



Seminario Permanente de Formación en Inteligencia Artificial Aplicada a la Defensa



Regresión y Series temporales

Julián Luengo

**Instituto Andaluz de Investigación en Data Science
and Computational Intelligence (DaSCI)**

Dpto. Ciencias de la Computación e I.A.

Universidad de Granada

julianlm@decsai.ugr.es

<http://sci2s.ugr.es>



DaSCI



**UNIVERSIDAD
DE GRANADA**

Resumen

1. Regresión
2. Series Temporales
3. Comentarios Finales

Resumen

1. Regresión

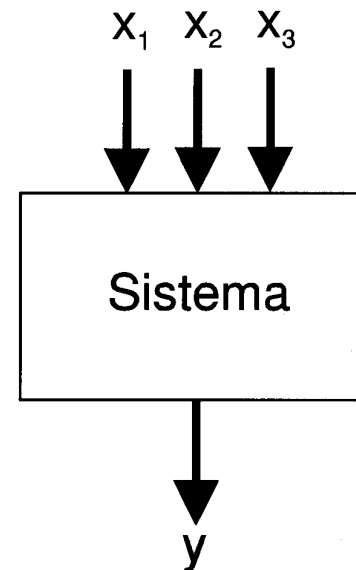
2. Series Temporales

3. Comentarios Finales

1. Definición del problema

- Objetivo: predecir el valor numérico para una variable a partir de los valores para otras
- La definición del problema es parecida a la del problema de clasificación: tenemos variables predictoras y una variable de regresión que en este caso es numérica
- En regresión la mayoría (e incluso todas) las variables predictoras son numéricas

El problema fundamental de la predicción está en modelar la relación entre las variables de estado para obtener el valor de la variable a predecir.



1. Definición del problema

- Objetivo: predecir el valor numérico para una variable a partir de los valores para otras
- La definición del problema es parecida a la del problema de clasificación: tenemos variables predictoras y una variable de regresión que en este caso es numérica
- En regresión la mayoría (e incluso todas) las variables predictoras son numéricas
- **Ejemplos:**
 - **¿qué consumo tendrá un coche en autovía en función de su peso, cilindrada, potencia,...?**
 - **¿qué número de artículos tendremos para el próximo pedido?**
 - **¿cuántos meses necesitaremos para desarrollar un proyecto software?**
 - **¿cuál es la probabilidad de que un cliente determinado sea receptivo a un envío publicitario?**
 - **¿cuántos enfermos tendremos en urgencias la próxima nochebuena?**

2. Validación en algoritmos de regresión

- Todas las técnicas de validación estudiadas en clasificación son válidas para predicción numérica
- La diferencia está en que ahora debemos medir el error de otra forma
- Debemos medir el error cometido al aproximar un conjunto de valores $\{v_1, \dots, v_n\}$ por su estimación $\{v'_1, \dots, v'_n\}$

Error cuadrático medio (ECM)

$$ECM = \frac{\sum_{i=1}^n (v_i - v'_i)^2}{n}$$

ECM estandarizado (ECME)

$$ECME = \sqrt{\frac{\sum_{i=1}^n (v_i - v'_i)^2}{n}}$$

Error medio absoluto (EMA)

$$EMA = \frac{\sum_{i=1}^n |v_i - v'_i|}{n}$$

Error absoluto relativo (EAR)

$$EAR = \frac{\sum_{i=1}^n |v_i - v'_i|}{\sum_{i=1}^n |v_i - \bar{v}|}$$

Coeficiente de correlación $r_{vv'} = \frac{\sum_{i=1}^n (v_i - \bar{v})(v'_i - \bar{v}')}{(n-1)\sigma_v\sigma_{v'}}$

3. Técnicas de clasificación válidas para regresión: kNN, RNN

- **Métodos basados en ejemplos/instancias:**

Al utilizar kNN, si los k vecinos más próximos $\{e_1, \dots, e_k\}$ tienen valores $\{v_1, \dots, v_k\}$ para la variable objetivo, entonces el valor a devolver para el objeto analizado e' sería

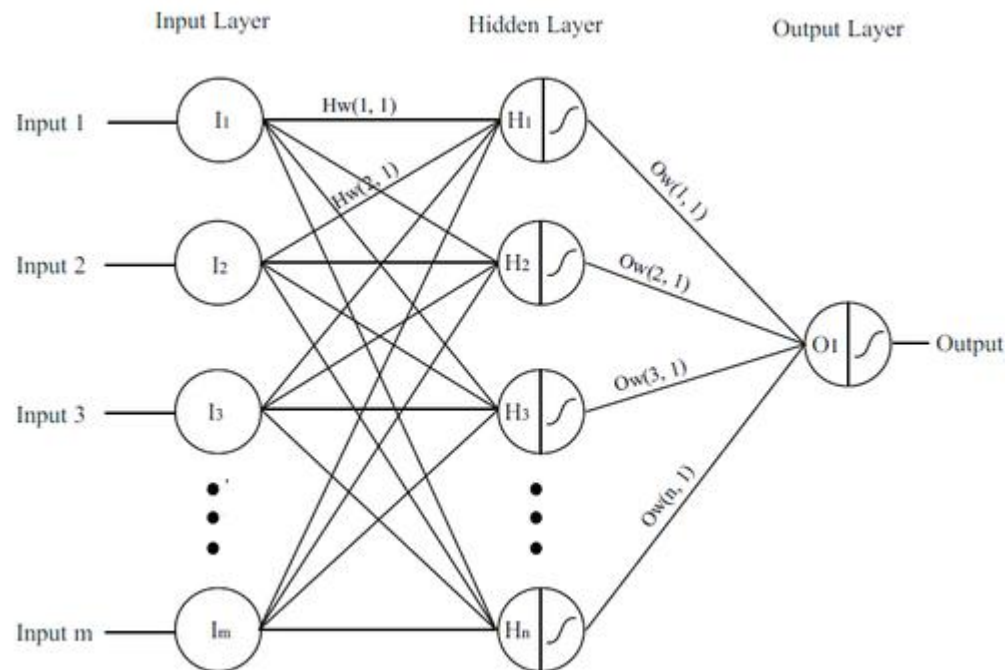
$$v = \begin{cases} \frac{\sum_{i=1}^k v_i}{k} & \text{si todos cuentan igual} \\ \frac{\sum_{i=1}^k w_i \cdot v_i}{\sum_{i=1}^k w_i} & \text{si se hace un voto ponderado} \end{cases}$$

por ejemplo, con $w_i = 1/d(e_i, e')$

3. Técnicas de clasificación válidas para regresión: kNN, RNN

■ Métodos basados en redes neuronales:

- La capa de salida sería una única neurona
- Como los pesos se adaptan en función del error cometido, es suficiente con medir de forma adecuada el error



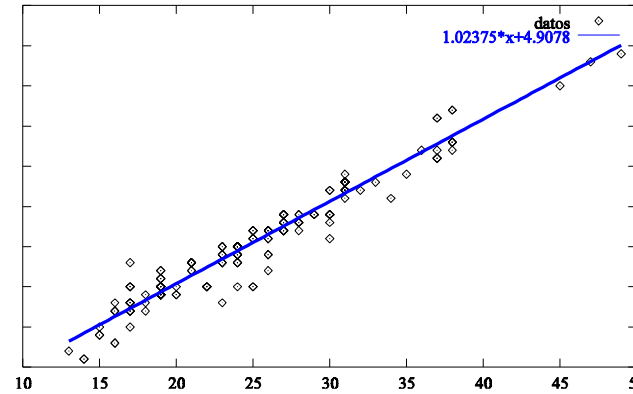
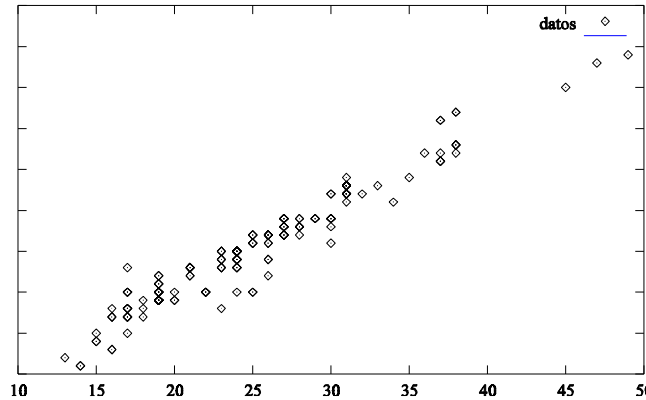
4. Análisis de regresión

- El análisis de regresión es el método más utilizado para realizar la tarea de predicción numérica
- **Objetivo:** estimar la variable objetivo (y) como una ecuación que contiene como incógnitas al resto de las variables ($x_1, \dots, x_n$)
- El modelo más sencillo es la **regresión lineal** que reducida a una sola variable predictora tiene la forma:
$$y = a + b \cdot x$$
- Estos coeficientes pueden obtenerse fácilmente mediante el método de los mínimos cuadrados

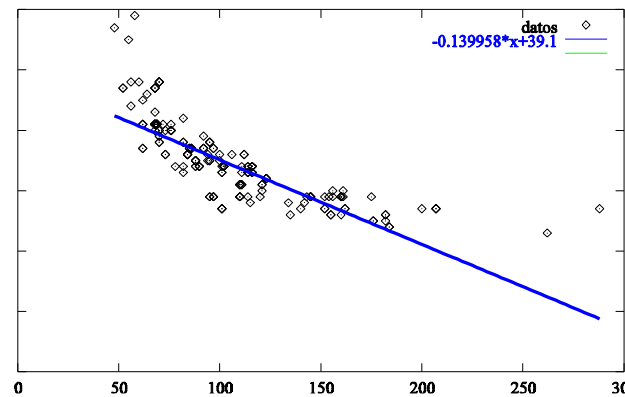
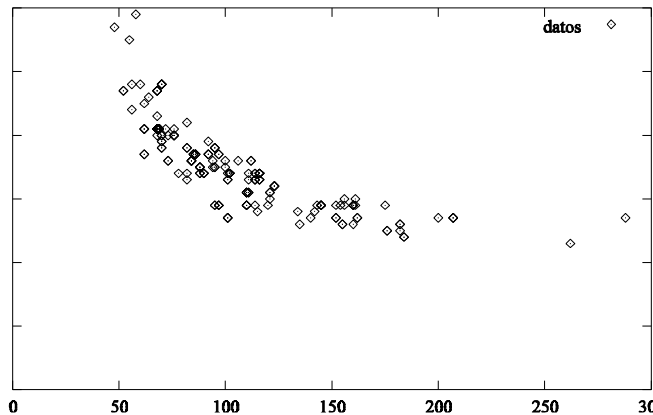
$$b = \frac{\sum_{i=1}^s (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^s (x_i - \bar{x})^2}$$

$$a = \bar{y} - b \cdot \bar{x}$$

4. Análisis de regresión



Bastante apropiado



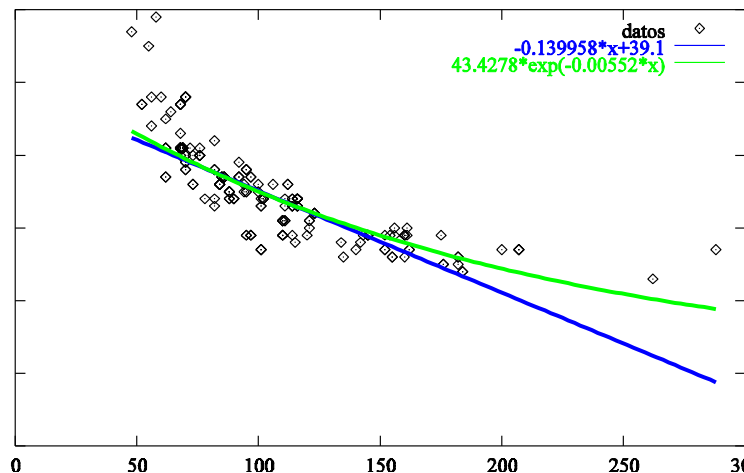
No tan apropiado

4. Análisis de regresión

- Para estimar curvas es necesario utilizar otra regresión, por ejemplo, **regresión exponencial**: $y = a \cdot e^{bx}$
- ¿Cómo estimamos ahora a y b ? Tomando logaritmos

$$\ln(y) = \ln(a \cdot e^{bx}) \Rightarrow \ln(y) = \ln(a) + \ln(e^{bx}) \Rightarrow y^* = a^* + bx$$

- Es decir, tenemos un problema de regresión lineal entre $y^* = \ln(y)$ y x . Una vez estimados a^* y b podemos calcular $a = e^{a^*}$



4. Análisis de regresión

Regresión lineal múltiple

- Cuando hay más de una variable predictora, la ecuación de predicción se transforma en

$$y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n$$

Este problema se conoce como regresión lineal múltiple

- La estimación de los coeficientes es algo más compleja y requiere operar con matrices (y sus inversas)
- Por ejemplo, continuando con el ejemplo CPU.arff, tendríamos

$$\text{Class} = 0.066 * \text{MYCT} + 0.0143 * \text{MMIN} + 0.0066 * \text{MMAX} + 0.4945 * \text{CACH} - 0.1723 * \text{CHMIN} + 1.2012 * \text{HMAX} - 66.4814$$

$$\text{EAM} = 34.31 \%$$

$$\text{EAR} = 39.26 \%$$

- Existen técnicas más complejas de regresión que aproximan de forma más precisa los datos de entrada

Regresión no-lineal, regresión logística, ...

