

Multivariate forecast of Ecuadorian financial indexes using Gaussian DBNs with an extension of bnlearn.

David Quesada López
Gabriel Antonio Valverde Castilla
Pedro Larrañaga
Concha Bielza

16/11/2019



UNIVERSIDAD
POLITÉCNICA
DE MADRID

- Introducción
 - Serie temporal multivariante
 - Dataset de índices econométricos ecuatorianos
- Redes bayesianas dinámicas
 - Funcionamiento y representación
 - Herramientas: bnlearn y dbnR
- Caso práctico
 - Construcción de la red
 - Visualización
- Resultados
- Conclusiones
- Referencias

INTRODUCCIÓN

Definición de serie temporal

Podemos definir una serie temporal multivariante como:

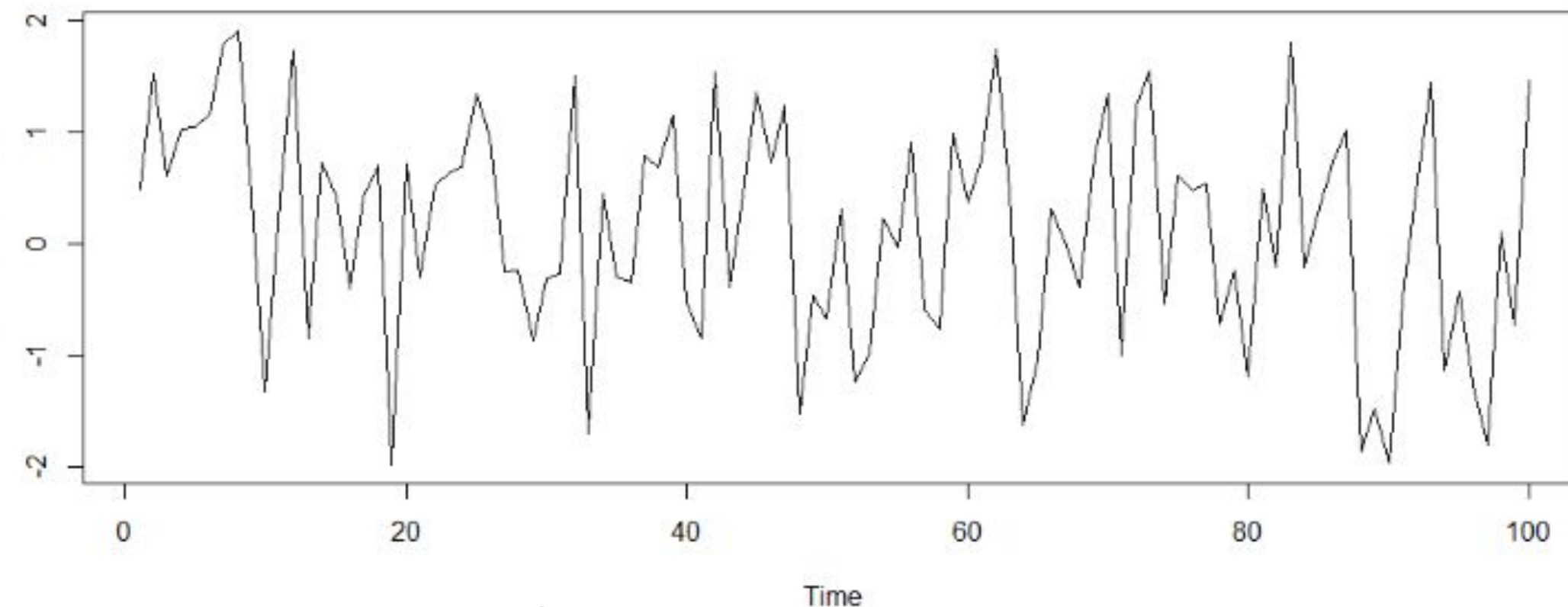
- Secuencia de N observaciones de varias características
- Ordenadas y equidistantes cronológicamente

Suelen representarse como:

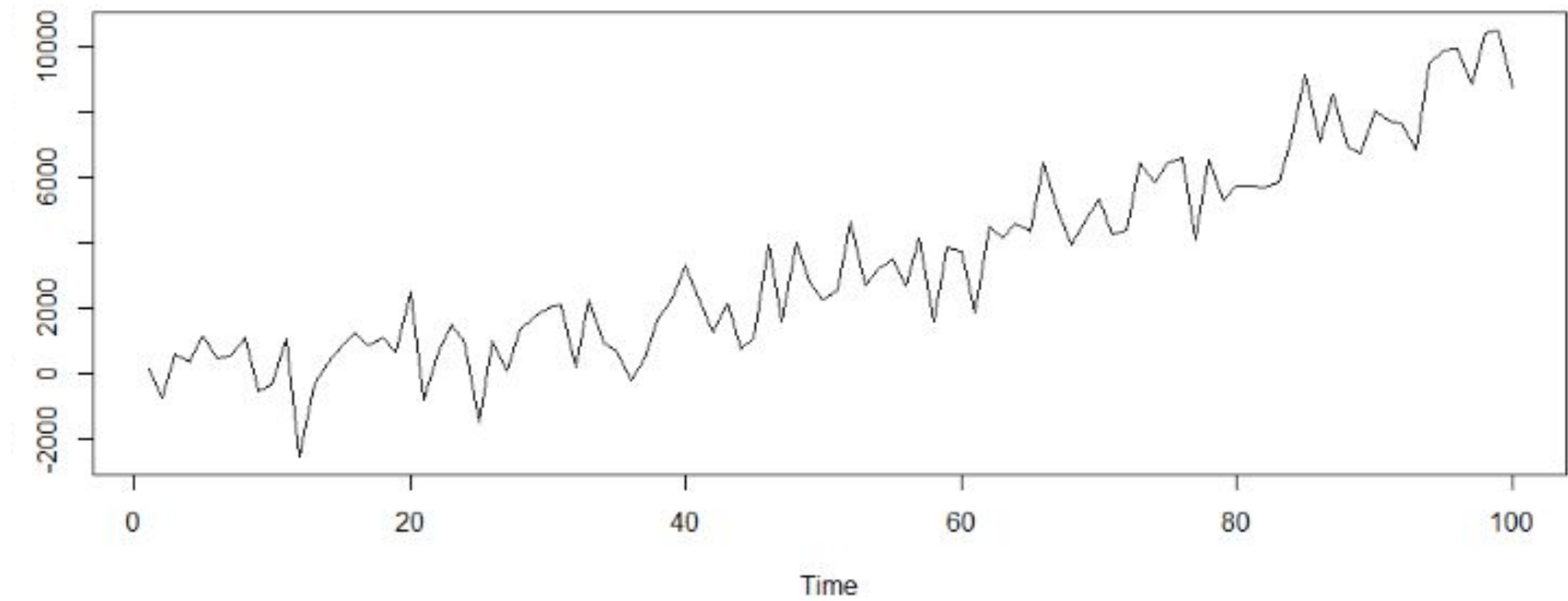
$$y_1, \dots, y_N; (y_t)_{t=1}^N; (y_t: t = 1 \dots N),$$
$$y_t \equiv [y_{t1}, y_{t2}, \dots, y_{tM}]' (M \geq 2)$$

Las series temporales pueden ser además estacionarias, si sus propiedades estadísticas se mantienen a lo largo del tiempo, y estacionales, si presentan una componente cíclica.

