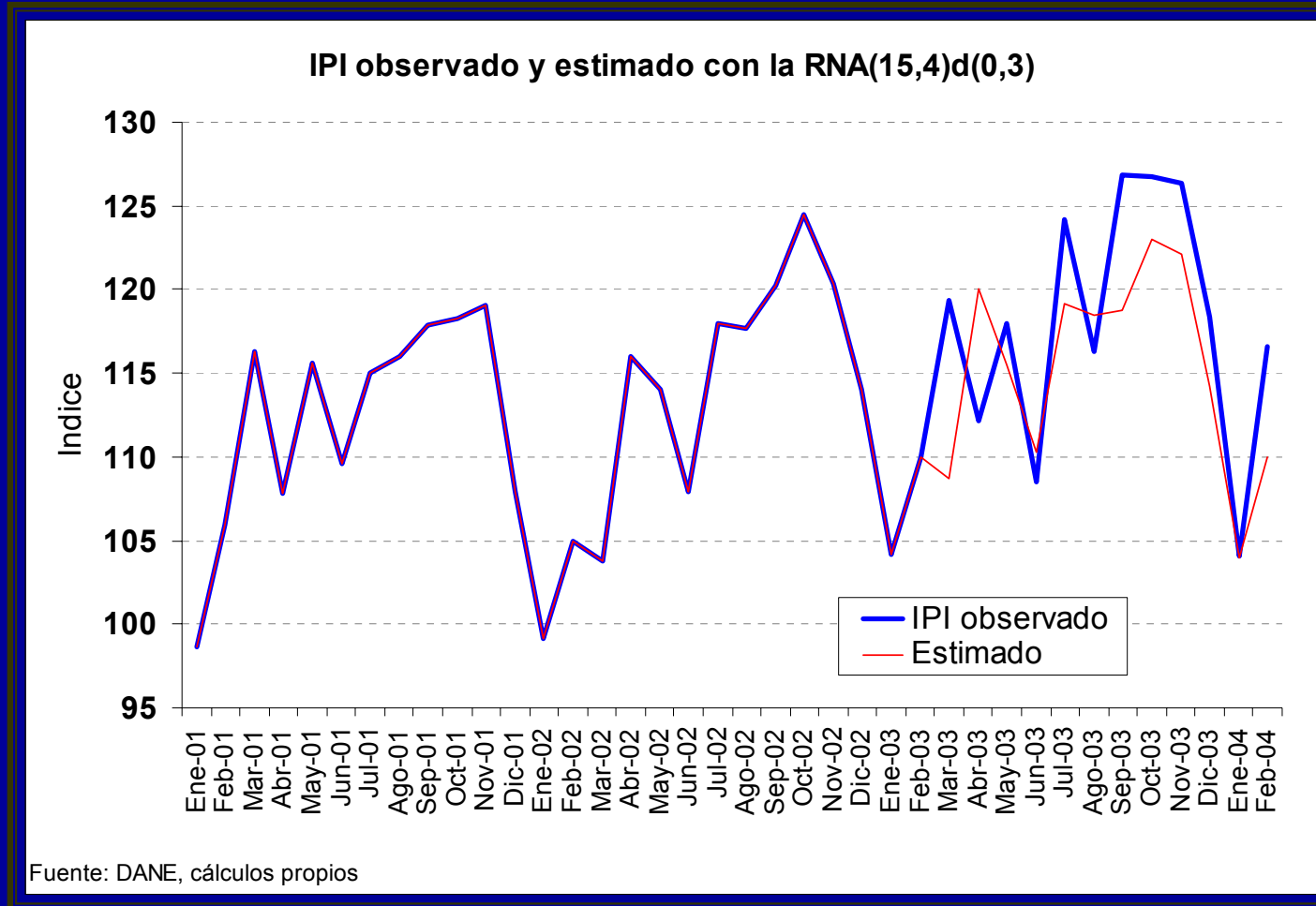
Selección de la arquitectura

- Se escogió la arquitectura (15,4) como el mejor modelo según los criterios descritos anteriormente.
- Nótese que la arquitectura (0,0) no resulto entre las mejores ni dentro ni fuera de muestra.
- A la red escogida se le practicaron desconexiones dentro y fuera de muestra, los resultados indicaron que si se desconectaba el 3er. parámetro de la parte no lineal (el rezago 8 del IPI), los pronósticos de la red mejoraban.

Contenido

1. Importancia de un buen pronóstico del IPI
2. Variables clave
3. Metodología usada para el pronóstico
- 4. Resultados**
 - A. Selección de la arquitectura
 - B. El pronóstico fuera de muestra**

El pronóstico fuera de muestra



El ECM del pronóstico a 12 meses, de la tasa de crecimiento anual del IPI es de 0,04%.

FIN

Aunque se reconoce que no es clara la dirección de causalidad entre las dos variables, ya que existe una aparente endogeneidad entre las dos se supuso independencia del IPI

Test de causalidad bajo el criterio de Granger

Muestra 1991:1-2004:5

Rezagos: 2

Hipótesis nula	Obs	Estad.-F	Probabilidad
El IPI no causa según Granger la demanda de energía	159	2.5521	0.0812
La demanda de energía no causa según Granger el IPI		8.1786	0.0004

Fuente: cálculos propios