5. Árboles de regresión

- Un **árbol de regresión** es un árbol de decisión cuyas hojas predicen una cantidad numérica
- Ese valor numérico se calcula como la media del valor para la variable clase de todos los ejemplos que han llegado a esa hoja durante el proceso de construcción del árbol
- La evaluación de un nuevo ejemplo es idéntico a los árboles de decisión
- Durante el proceso de predicción es posible utilizar un suavizado de los valores del ejemplo a tratar, con el fin de salvar las posibles discontinuidades presentes en los datos
- El criterio de selección de una variable en la construcción del árbol está basado en una reducción del error esperado: reducción de la desviación/varianza en la variable objetivo (standard reduction deviation):

$$SDR = sd(T) - \sum_i \frac{|T_i|}{|T|} \cdot sd(T_i)$$

SDR=Standard Deviation Reduction

- Al final el árbol se poda para evitar el sobreajuste

5. Árboles de regresión

```
MMAX <= 14000 :
|  CACH <= 8.5 :
|  |  MMAX <= 6100 : LM1 (75/5.4%)
|  |  MMAX > 6100 :
|  |  |  MYCT <= 83.5 :
|  |  |  |  MMAX <= 10000 : LM2 (8/4.5%)
|  |  |  |  MMAX > 10000 : LM3 (3/3.78%)
|  |  |  MYCT > 83.5 : LM4 (22/6.73%)
|  CACH > 8.5 :
|  |  CHMIN <= 7 :
|  |  |  MYCT <= 95 : LM5 (7/13.3%)
|  |  |  MYCT > 95 :
|  |  |  |  CACH <= 28 : LM6 (12/7.85%)
|  |  |  |  CACH > 28 : LM7 (6/7.27%)
|  |  CHMIN > 7 : LM8 (8/45%)
MMAX > 14000 :
|  MMAX <= 22500 :
|  |  CACH <= 27 :
|  |  |  CHMIN <= 5 : LM9 (14/6.92%)
|  |  |  CHMIN > 5 : LM10 (5/12.8%)
|  |  CACH > 27 : LM11 (18/25.1%)
```

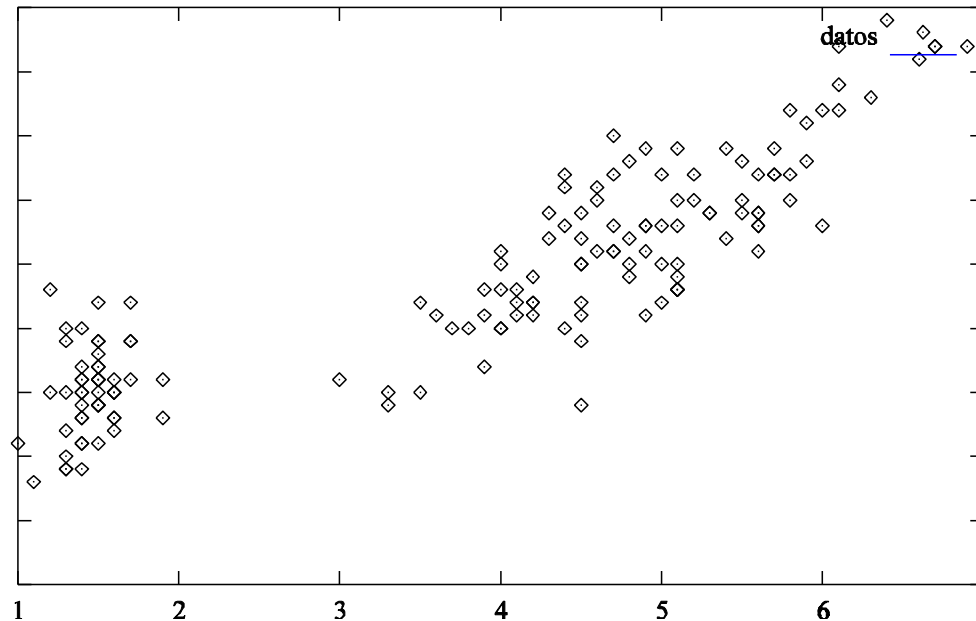
```
LM1: class = 24.3
LM2: class = 45.3
LM3: class = 62.7
LM4: class = 39.6
LM5: class = 63.6
LM6: class = 43.9
LM7: class = 55
LM8: class = 104
LM9: class = 75.4
LM10: class = 94.4
LM11: class = 128
LM12: class = 262
LM13: class = 174
LM14: class = 355
LM15: class = 971
```

```
EAM    18.1275
EAR    20.7449%
```

WEKA: sin suavizar
poda=1

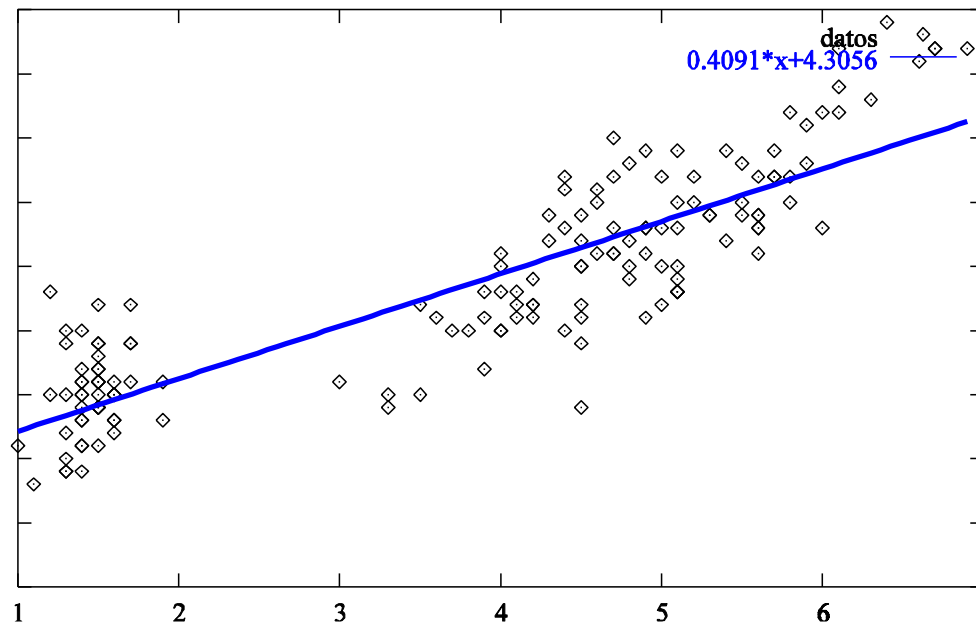
6. Árboles de modelos

- Son **árboles de regresión** en los que la poda se realiza en mayor medida y en las **hojas** en lugar de un valor numérico **contienen una ecuación de regresión local** a esa partición del espacio



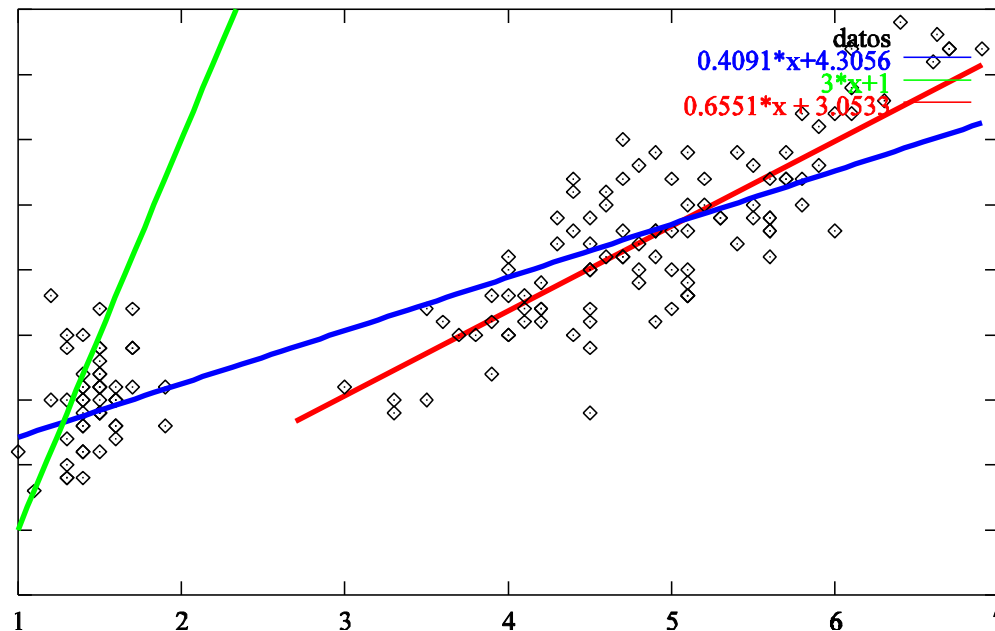
6. Árboles de modelos

- Son árboles de regresión en los que la poda se realiza en mayor medida y en las hojas en lugar de un valor numérico contienen una ecuación de regresión local a esa partición del espacio

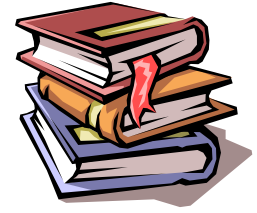


6. Árboles de modelos

- Son árboles de regresión en los que la poda se realiza en mayor medida y en las hojas en lugar de un valor numérico contienen una ecuación de regresión local a esa partición del espacio

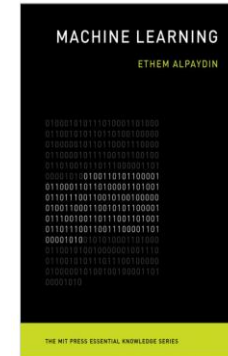


Bibliografía



Alpaydin, Ethem. Machine learning: the new AI. MIT press, 2016.

Alpaydin, Ethem. Introduction to machine learning. MIT press, 2020.



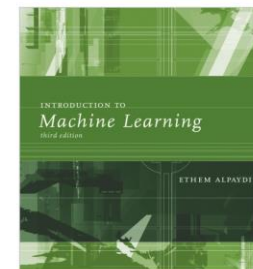
From [The MIT Press Essential Knowledge series](#)

Machine Learning

The New AI

By [Ethem Alpaydin](#)

A concise overview of machine learning – computer programs that learn from data – which underlies applications that include recommendation systems, face recognition, and driverless cars.



From [Adaptive Computation and Machine Learning series](#)

Introduction to Machine Learning, Third Edition

By [Ethem Alpaydin](#)

A substantially revised third edition of a comprehensive textbook that covers a broad range of topics not often included in introductory texts.

Resumen

1. Regresión

2. Series Temporales

3. Comentarios Finales

Series temporales

- Predicción (forecasting)
- Herramientas básicas para predicción
- Regresión simple
- Regresión multivariada
- Descomposición de series temporales
- Modelos avanzados

Predicción - Forecasting

Definición

- **Forecasting**: Predecir el futuro con la mayor precisión posible, dada toda la información disponible, incluidos los datos históricos y el conocimiento de cualquier acontecimiento futuro que pueda afectar a las previsiones
- Por lo general, es una parte integral de la toma de decisiones.

Ejemplos

- Previsión de la demanda de electricidad: muy precisa
- Previsiones sobre los tipos de cambio: aproximadas



Factores que afectan al pronóstico

- Horizonte temporal

Marco temporal (¿Hasta dónde podemos predecir?)

a corto plazo (1 - 2 periodos)

a mediano plazo (5 - 10 períodos)

a largo plazo (más de 12 períodos)

- Tipos de patrones de datos

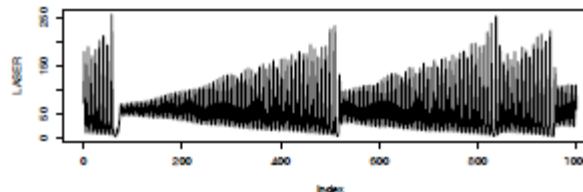
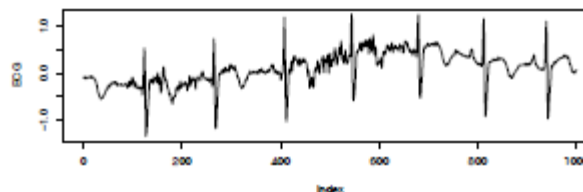
Forecasting cuantitativo

- Puede aplicarse cuando:
 - **Se dispone de datos numéricos sobre el pasado**
 - Es razonable suponer que algunos aspectos de las pautas pasadas continuarán en el futuro

Series temporales

- Todo lo que pueda ser medido a lo largo del tiempo es una serie temporal
- Series de tiempo observadas a intervalos regulares de tiempo (cada minuto, cada hora, diariamente, semanalmente, ...)

$$\{X_{t_1}, X_{t_2}, X_{t_3}, \dots, X_{t_n}\}$$



Forecasting de series temporales

- Los datos de las series temporales son útiles cuando se pronostica algo que cambia con el tiempo (por ejemplo, los precios de las acciones, las ventas, los beneficios, ...)
- La previsión de series temporales tiene por objeto estimar cómo continuará la secuencia de observaciones en el futuro