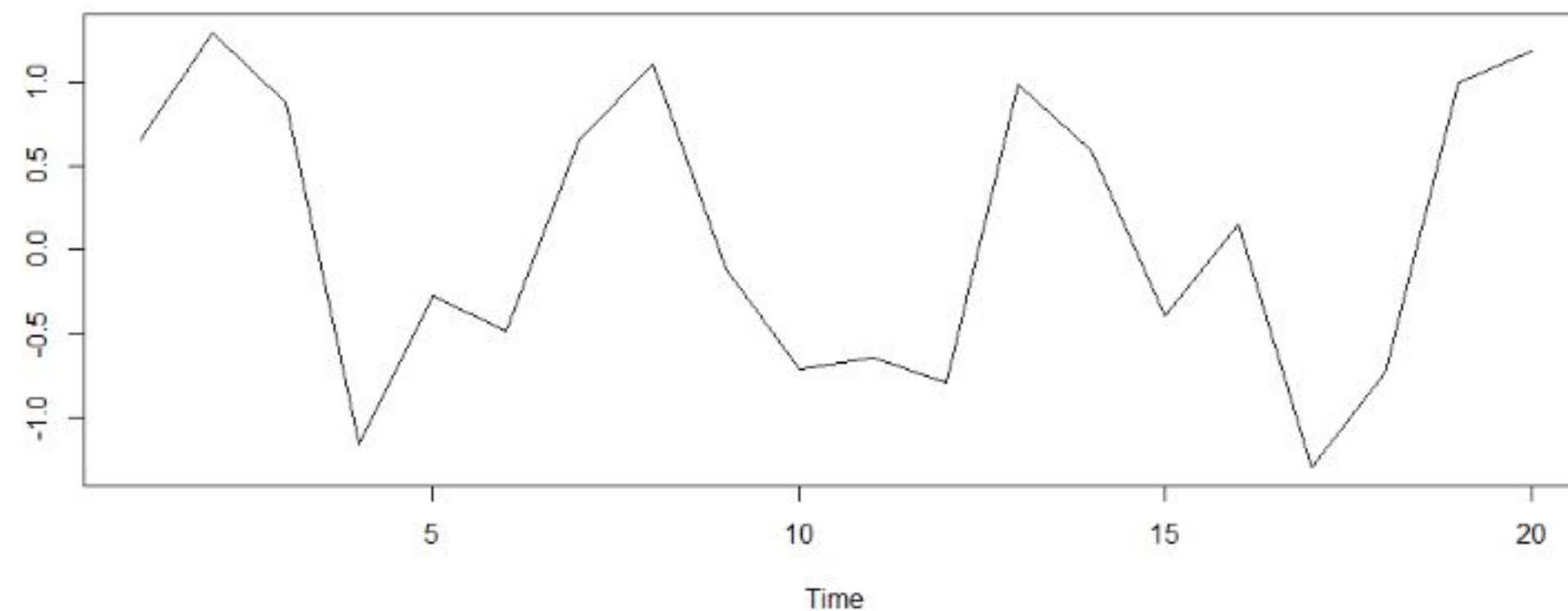
Ejemplos de series estacionarias y estacionales



Serie estacionaria



Serie no estacionaria con tendencia



Serie estacional

Modelos típicos de series temporales multivariantes:

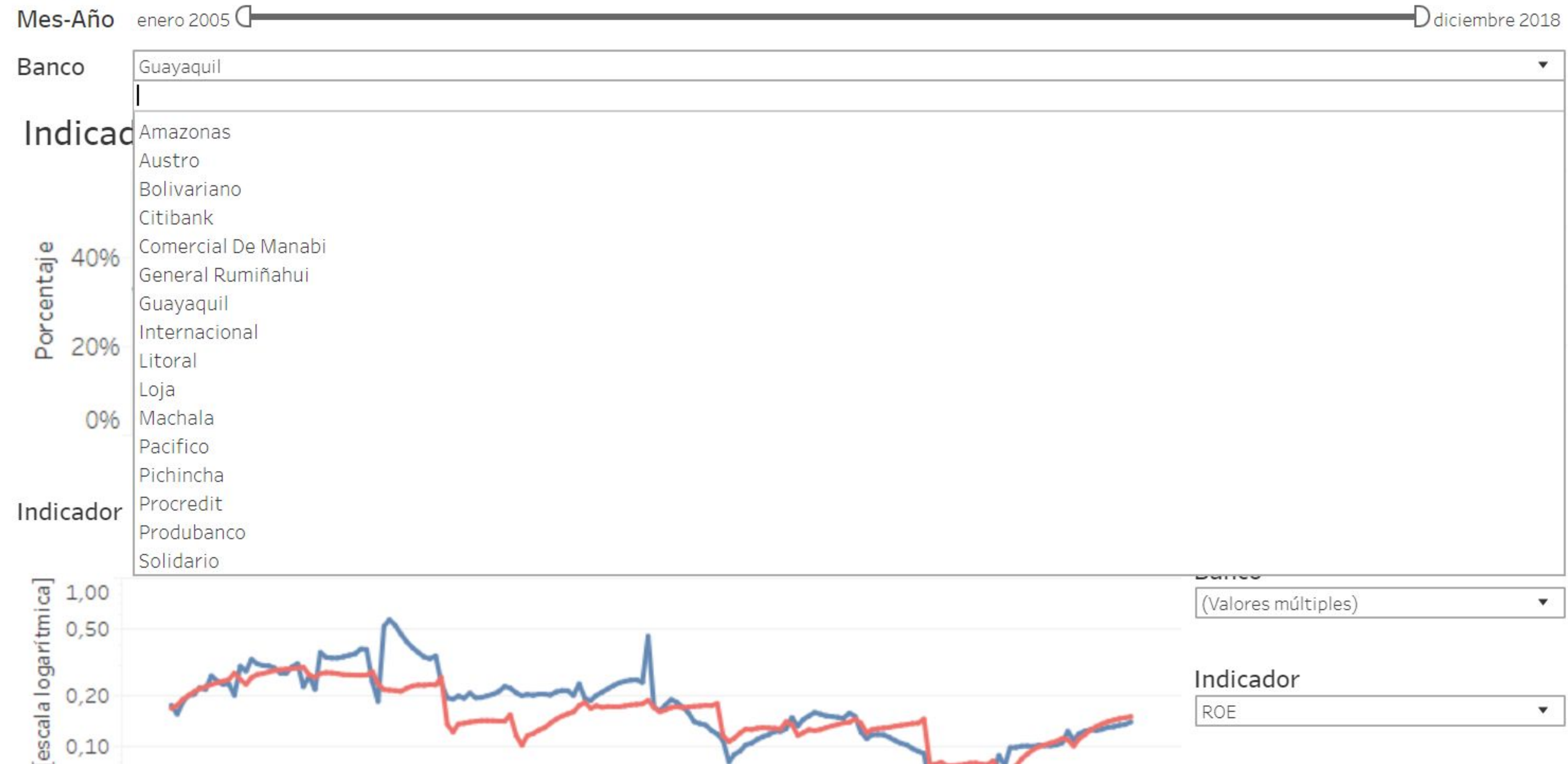
- VARIMA
- BSTS
- LSTM
- DBN

Dataset sobre índices descriptivos de los principales bancos ecuatorianos



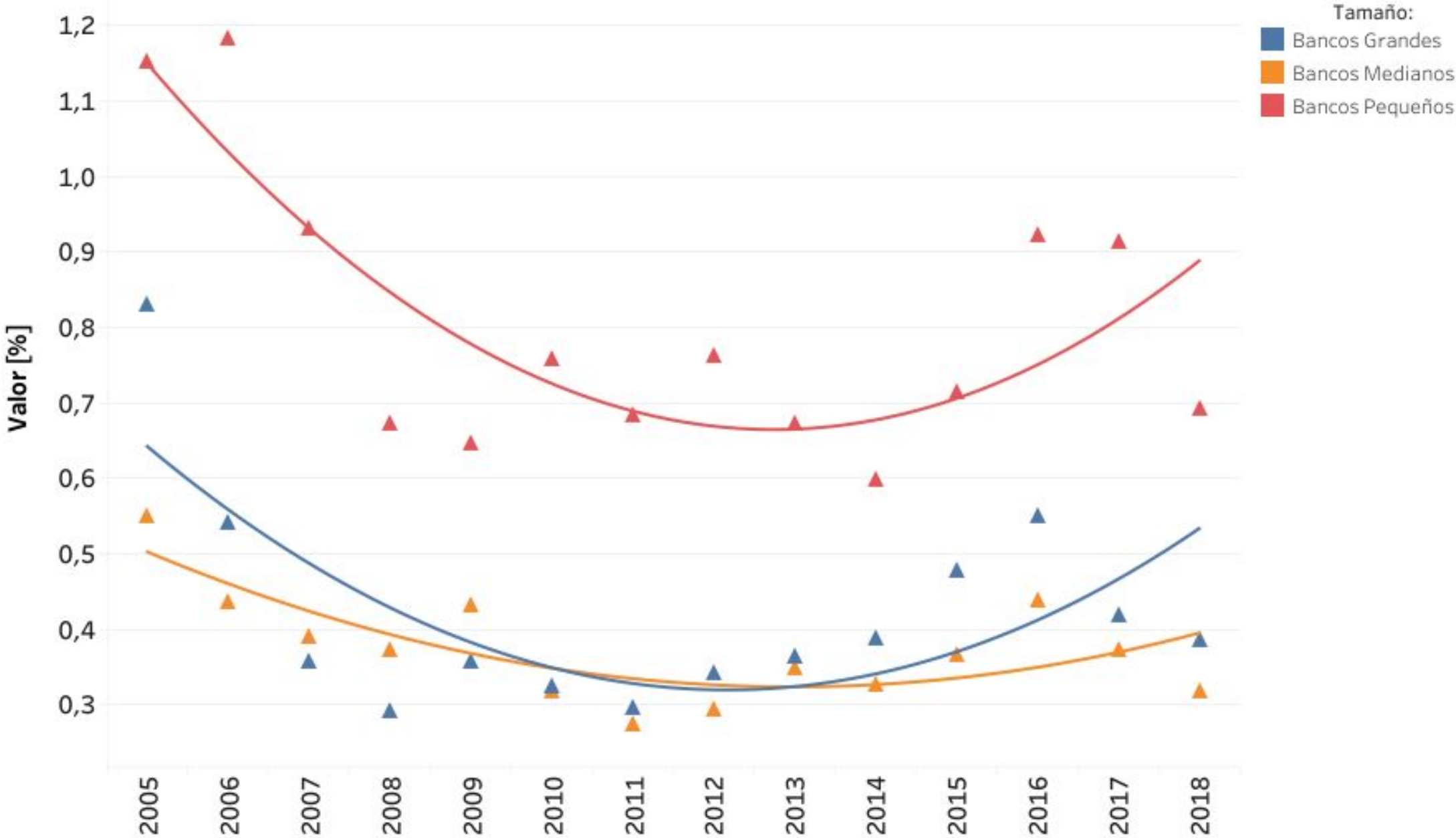
Indicador	Forma de cálculo
Rentabilidad	Utilidad Neta / Patrimonio
Liquidez	Fondos Disponibles / Total depósitos a corto plazo
Cobertura	Provisiones / Cartera Improductiva
Calidad de Activos	Activos Productivos / Pasivos con costo
Morosidad	Cartera Improductiva / Cartera Bruta



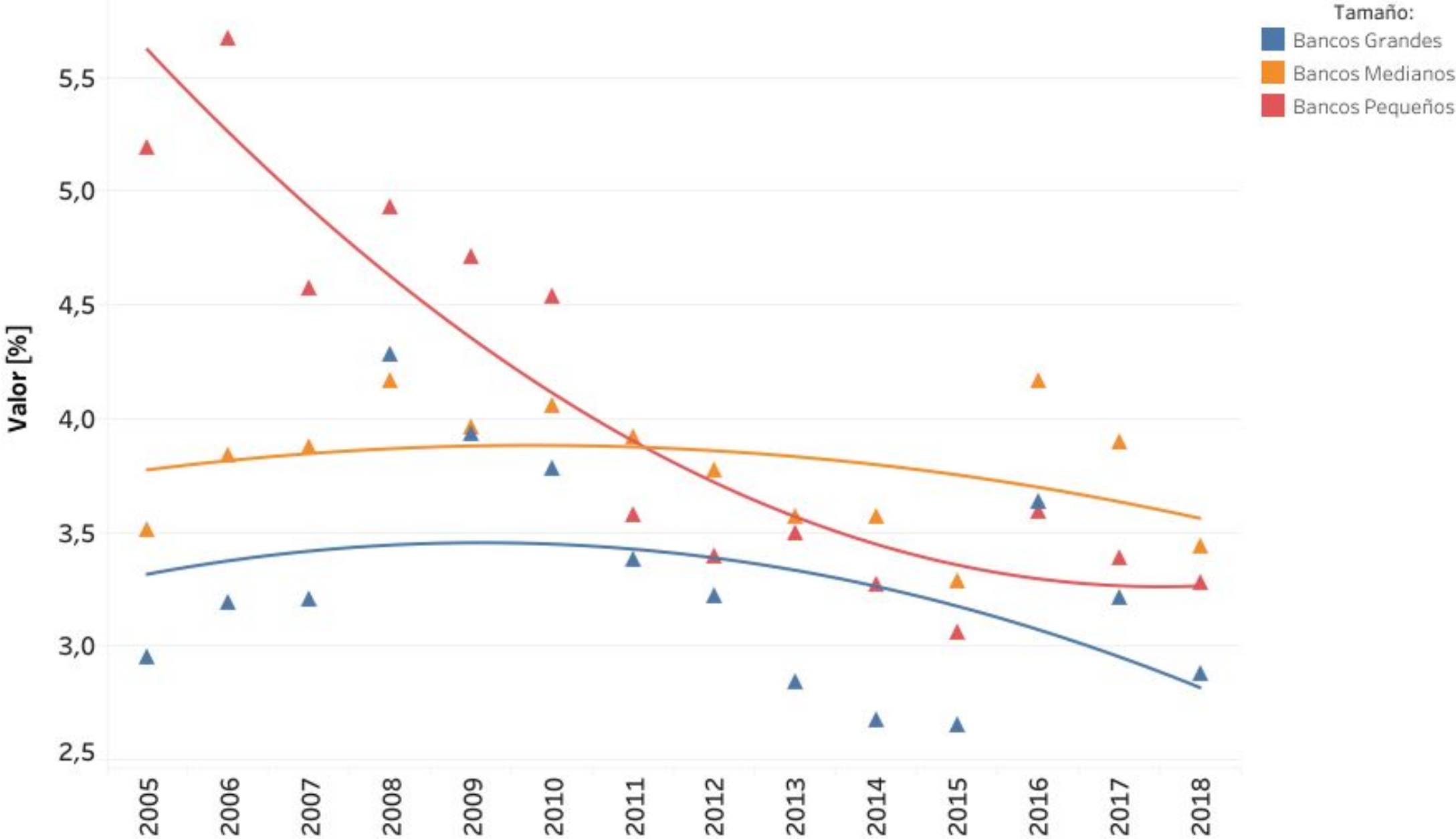


Análisis Descriptivo: Morosidad y Liquidez

Evolución anual por tamaño de banco



Evolución anual por tamaño de banco



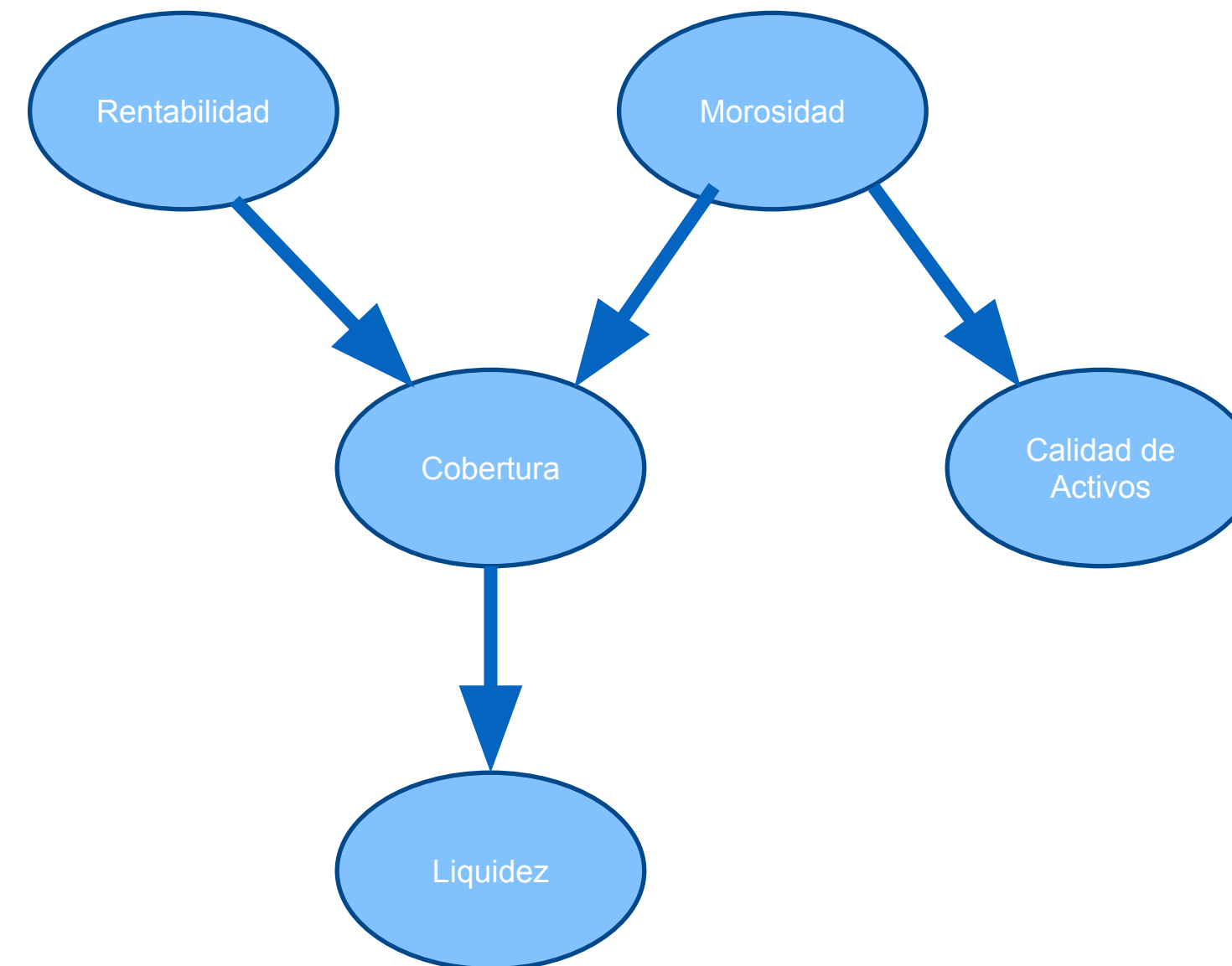
Propuesta de modelado:
Redes bayesianas dinámicas gaussianas