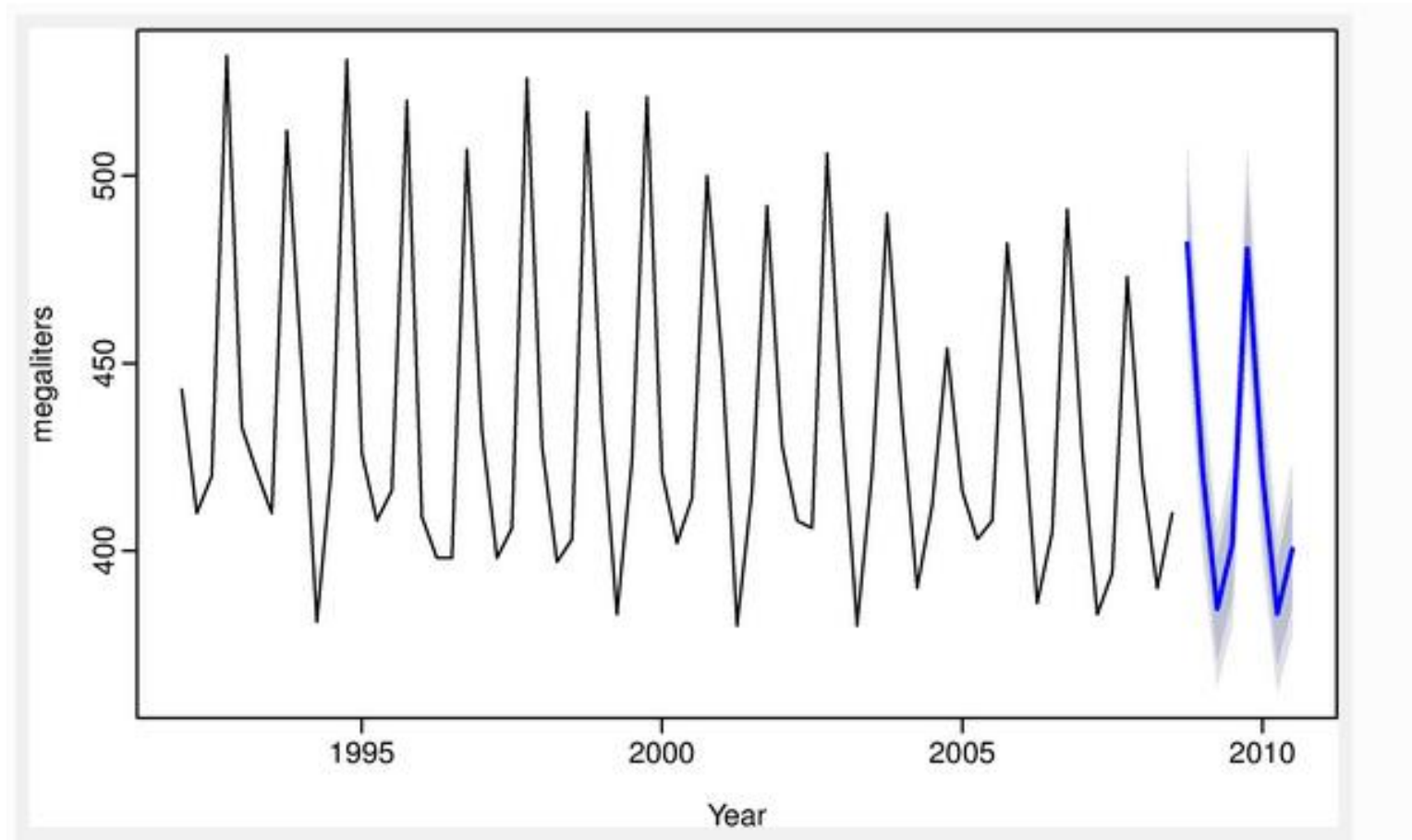
Marco temporal (¿Hasta dónde podemos predecir?)

a corto plazo (1 - 2 periodos)

a mediano plazo (5 - 10 períodos)

a largo plazo (más de 12 períodos)

Predicción de producción de cerveza



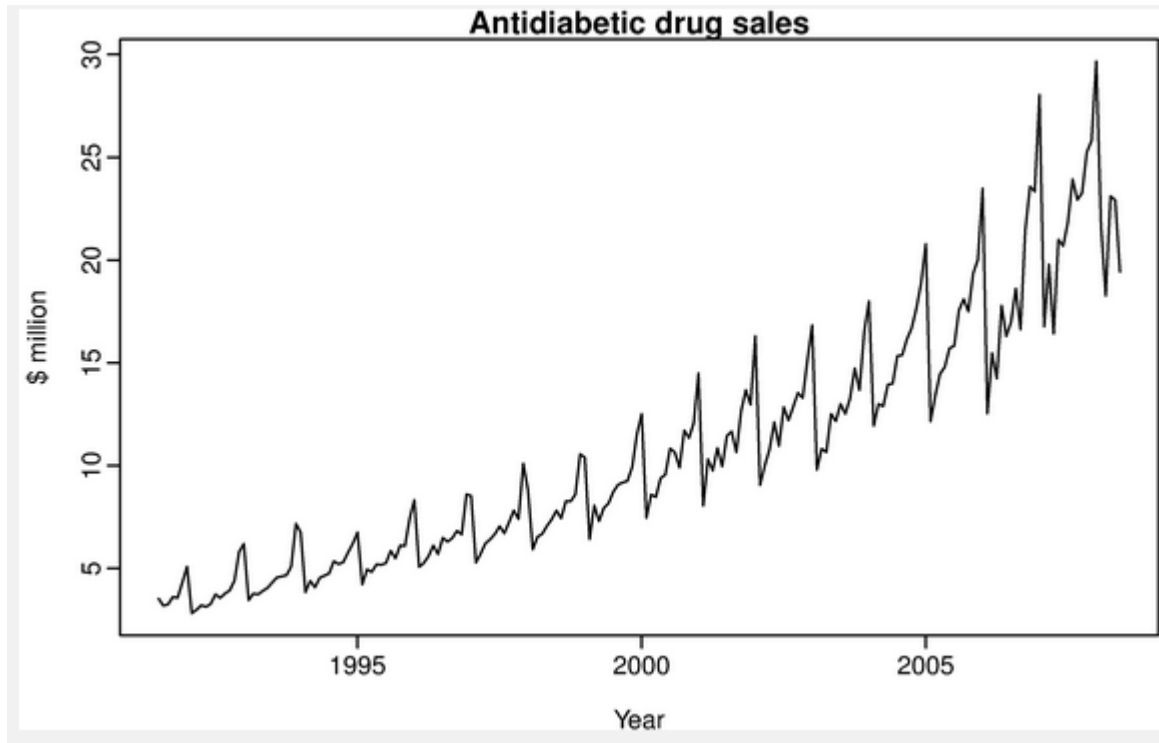
Series temporales

- Predicción (forecasting)
- Herramientas básicas para predicción
- Regresión simple
- Regresión multivariada
- Descomposición de series temporales
- Modelos avanzados

Gráfico en el tiempo

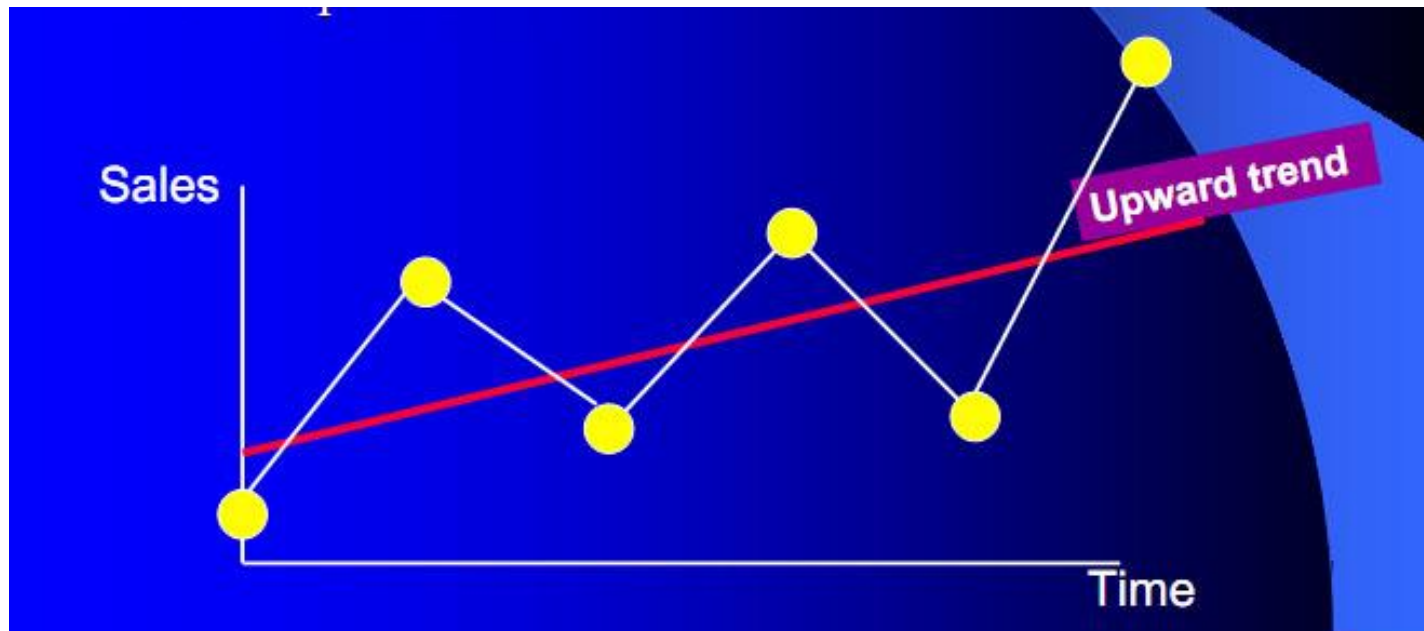


Gráfico en el tiempo, ej. 2



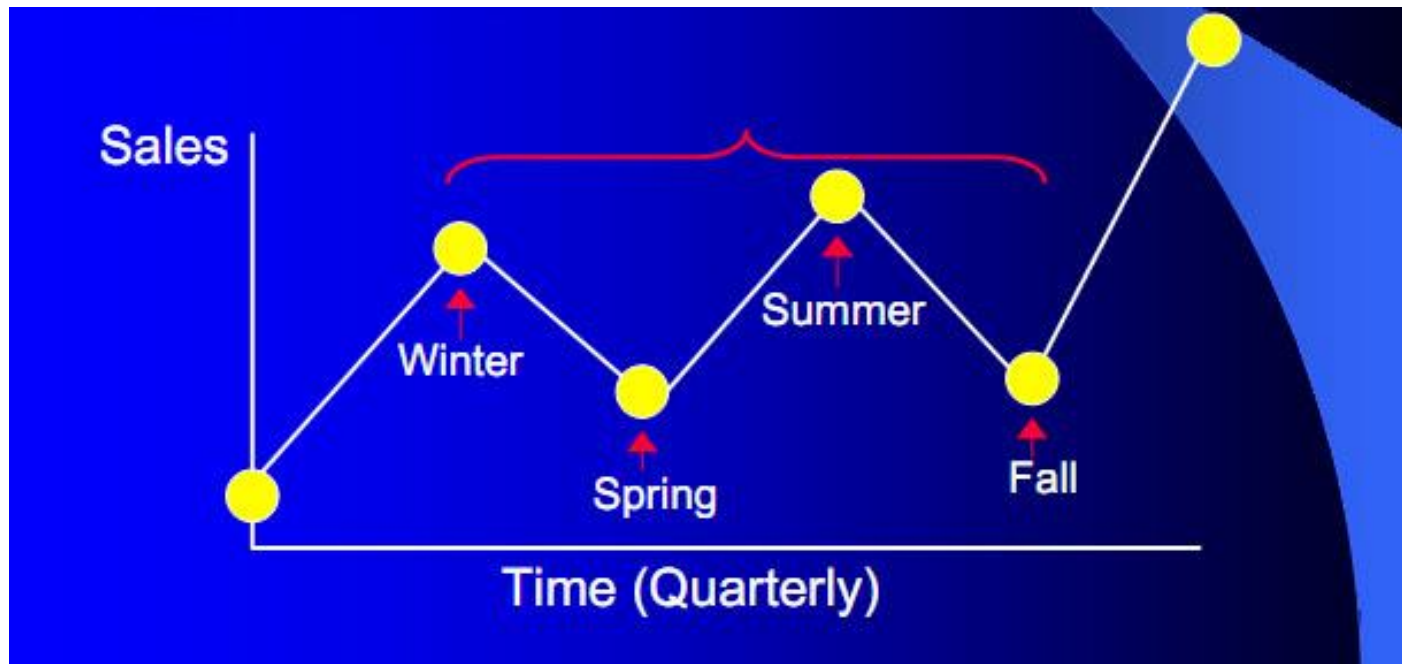
Tipos de patrones en series temporales

- *Tendencia*: aumento o disminución a largo plazo de los datos



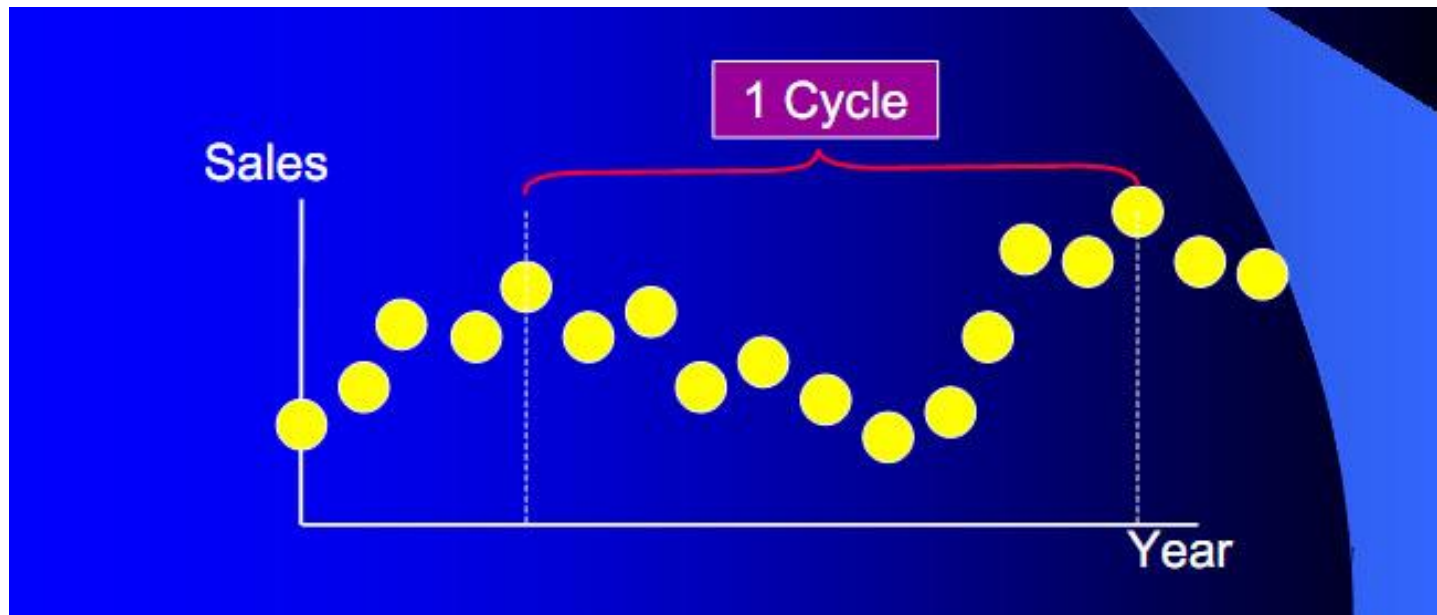
Tipos de patrones en series temporales

- *Patrón estacional*: datos afectados por factores estacionales como la época del año o el día de la semana



Tipos de patrones en series temporales

- *Ciclo*: los datos exhiben subidas y bajadas que no son de un período fijo; duración variable y desconocida



Resúmenes de datos numéricos

■ Estadísticas univariadas

- Promedio
- Mediana
- Percentiles
- Rango Inter cuartil (IQR)
- Desviación típica

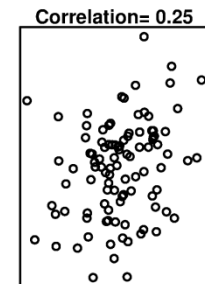
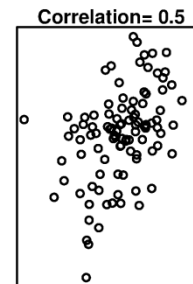
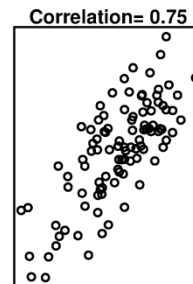
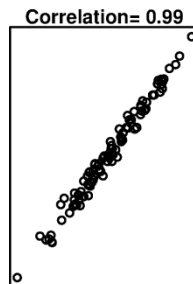
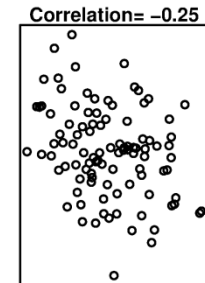
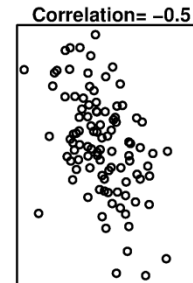
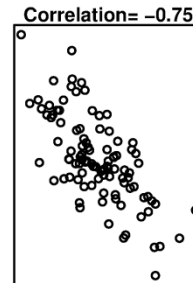
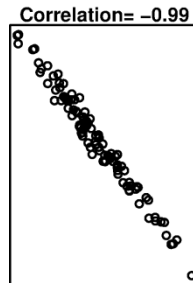
$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = (x_1 + x_2 + x_3 + \cdots + x_N)/N.$$

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}.$$

Coeficiente de correlación

Fuerza de la relación
lineal

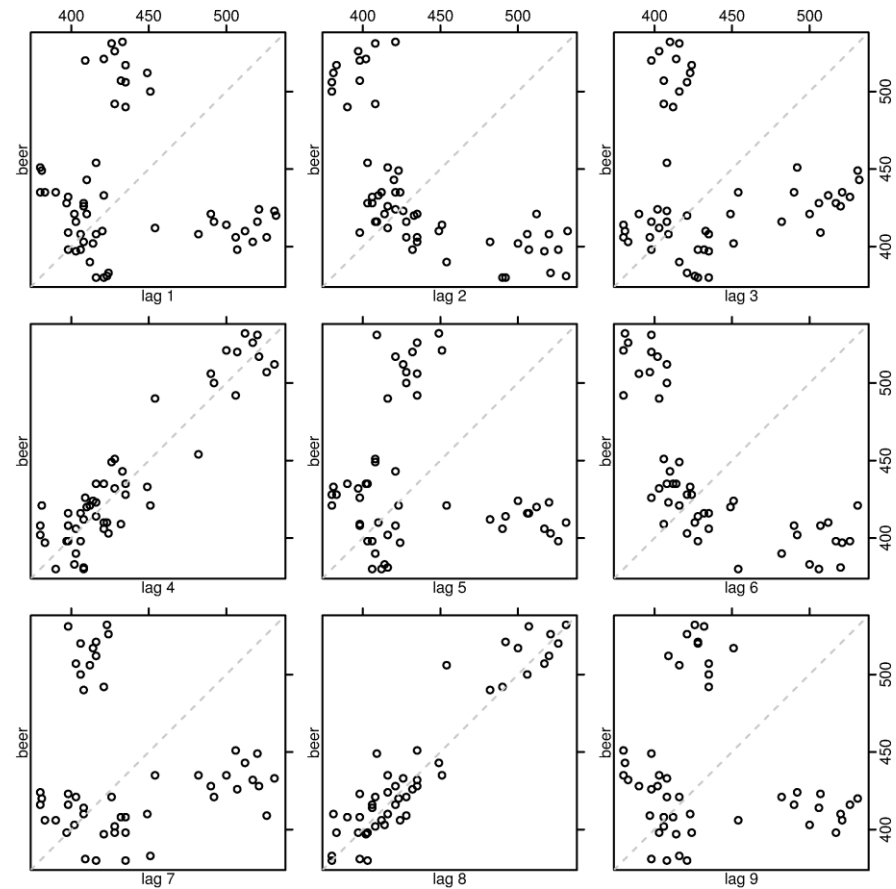
$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}},$$



Autocorrelación

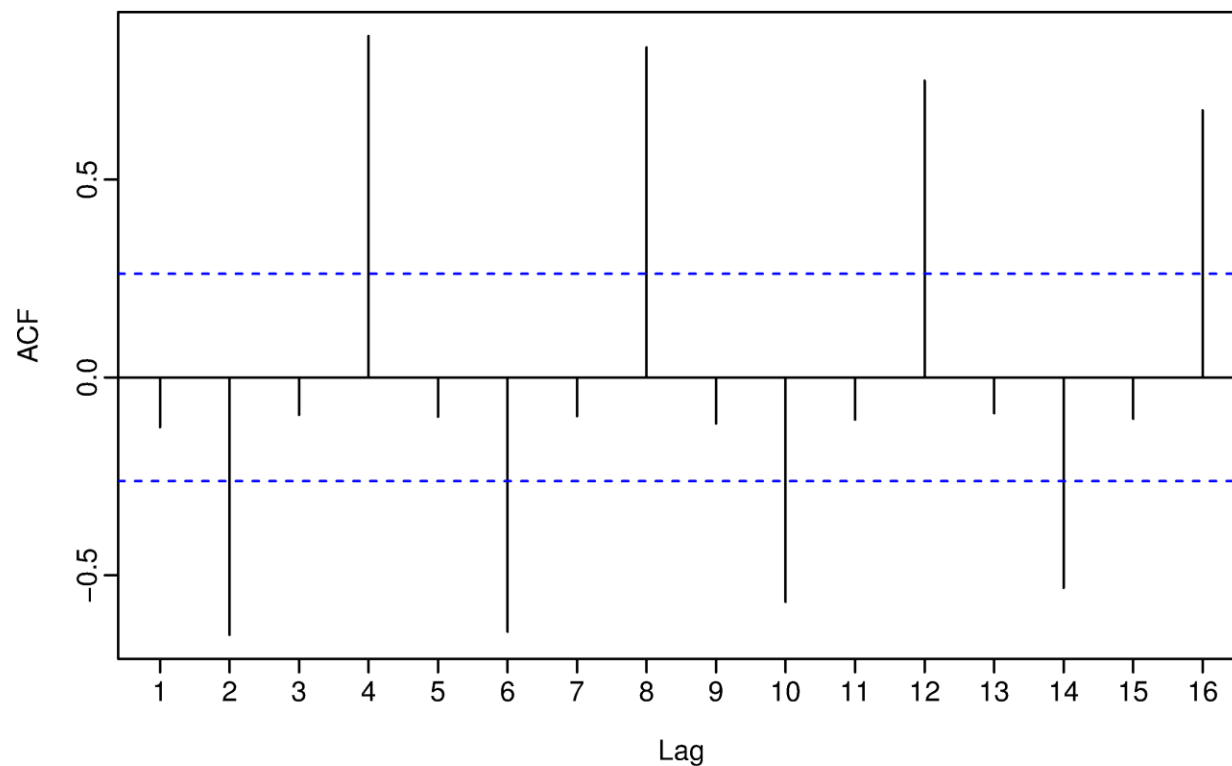
- Relación entre los valores retardados de una serie temporal

$$r_k = \frac{\sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2},$$



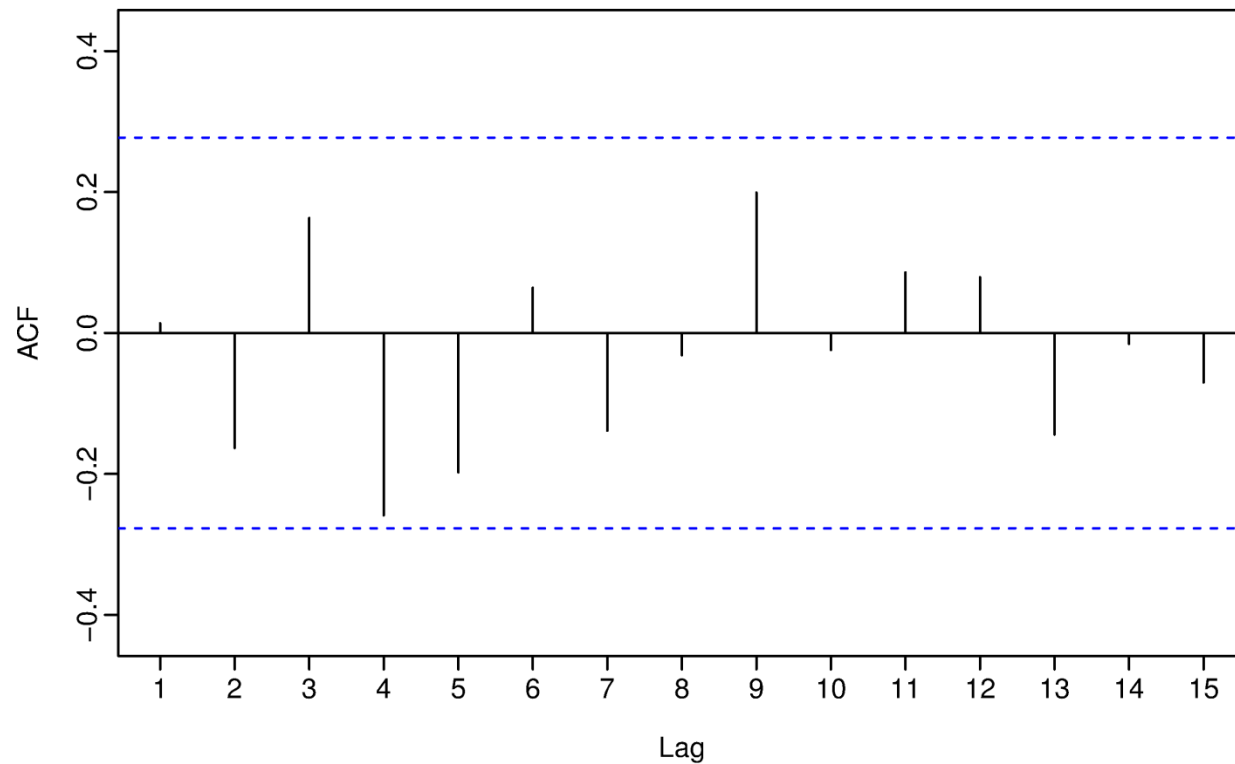
Función de autocorrelación (ACF)

r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9
-0.126	-0.650	-0.094	0.863	-0.099	-0.642	-0.098	0.834	-0.116



Ruido blanco

- Serie que no muestra autocorrelación



Métodos sencillos de forecasting

- Método del promedio

$$\hat{y}_{T+h|T} = \bar{y} = (y_1 + \cdots + y_T)/T.$$

- Método simplista (Naïve)

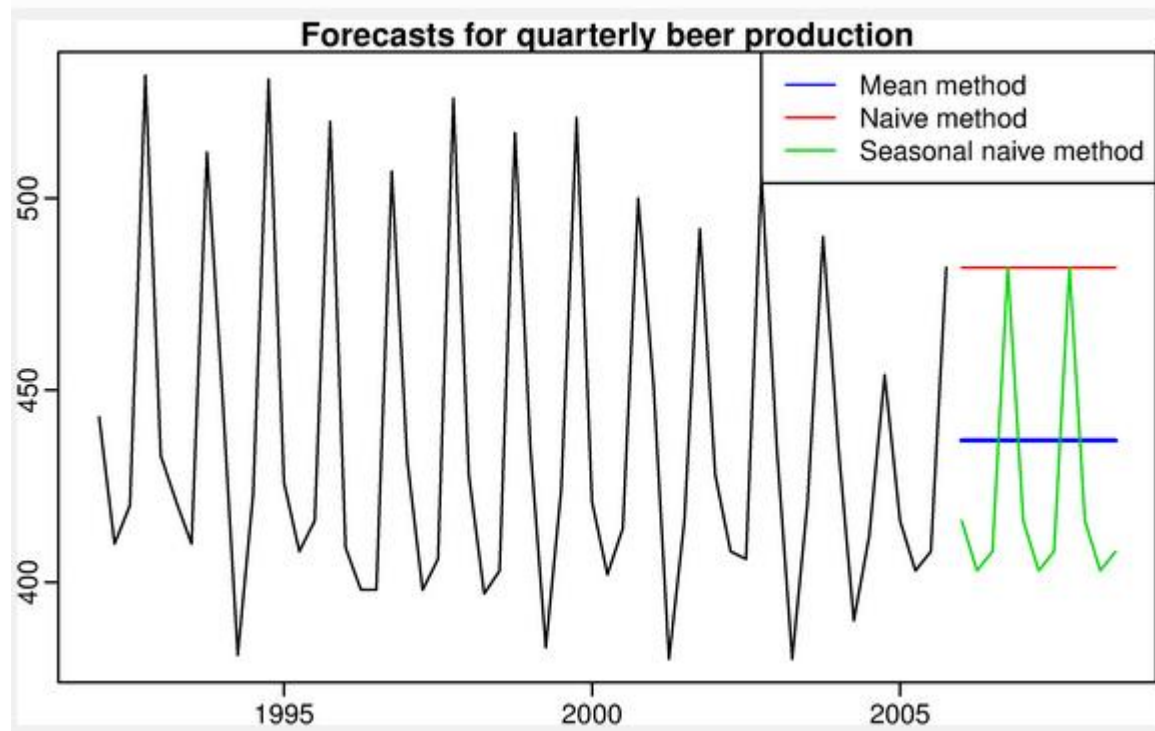
- Predicción: último valor observado

- Método simplista por estaciones

- Método de deriva (drift)