Definición de red bayesiana

Las redes bayesianas son:

- Un grafo acíclico dirigido (DAG) G donde cada nodo representa las variables aleatorias X_1, \dots, X_n
- Las aristas representan relaciones de dependencia condicional entre las variables

En el caso de variables continuas se emplean redes bayesianas gaussianas, donde cada variable X se representa como una normal con media y varianza influenciada por sus padres. En este caso, la distribución conjunta se presenta como una gaussiana multivariante.



$$P(L, Co, R, M, Ca) = P(R)P(M)P(Ca|M)P(Co|M, R)P(L|Co)$$

Para modelar series temporales, se utilizan redes bayesianas dinámicas

Sea Y_1, \dots, Y_n un proceso estocástico, con $Y_i \sim N$

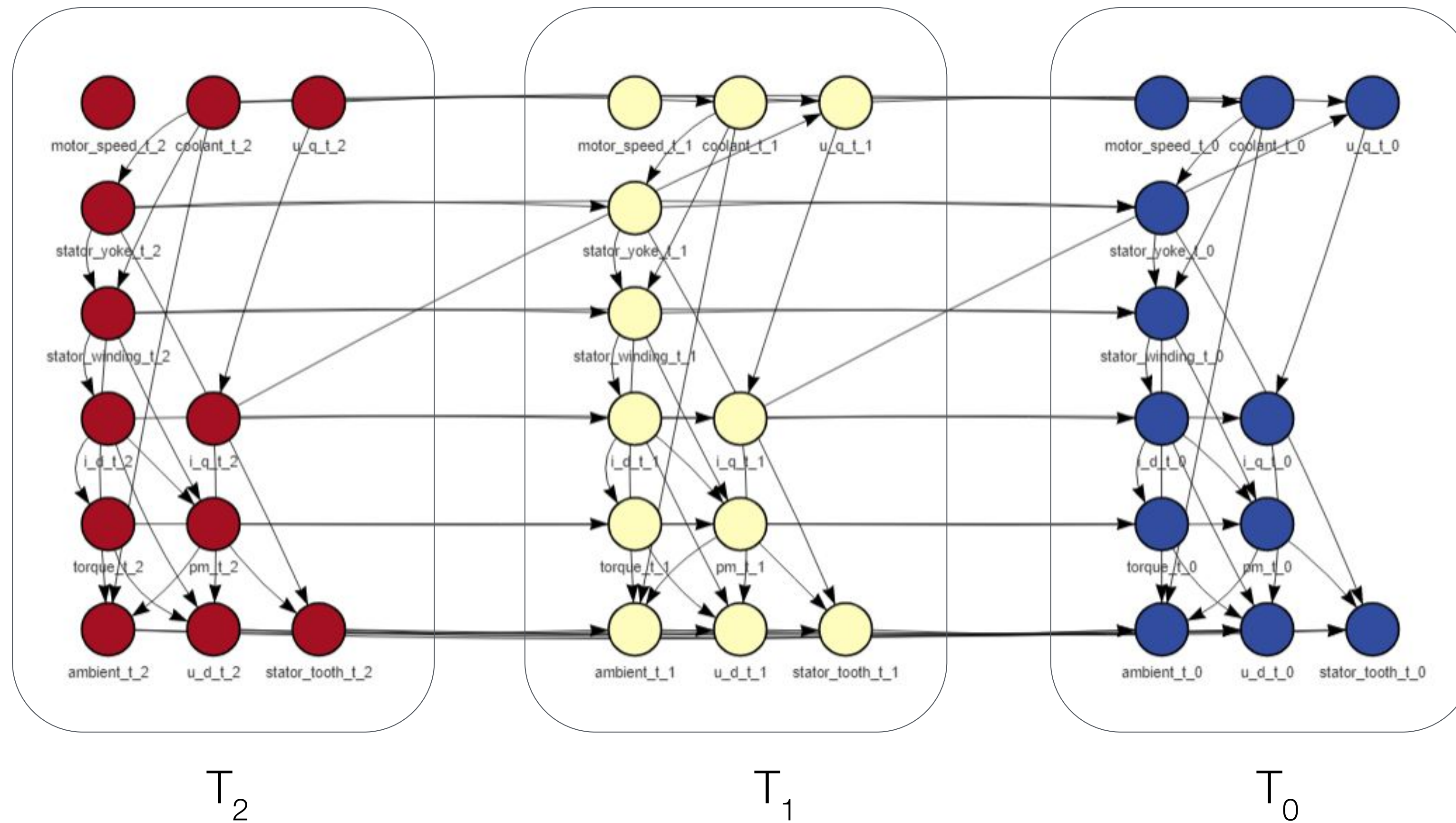
$$P(Y_{0:T}) = P(Y_0) \prod_{t=0}^{T-1} P(Y_{t+1} | Y_{0:t})$$

De esta forma, obtenemos redes separadas en instantes temporales (time slices), donde el presente es condicionalmente independiente del pasado dados n instantes anteriores. El orden markoviano de la red define este orden autorregresivo n .