- El pronóstico aumenta o disminuye con el tiempo, por el promedio de los datos históricos

Métodos sencillos de forecasting (2)



Transformaciones

- El ajuste de los datos históricos puede conducir a un modelo de previsión más simple
 - Transformación matemática
 - Ajustes del calendario
 - Ajustes de la población
 - Ajustes de la inflación
- **Objetivo:** conseguir una serie temporal estacionaria

Evaluación de la precisión del forecasting

- Forecast error: $e_i = y_i - \hat{y}_i$
- Errores dependientes de la escala errors

$$\text{MAE} = \text{mean}(|e_i|), \quad \text{RMSE} = \sqrt{\text{mean}(e_i^2)}.$$

- Porcentaje de error:

$$p_i = 100e_i/y_i$$

- Errores escalados

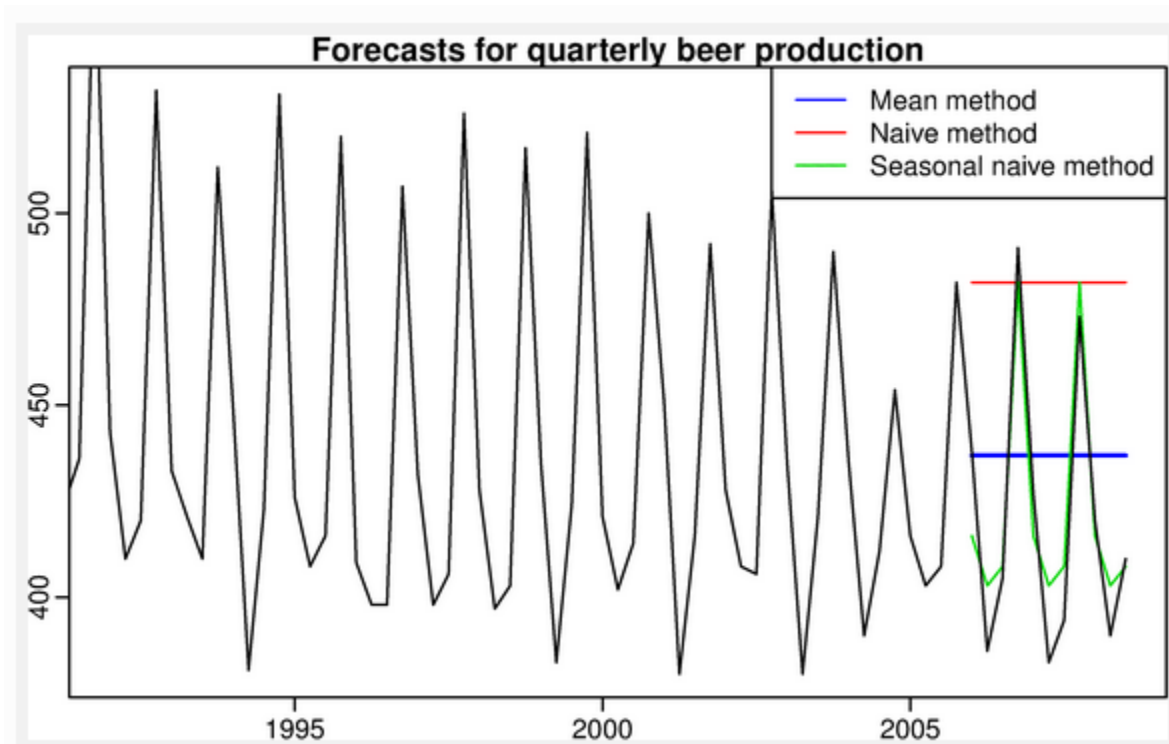
$$\text{MAPE} = \text{mean}(|p_i|).$$

$$\text{sMAPE} = \text{mean}(200|y_i - \hat{y}_i|/(y_i + \hat{y}_i))$$

$$q_j = \frac{e_j}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|}.$$

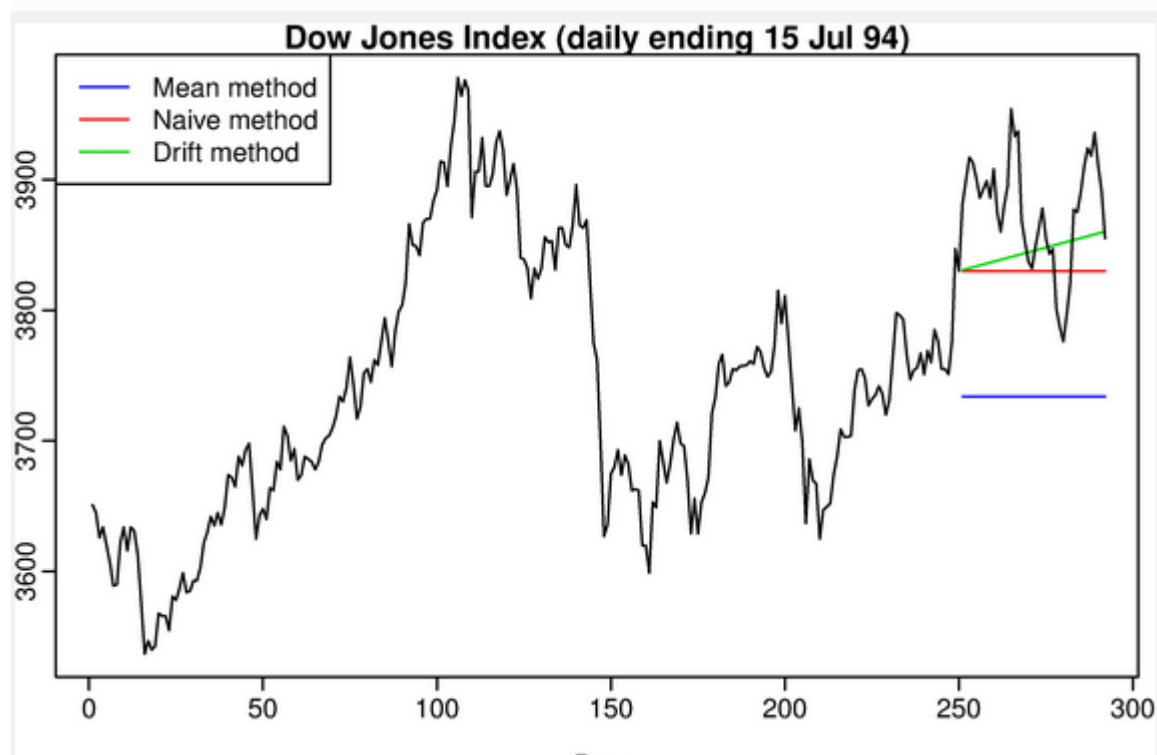
$$\text{MASE} = \text{mean}(|q_j|).$$

Series temporales estacionarias



Method	RMSE	MAE	MAPE	MASE
Mean method	38.01	33.78	8.17	0.61
Naïve method	70.91	63.91	15.88	1.15
Seasonal naïve method	12.97	11.27	2.73	0.20

Series temporales no estacionarias

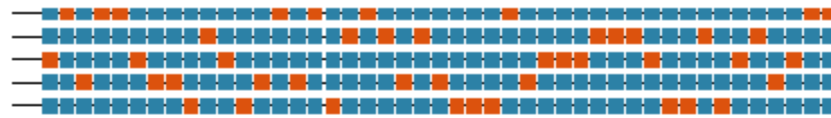


Method	RMSE	MAE	MAPE	MASE
Mean method	148.24	142.42	3.66	8.70
Naïve method	62.03	54.44	1.40	3.32
Drift method	53.70	45.73	1.18	2.79

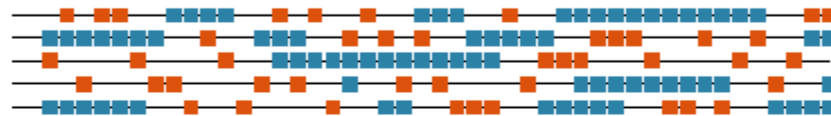
Metodología

- Como en cualquier otra tarea de modelación es esencial realizar una evaluación correcta
- Los datos deben dividirse en partes de entrenamiento y de prueba
- Mejorado a través de la validación cruzada
- Mejorado aún más a través de la validación cruzada bloqueada

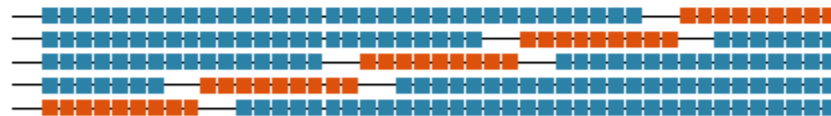
Procedimientos de selección de modelos



cross-validation



non-dep. cross-validation



blocked cross-validation

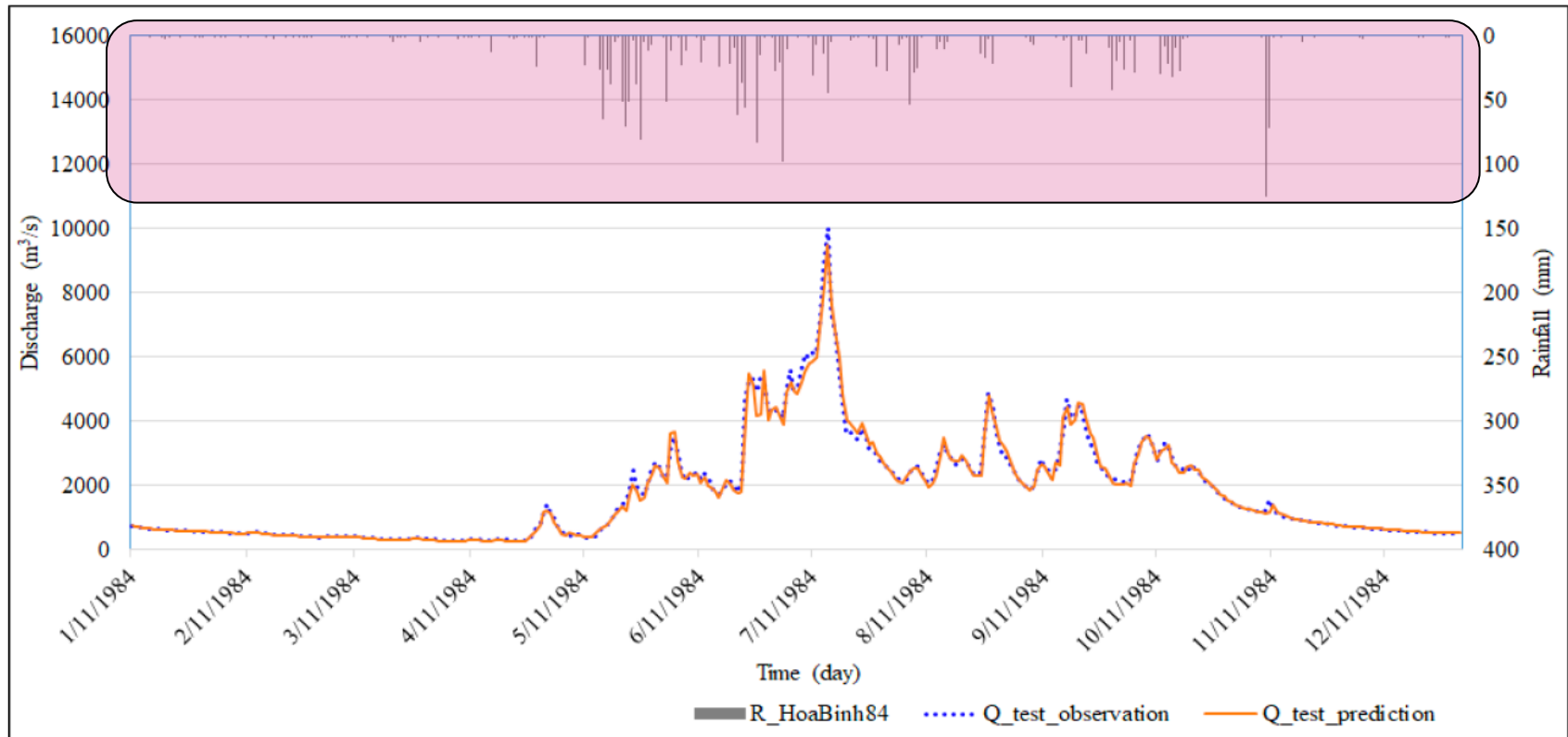


last block

Diagnóstico en base al residuo (error)

- Residuo (error): $e_i = y_i - \hat{y}_i$
- *Buen* método de previsión:
 - Residuos no correlacionados
 - Los residuos tienen una media de cero
- Si el método no los cumple, puede ser mejorado
- Propiedades adicionales:
 - Los residuos tienen una varianza constante
 - Los residuos se distribuyen normalmente

Diagnóstico en base al residuo (error)



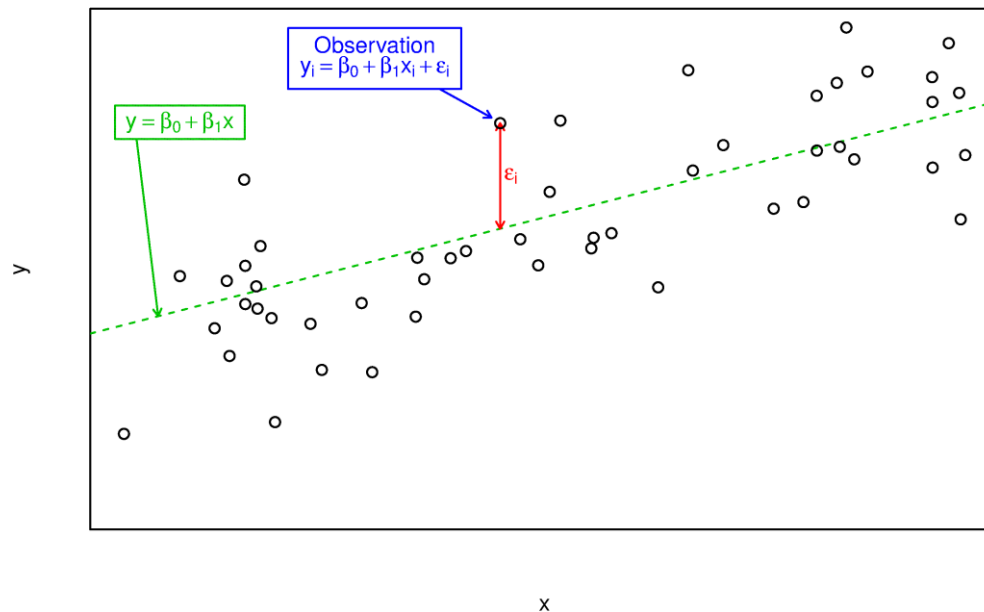
Series temporales

- Predicción (forecasting)
- Herramientas básicas para predicción
- Regresión simple
- Regresión multivariada
- Descomposición de series temporales
- Modelos avanzados

Modelo lineal simple

- Let's assume that the forecast and predictor variables are linearly related:

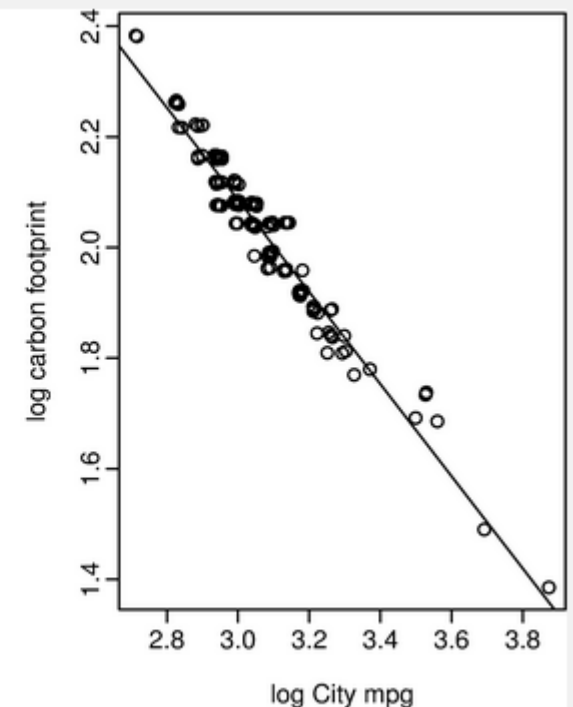
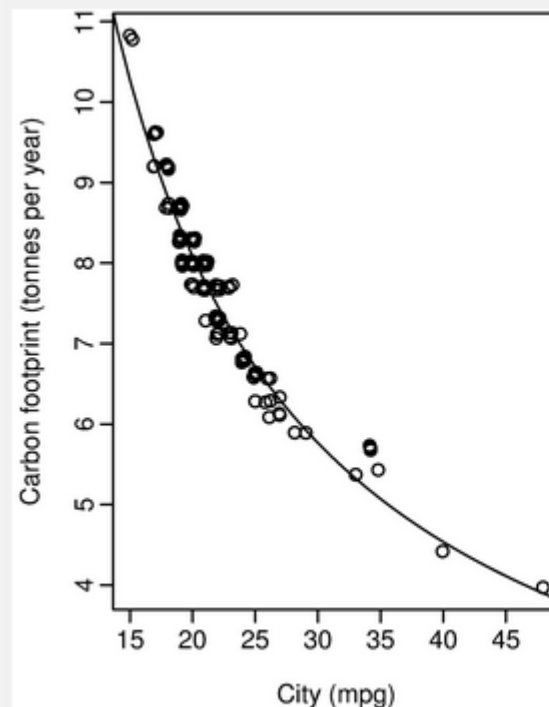
$$y = \beta_0 + \beta_1 x + \varepsilon$$



Regresión no lineal

- Una forma funcional no lineal puede ser más adecuada para un problema que una lineal
- Esto puede obtenerse mediante la transformación de y o x

$$\log y_i = \beta_0 + \beta_1 \log x_i + \varepsilon_i.$$



Series temporales

- Predicción (forecasting)
- Herramientas básicas para predicción
- Regresión simple
- Regresión multivariada
- Descomposición de series temporales
- Modelos avanzados

Regresión múltiple

Una variable a ser pronosticada y **varias variables predictoras**

$$y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \cdots + \beta_k x_{k,i} + e_i.$$

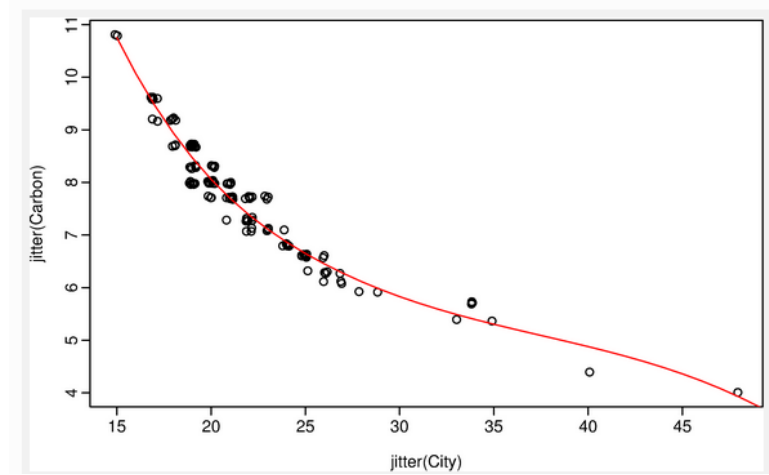
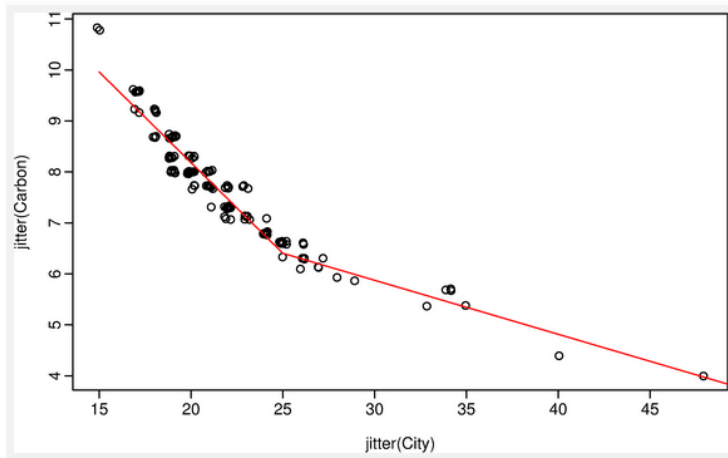
$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{1,i} - \cdots - \beta_k x_{k,i})^2.$$

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_k x_k$$

Predicción de la demanda de electricidad (ED)
ED = f(temperatura actual, fuerza de la economía, población, hora del día, día de la semana, error)

Regresión no lineal

$$y = f(x) + e_i$$



La correlación no es causalidad

- Una variable x puede ser útil para predecir una variable y , pero eso no significa que x sea la causa de y .
- Las correlaciones son útiles para la predicción, incluso cuando no hay una relación causal entre las dos variables

Series temporales

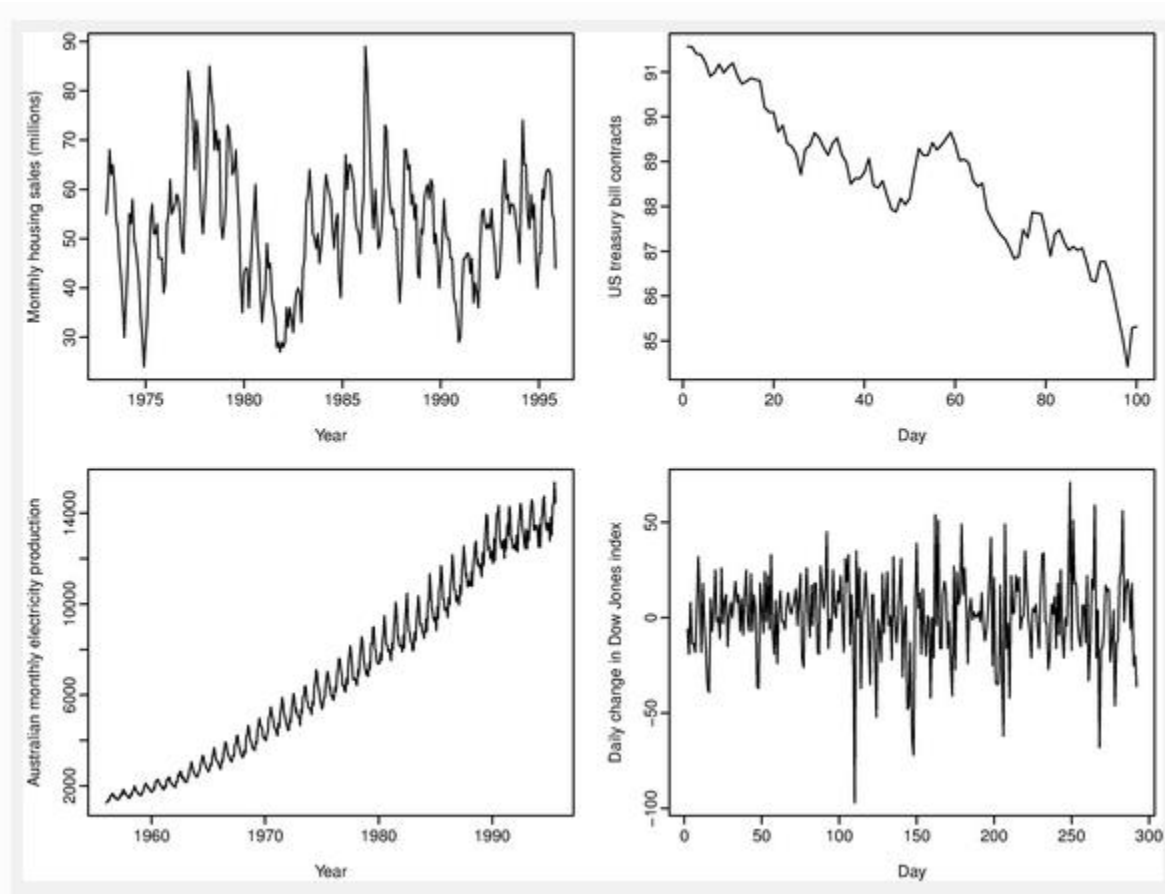
- Predicción (forecasting)
- Herramientas básicas para predicción
- Regresión simple
- Regresión multivariada
- Descomposición de series temporales
- Modelos avanzados

Descomposición de las series temporales

- Las series temporales pueden exhibir una gran variedad de patrones y es útil para categorizar algunos de los patrones y comportamientos que se pueden ver
- A veces también es útil tratar de dividir una serie temporal en varios componentes, cada uno de los cuales representa uno de los componentes subyacentes

Componentes de la serie de tiempo

Tendencia, Estacional, Cíclica



Descomposición de las series temporales

Descomposición aditiva

$$y_t = S_t + T_t + E_t$$

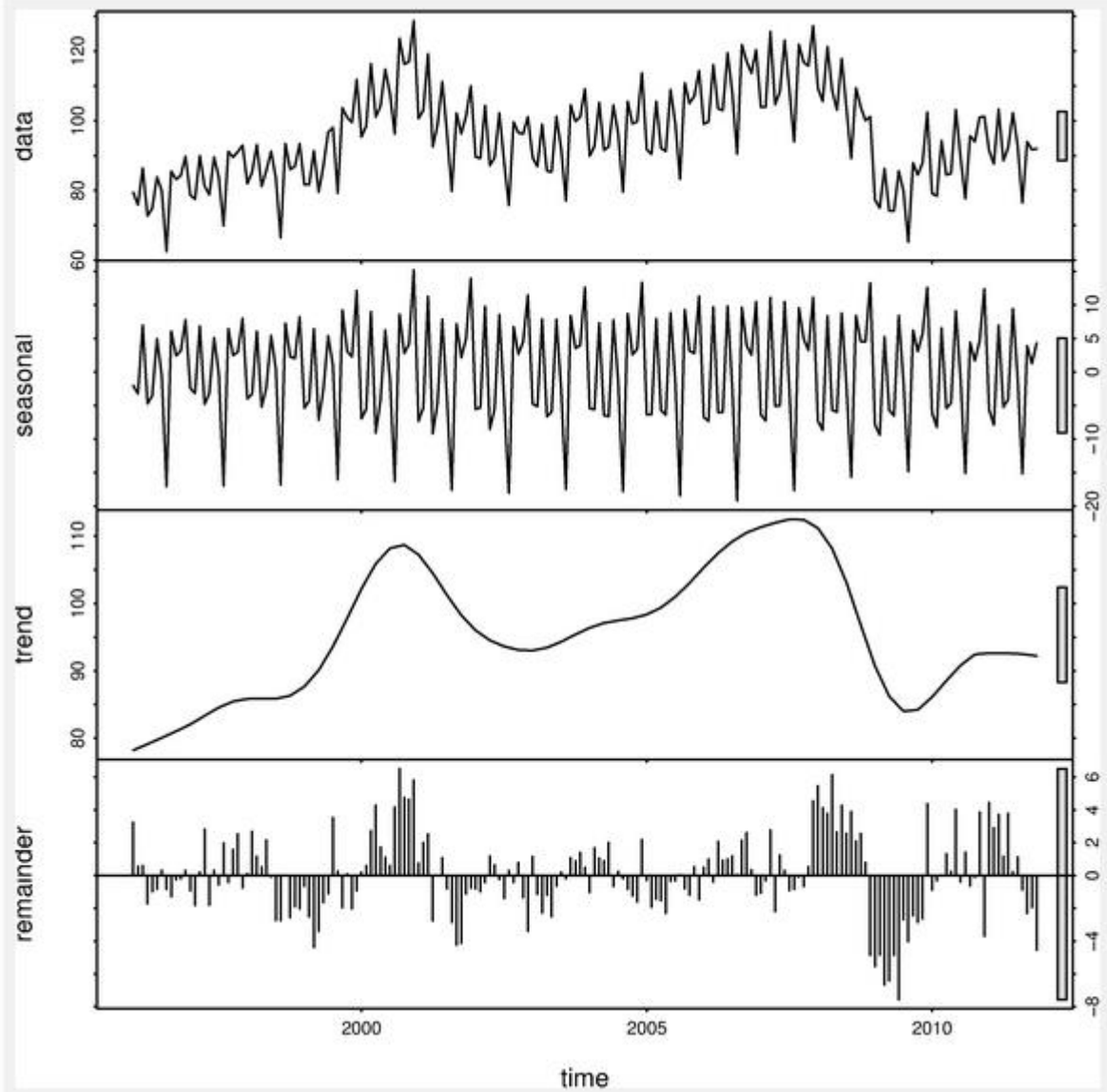
Adecuada cuando la magnitud de las fluctuaciones estacionales o la variación en torno al ciclo de tendencia no varía con el nivel de la serie temporal