$$(Y_{t+1} \perp Y_{0:t-1} | Y_t)$$



$$P(Y_{0:T}) = P(Y_0) \prod_{t=0}^{T-1} P(Y_{t+1} | Y_t)$$



Herramientas para BN y DBN

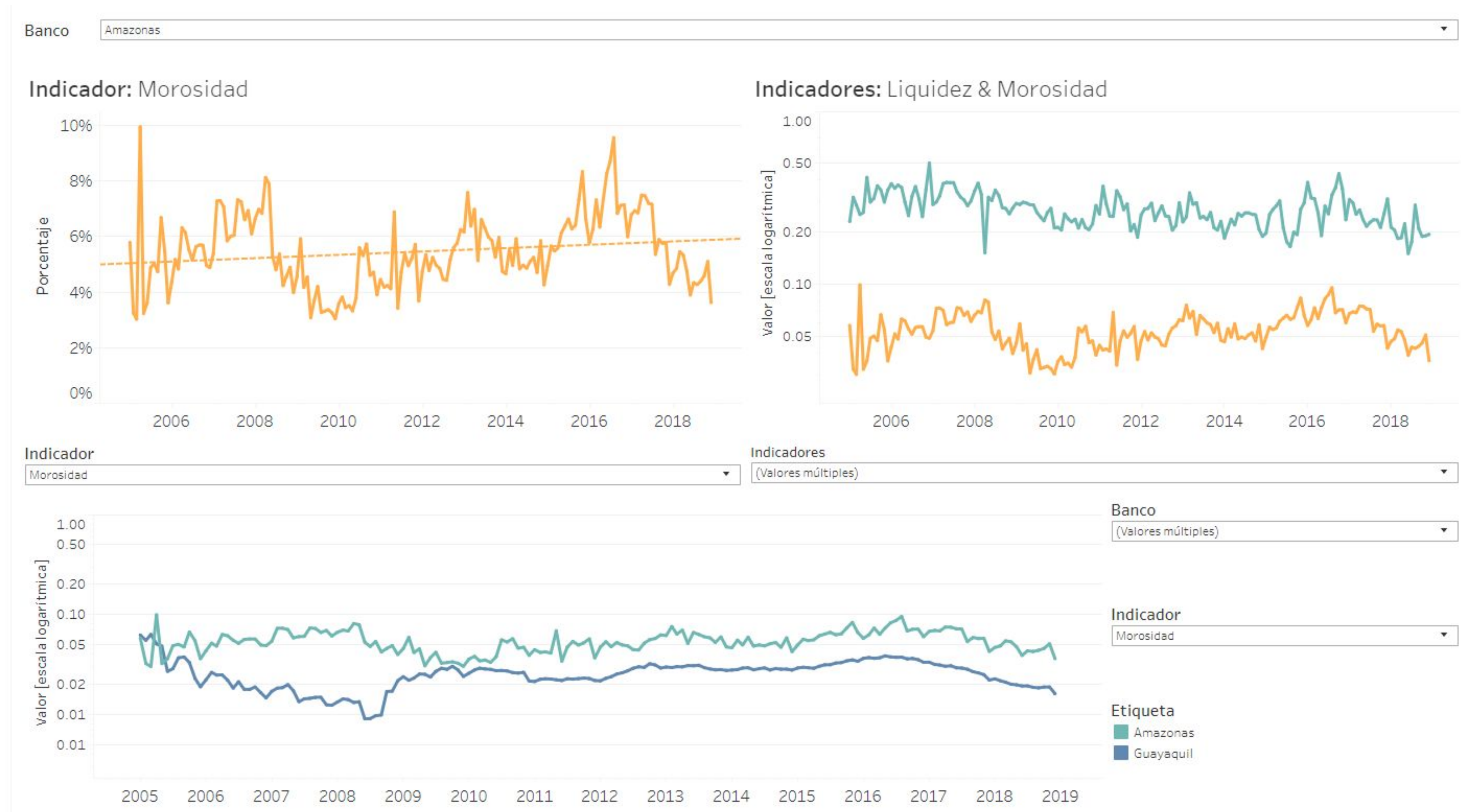
Para el uso de redes bayesianas hay múltiples paquetes disponibles tanto en R como en otros lenguajes. En concreto, en R recomendamos el uso del paquete *bnlearn* de Marco Scutari ya que:

- Ofrece soporte para redes discretas, gaussianas y mixtas
- Ofrece una amplia variedad de métodos para el aprendizaje de la estructura de la red
- Permite varios tipos de inferencia
- Su implementación parcial en C ofrece una buena eficiencia

De cara al uso de redes bayesianas dinámicas, hemos implementado el paquete *dbnR* basándonos en *bnlearn* y extendiendo su funcionalidad para:

- Aprender la estructura de redes bayesianas dinámicas gaussianas de orden markoviano ≥ 2
- Realizar forecasting e inferencia tanto aproximada como exacta sobre estas redes
- Ofrecer una herramienta de visualización de redes dinámicas y estáticas por medio de *visNetwork*

Caso Práctico
Predicción de indicadores en entidades financieras
Ecuatorianas



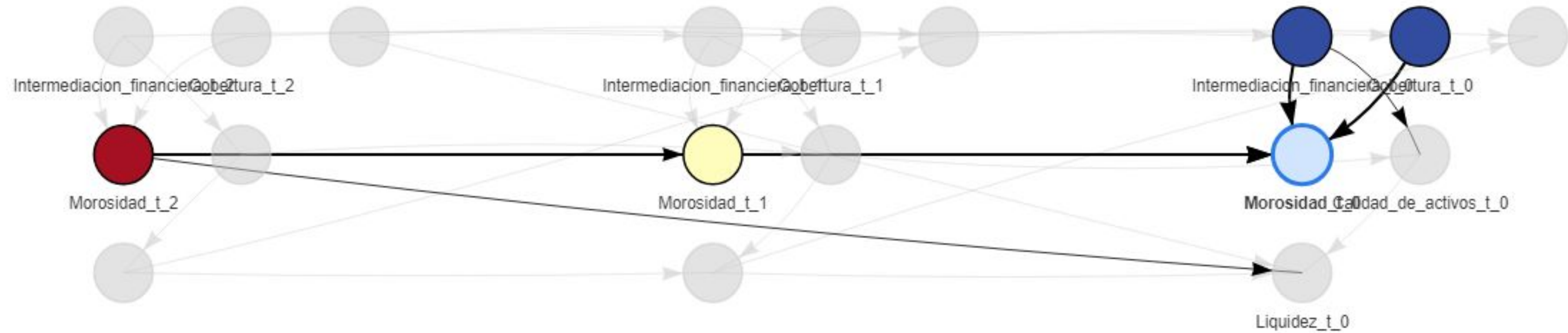
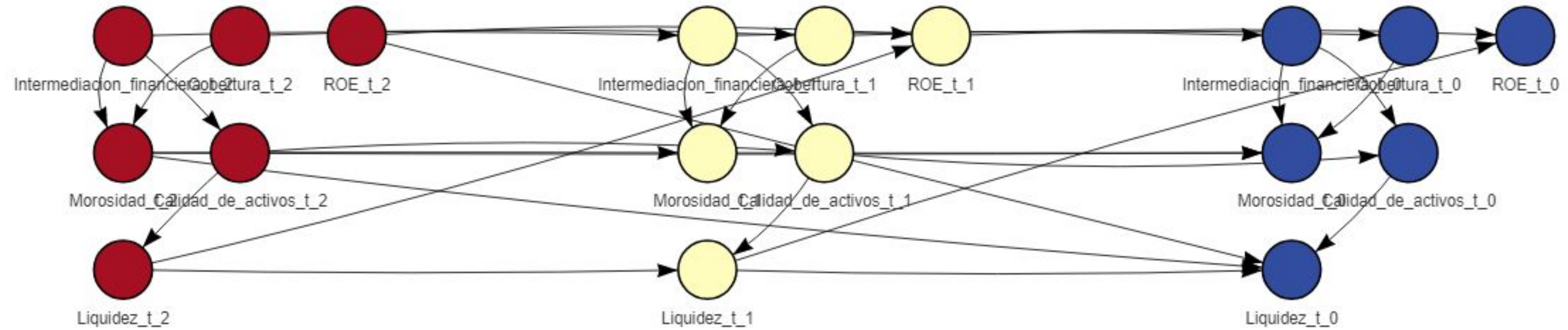
Aprendizaje de la red en 2 pasos: estructura y parámetros

Para construir la estructura de la red, primero se tiene en cuenta la relación entre las variables de manera estática y después la relación entre instantes temporales.

- Se ajusta la estructura que define la relación entre los índices econométricos en un mismo instante y frente al siguiente
- Este aprendizaje está contenido en *dbnR::learn_dbn_struct*

Una vez creada la estructura, hay que ajustar los parámetros de cada nodo. Para esto, se ajustan la media, varianza y pesos de cada variable por máxima verosimilitud con *dbnR::fit_dbn_params*

$$L(\hat{\boldsymbol{\theta}} : \mathcal{D}) = \max_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta} : \mathcal{D})$$

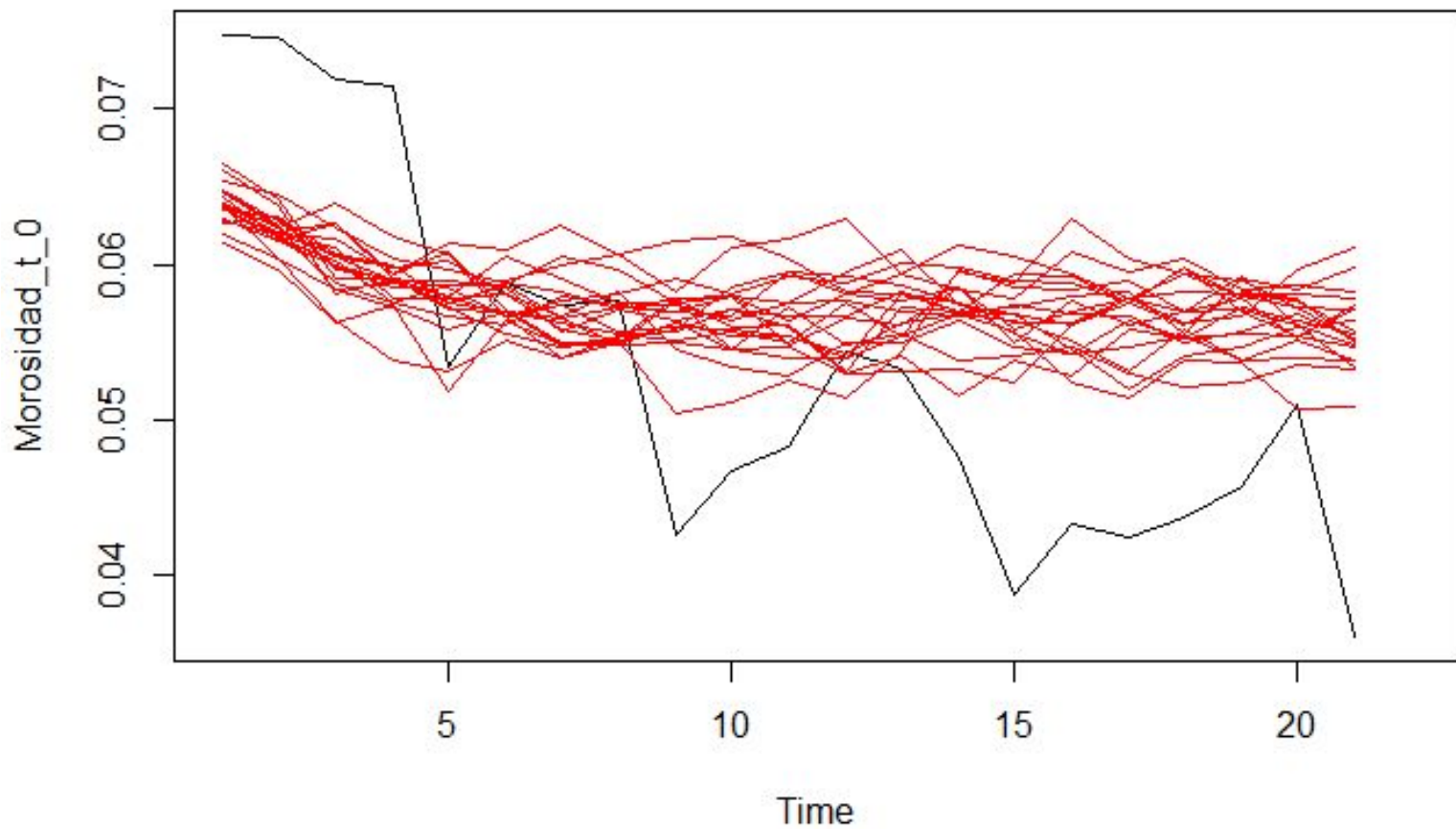
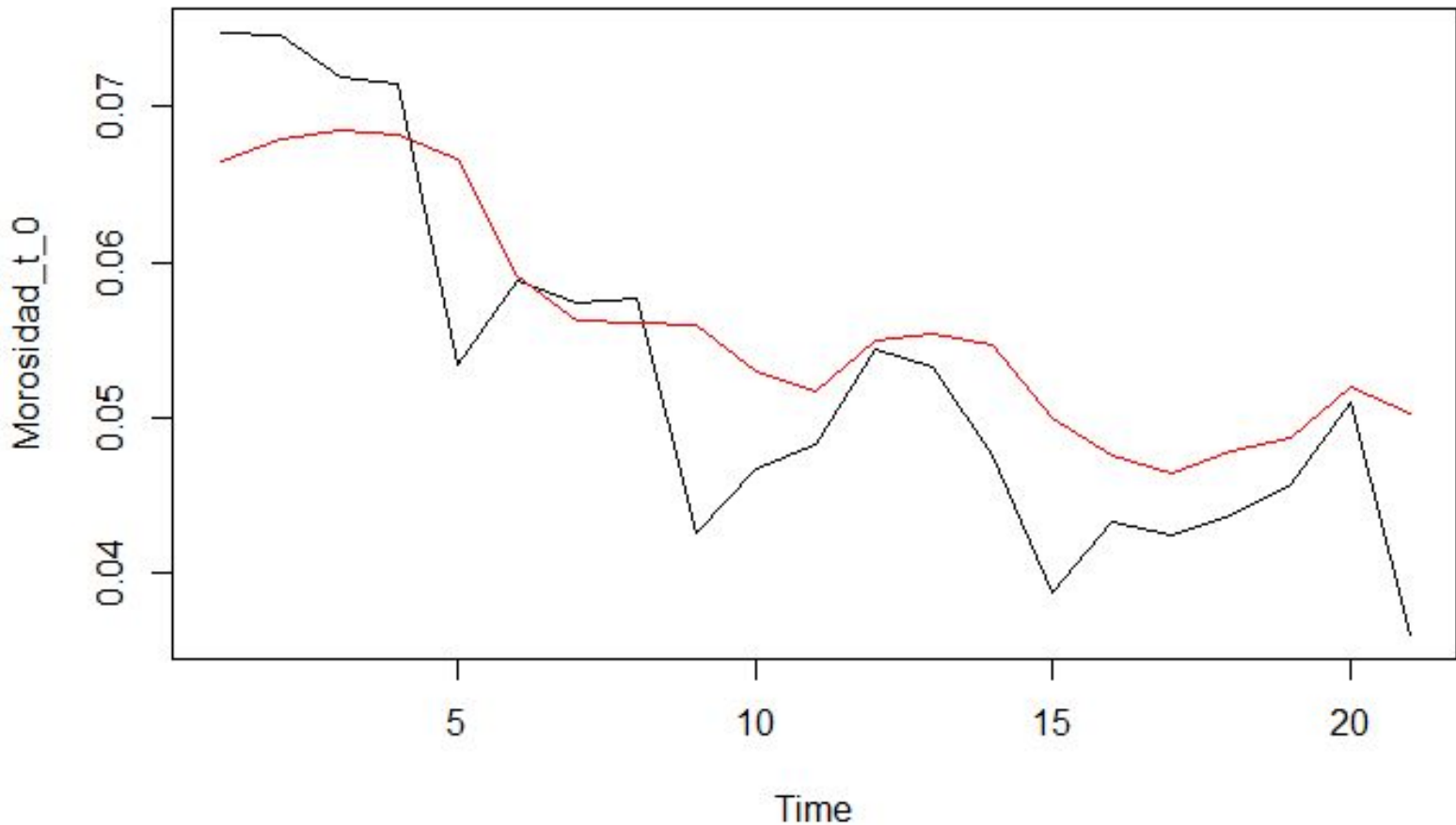
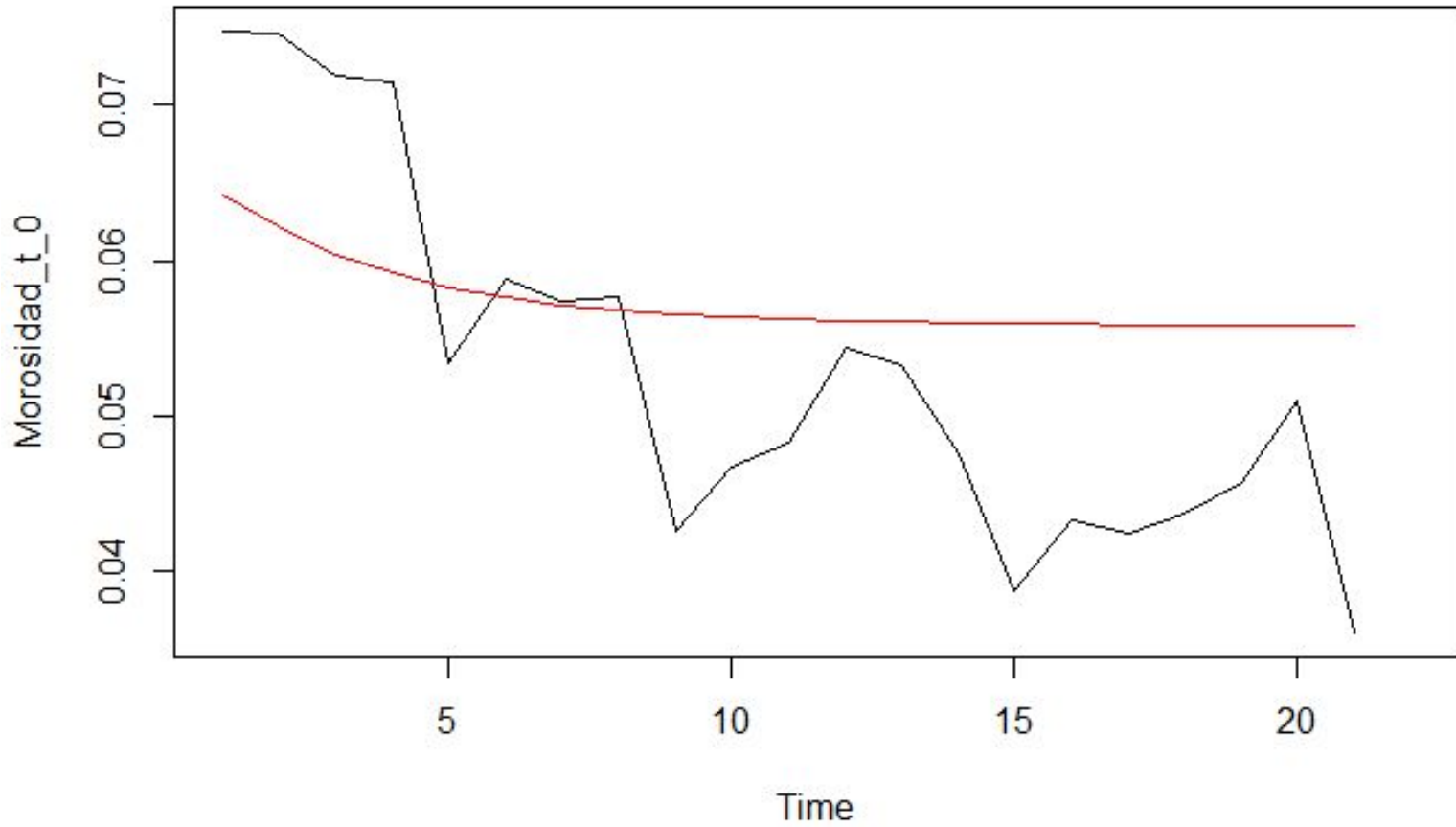