Descomposición multiplicativa

$$y_t = S_t \times T_t \times E_t$$

Descomposición STL

- STL es un método de descomposición robusto y versátil: Descomposición estacional y de tendencia usando Loess.
- Puede manejar cualquier tipo de estacionalidad
 - Se permite que el componente estacional cambie con el tiempo, dentro de un rango controlable por el usuario
 - La suavidad del ciclo de tendencia también puede ser controlada por el usuario
 - Es robusto a los valores atípicos (outliers)



Forecasting con descomposición

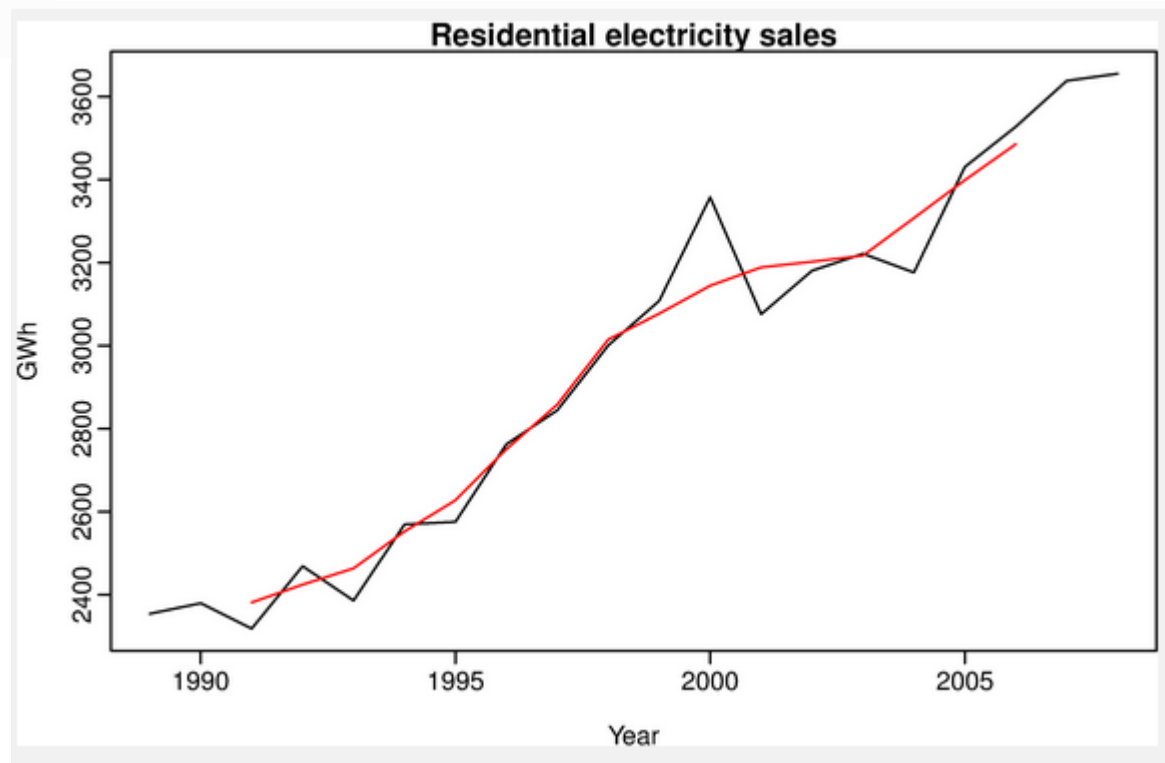
- Para pronosticar una serie temporal descompuesta, pronosticamos los componentes individuales, y luego calculamos el valor predicho

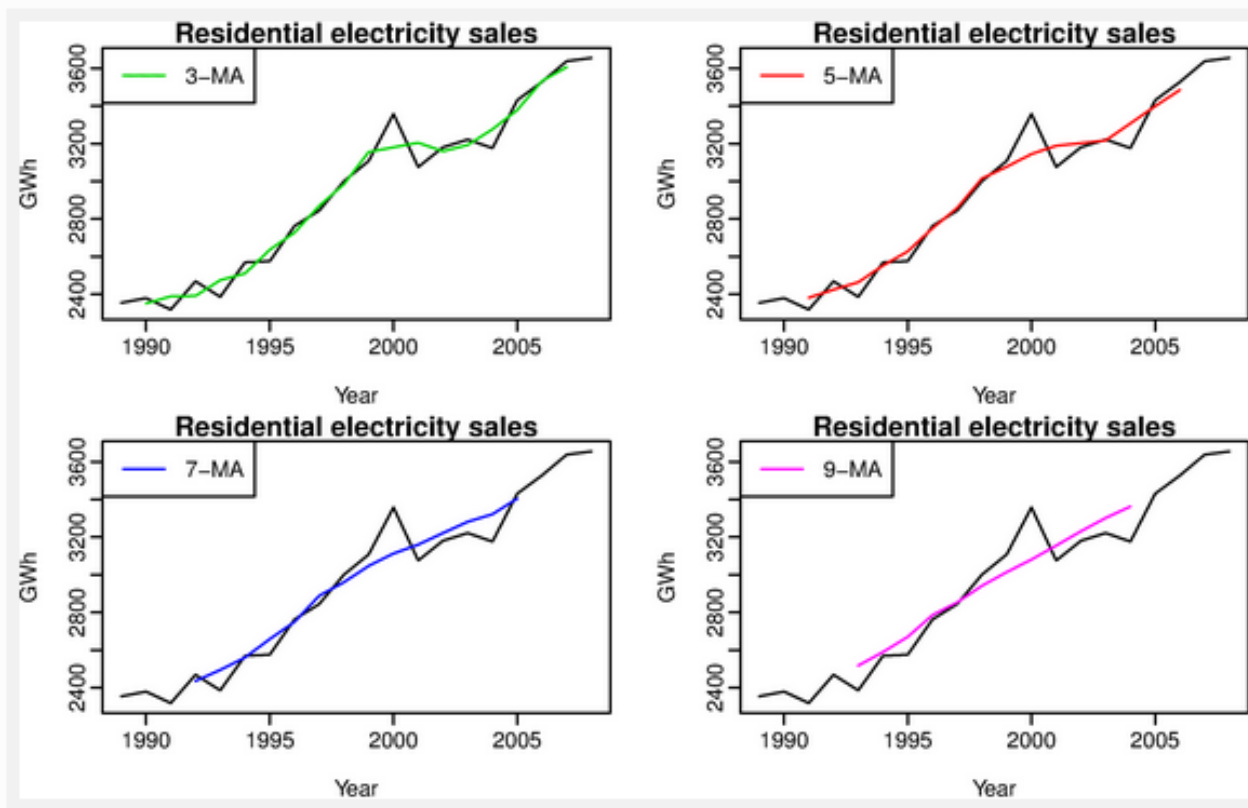
$$y_t = \hat{S}_t + \hat{A}_t$$

$$y_t = \hat{S}_t \hat{A}_t,$$

Medias móviles

$$\hat{T}_t = \frac{1}{m} \sum_{j=-k}^k y_{t+j}$$





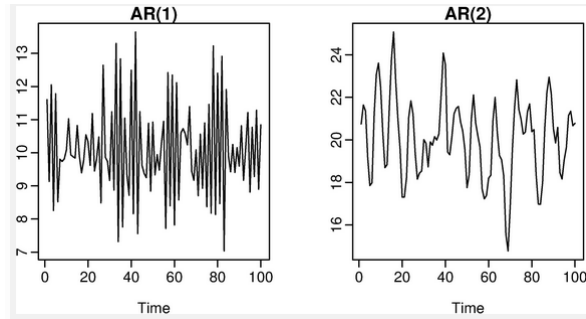
Se utiliza con frecuencia para estimar el ciclo de tendencia a partir de datos estacionales

Series temporales

- Predicción (forecasting)
- Herramientas básicas para predicción
- Regresión simple
- Regresión multivariada
- Descomposición de series temporales
- Modelos avanzados

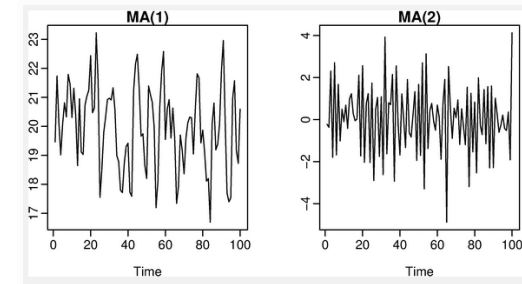
ARIMA

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t$$

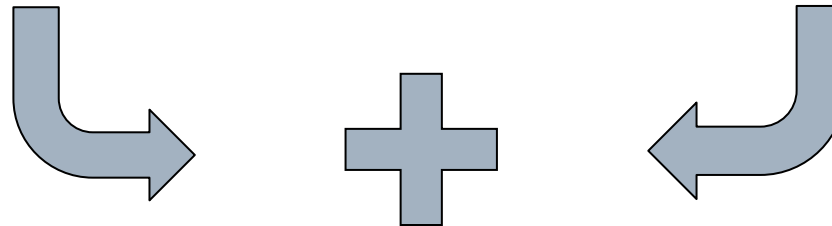


Parte autoregresiva

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q}$$



Parte media móvil



ARIMA(p,d,q) necesita serie estacionaria

p: orden de la parte autorregresiva

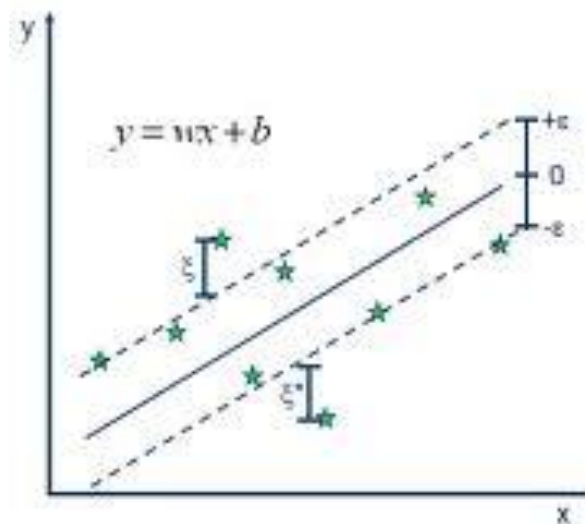
d: grado de la primera parte de la diferencia ← conseguir estacionariedad

q: orden de la parte media móvil

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t$$

Support Vector Regression

- Esta es la versión de las máquinas de kernel para tareas de regresión



• Minimize:

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*)$$

• Constraints:

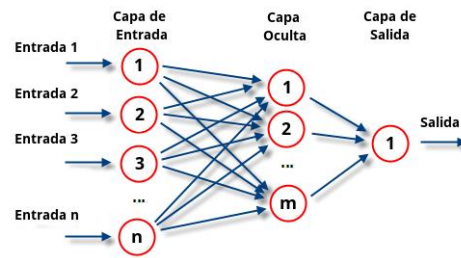
$$y_i - wx_i - b \leq \epsilon + \xi_i$$