CASO PRÁCTICO

Resultados Multivariantes

Banco Amazonas

Para la predicción exacta, lo que se obtiene es la media a posteriori de la variable objetivo, pero hay que tener en cuenta la incertidumbre que se refleja en su varianza.

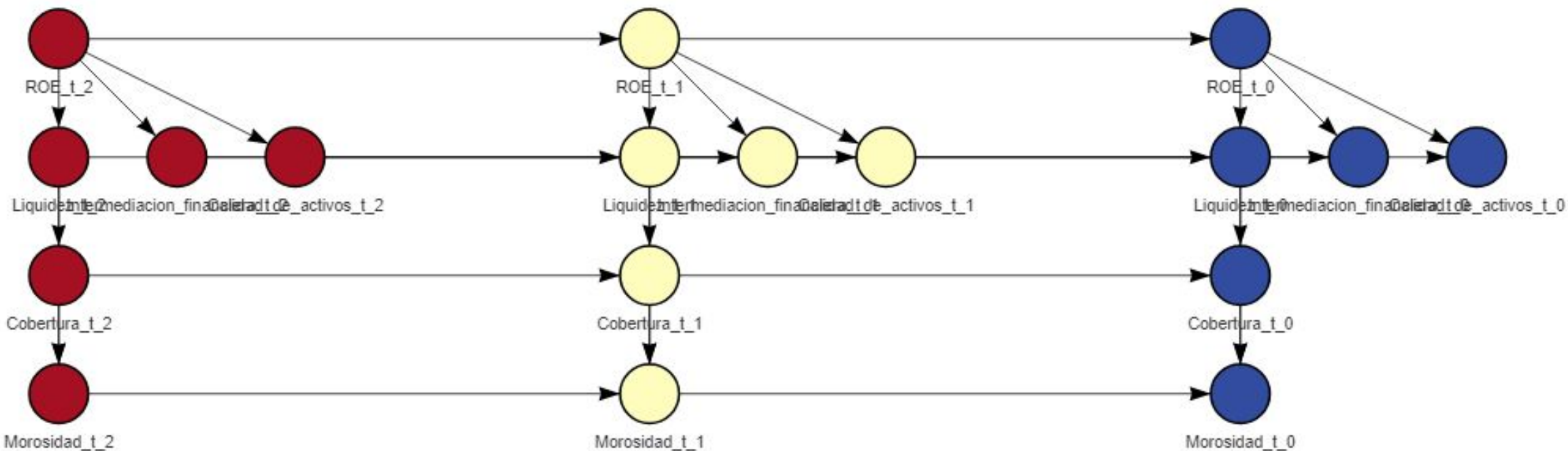
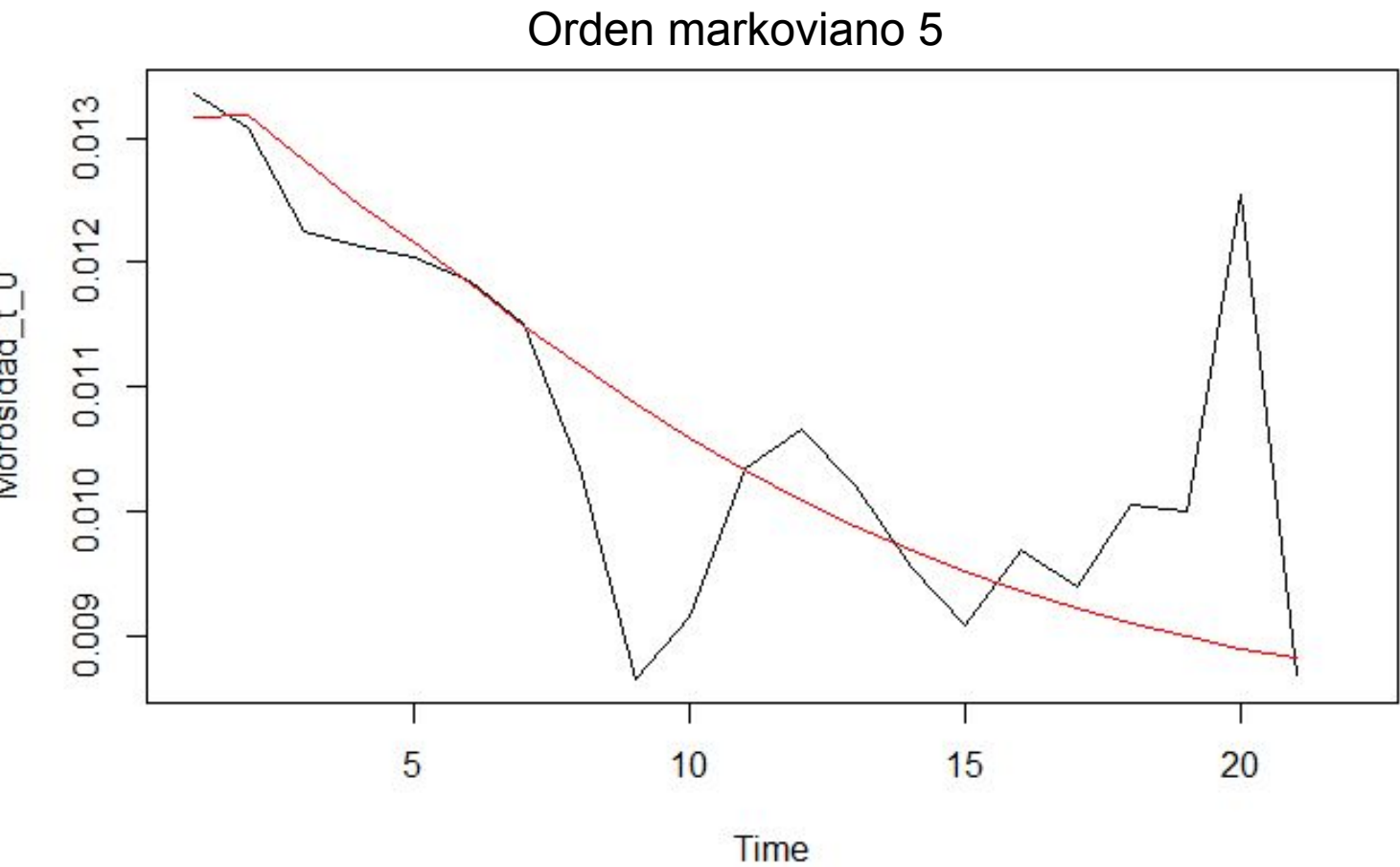
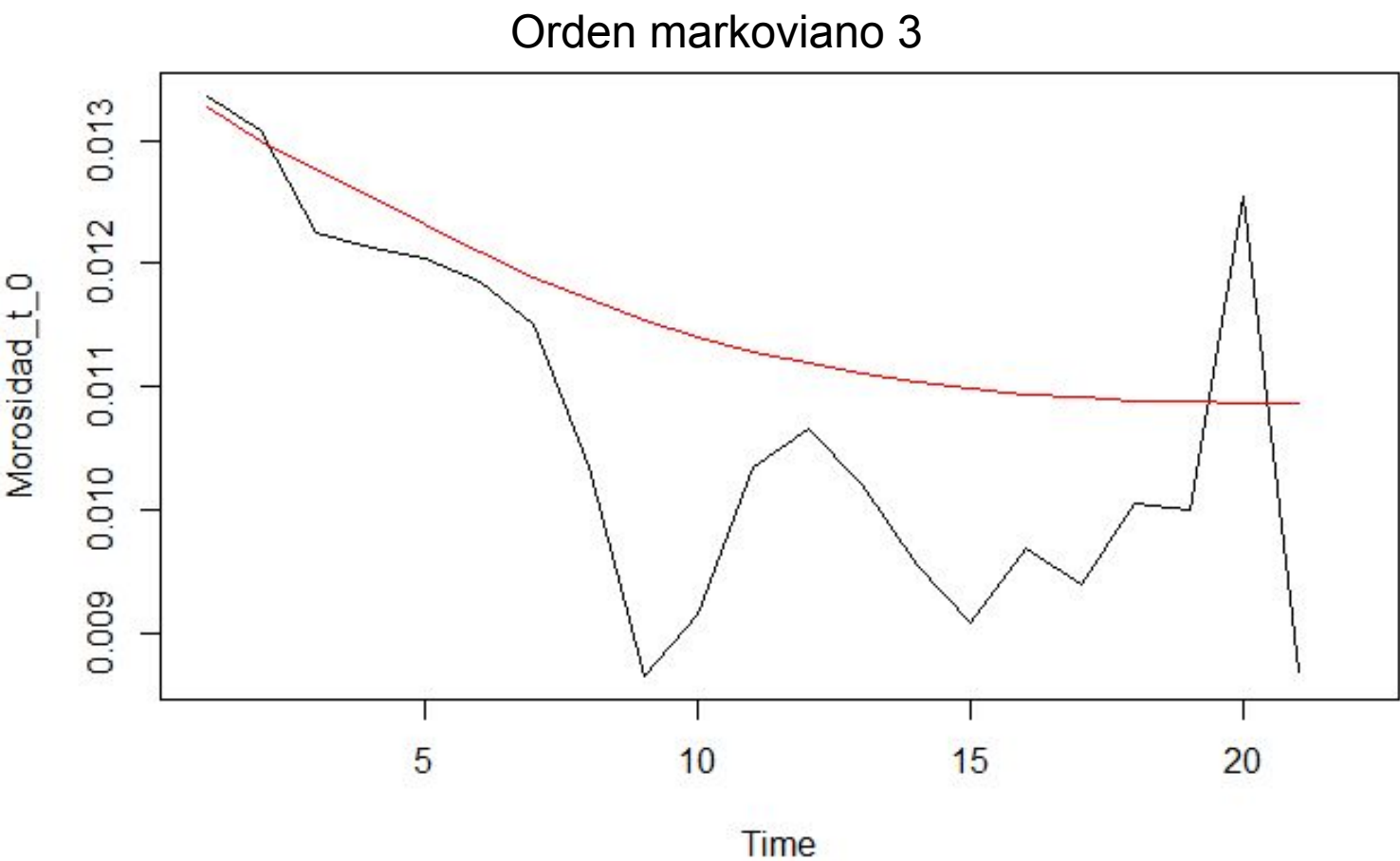


Morosidad	MAE	SD
Forecasting	9e-3	9.76e-3
Puntual	5.35e-3	6.29e-3

- Al hacer inferencia puntual, las predicciones son mes a mes y reflejan mejor el comportamiento inmediato de la serie

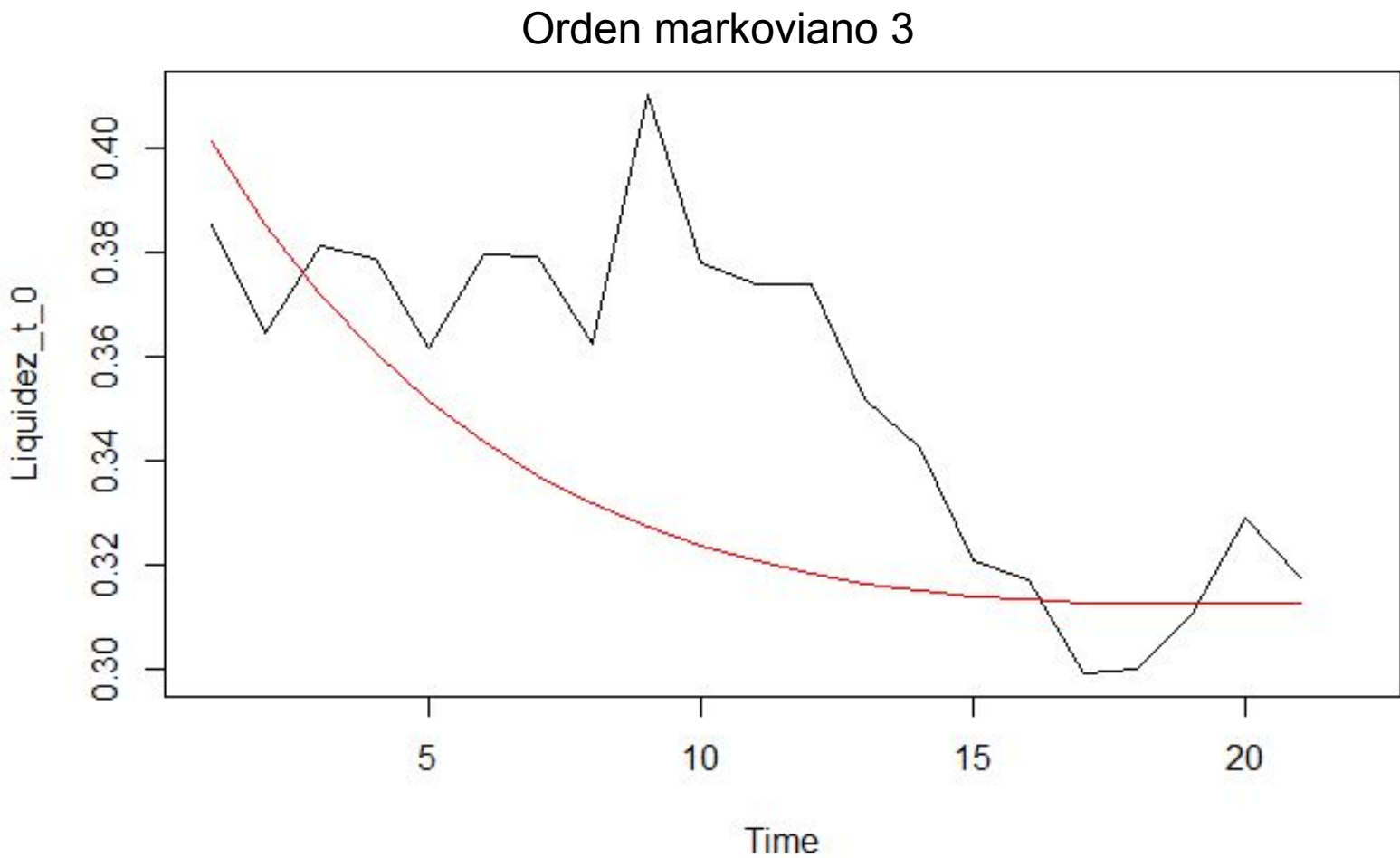
- Cuando usamos la inferencia aproximada, hay que lanzar varias repeticiones del sampleo para tener una idea de la varianza esperada en el forecasting

Banco Bolivariano



Morosidad	MAE	SD
Ord(3)	11e-4	10e-4
Ord(5)	6e-04	10e-4

Liquidez	MAE	SD
Ord(2)	0.0224	0.0263
Ord(3)	0.0262	0.0275



Conclusiones

- Las DBNs presentan una alternativa poco explorada en el análisis de series temporales multivariantes.
- La propia estructura de la red ofrece información sobre el funcionamiento del modelo y de las relaciones presentes entre las variables en los datos.
- Las redes permiten hacer estimaciones tanto puntuales como a largo plazo. En ambos casos, se obtiene una distribución de probabilidad sobre la predicción, con una media y una varianza esperadas sobre ella.
- Es posible predecir varios índices al mismo tiempo sobre un mismo banco, pudiendo verse su evolución conjunta y la relación entre ellas.

Referencias

- *bnlearn*: bnlearn.com
- *dbnR*: github.com/dkesada/dbnR
- Data World Bank: data.worldbank.org