$$wx_i + b - y_i \leq \epsilon + \xi_i^*$$

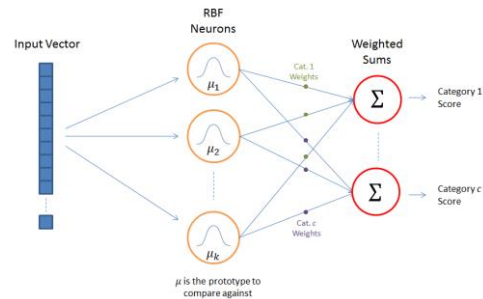
$$\xi_i, \xi_i^* \geq 0$$

Redes neuronales

- Perceptrón multicapa

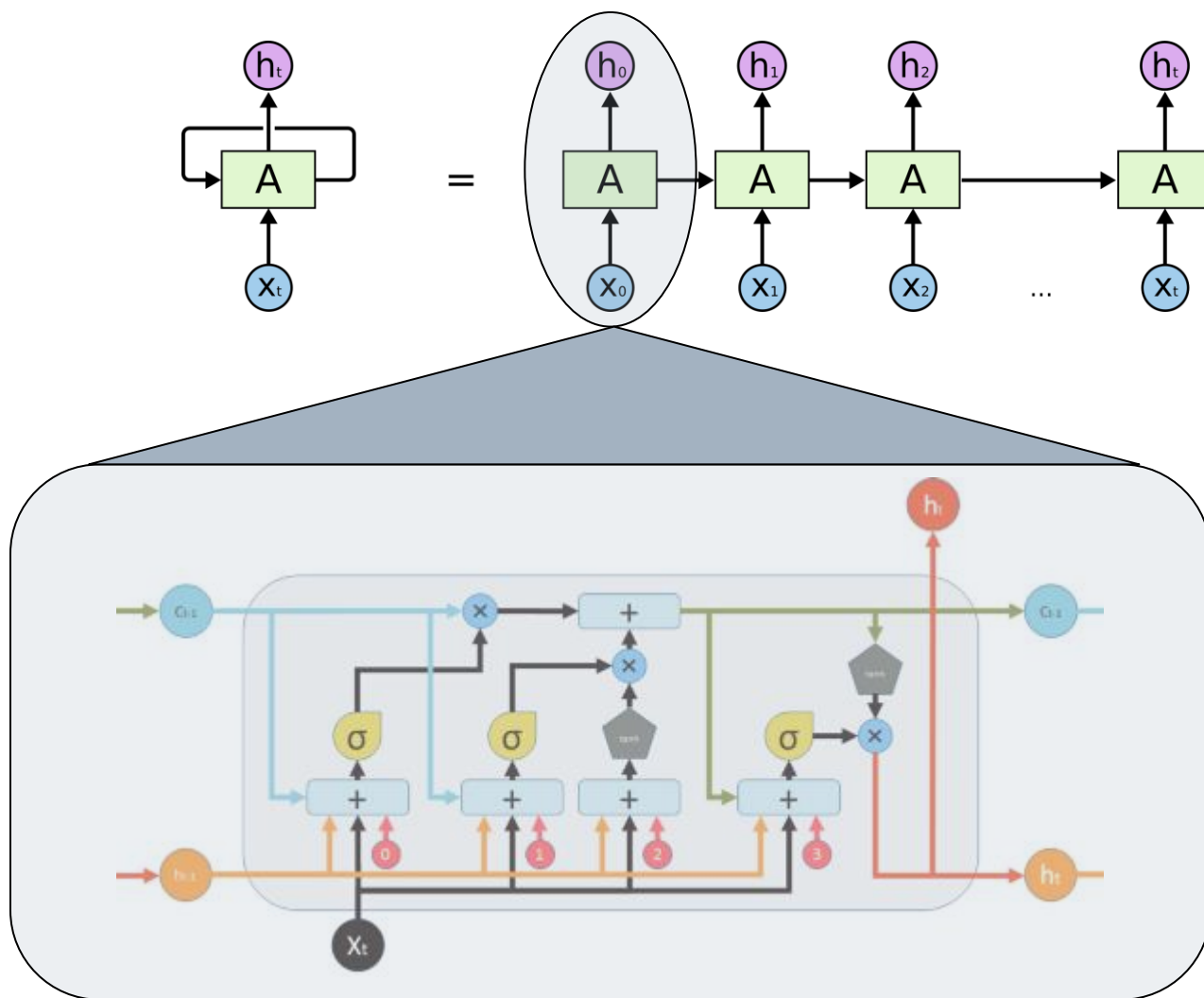


- RBF



- Redes neuronales recurrentes (LSTMs)

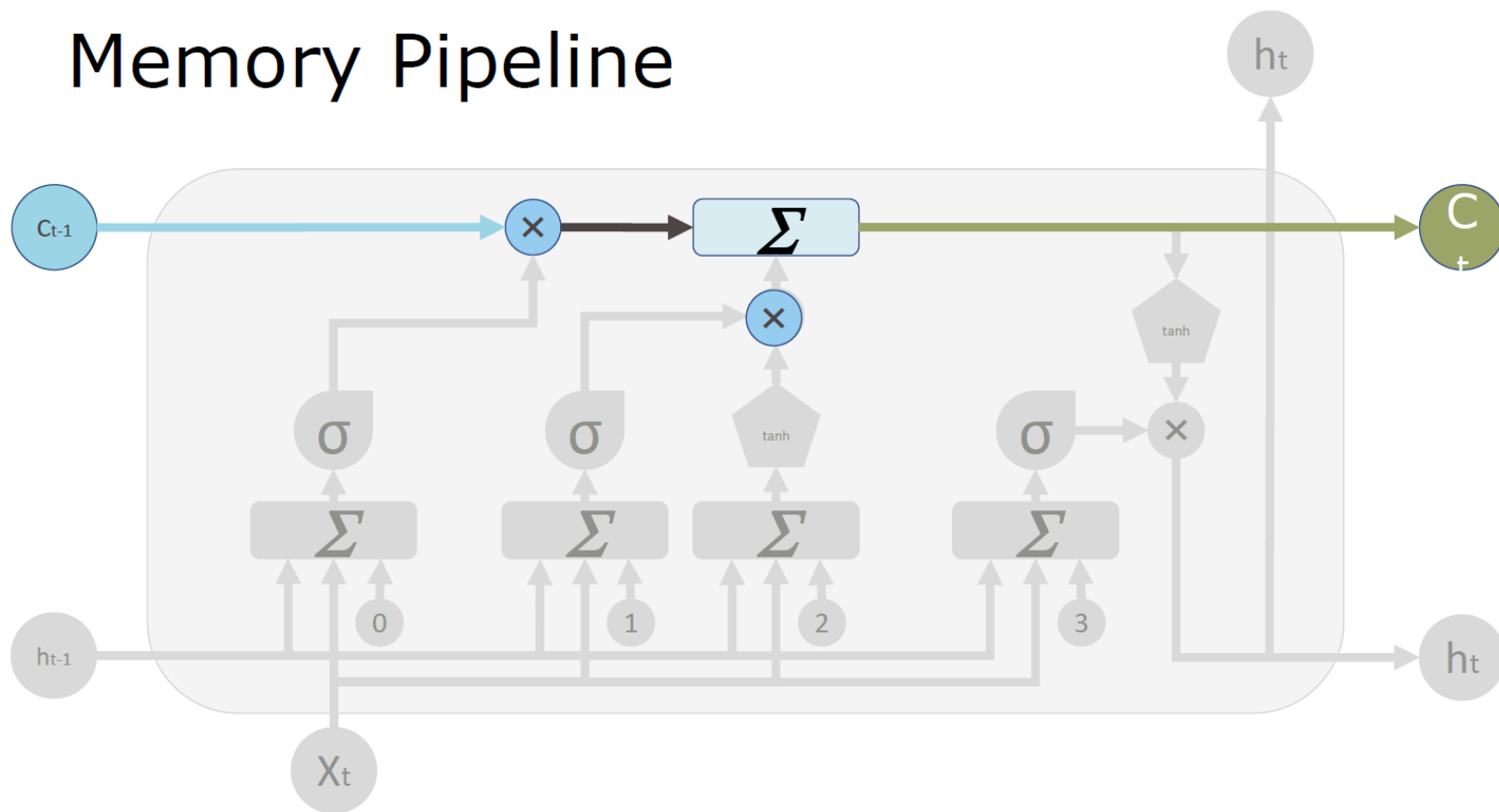
LSTMs



<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

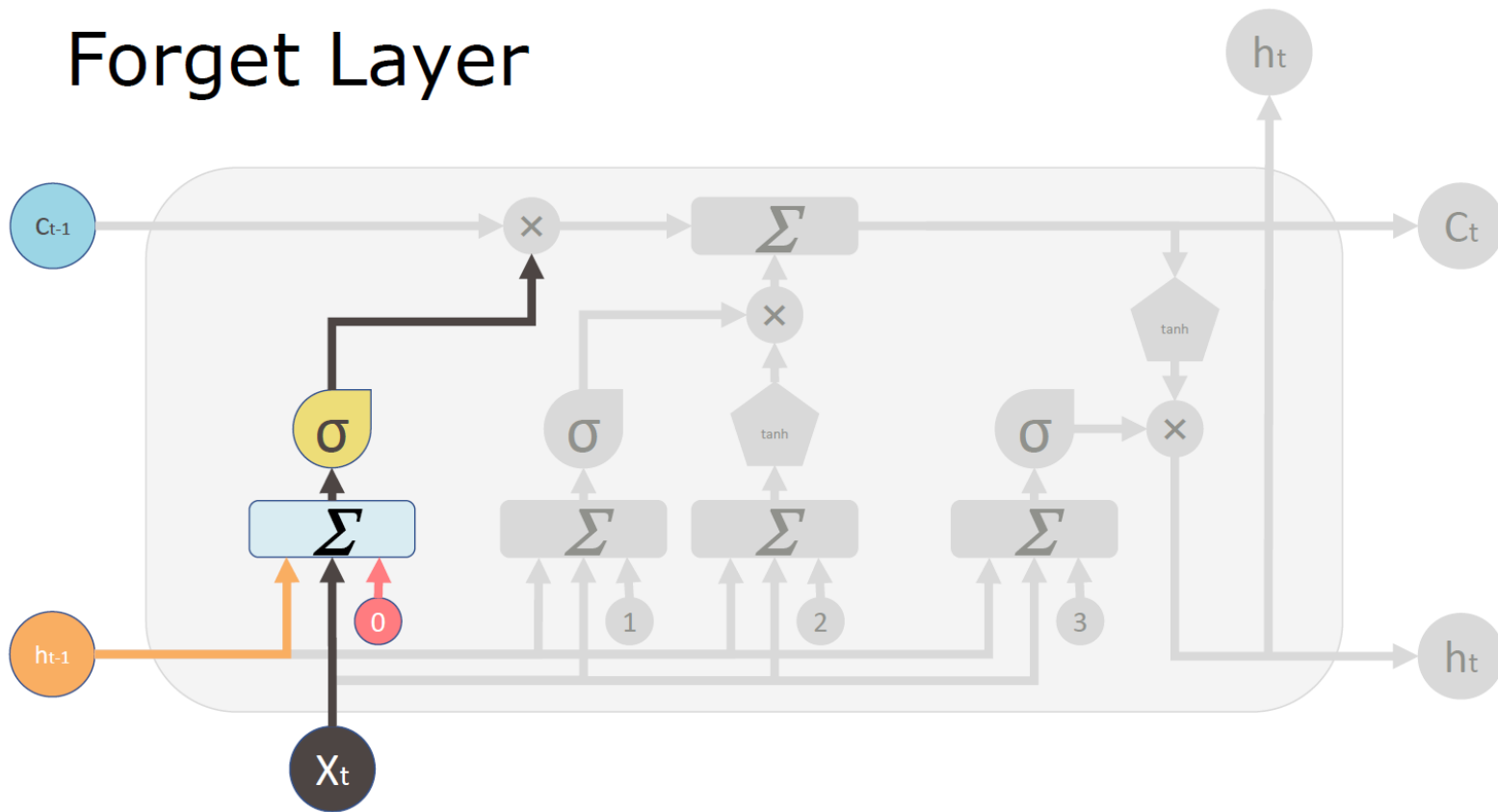
LSTMs

Memory Pipeline



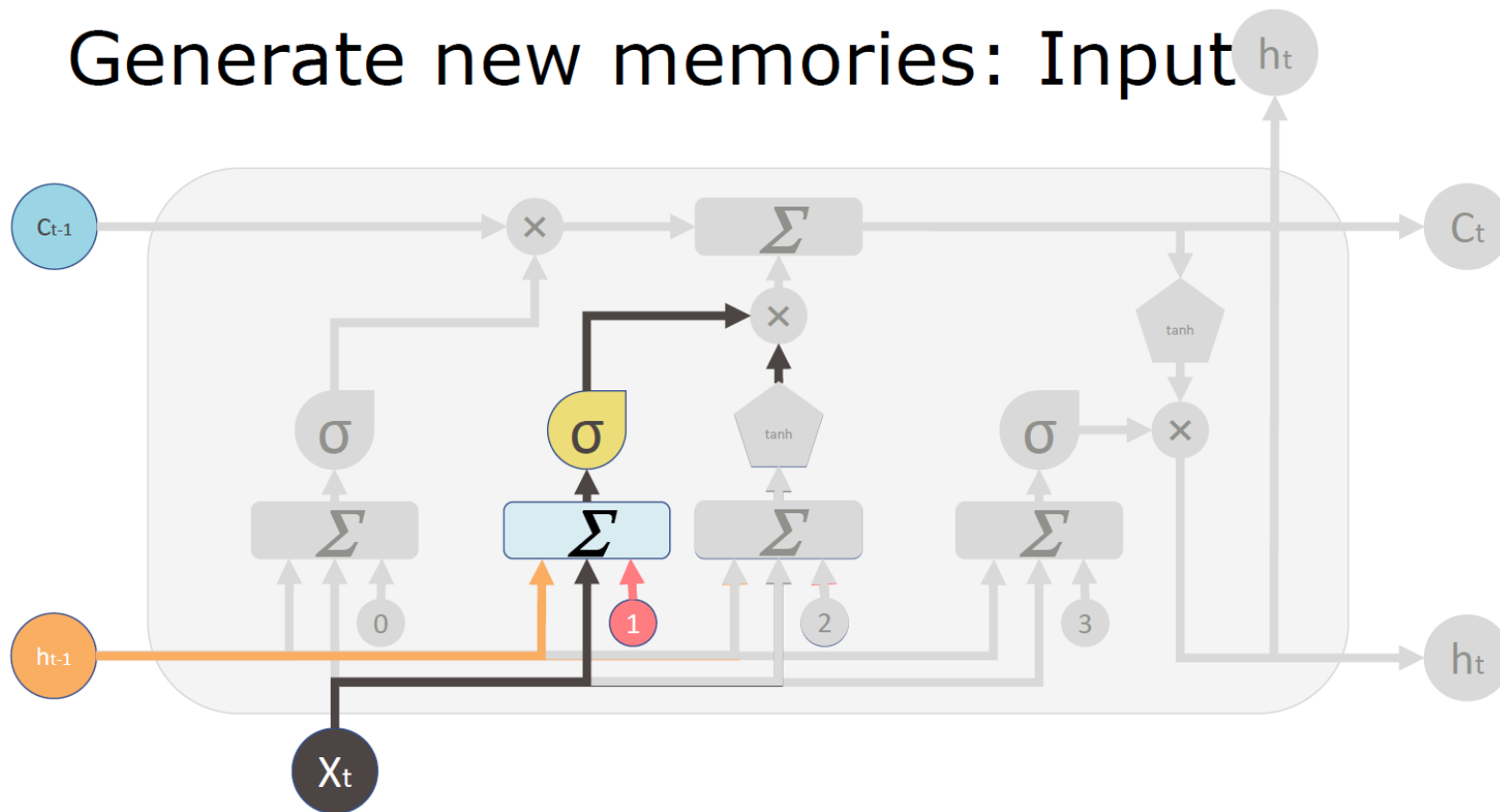
LSTMs

Forget Layer



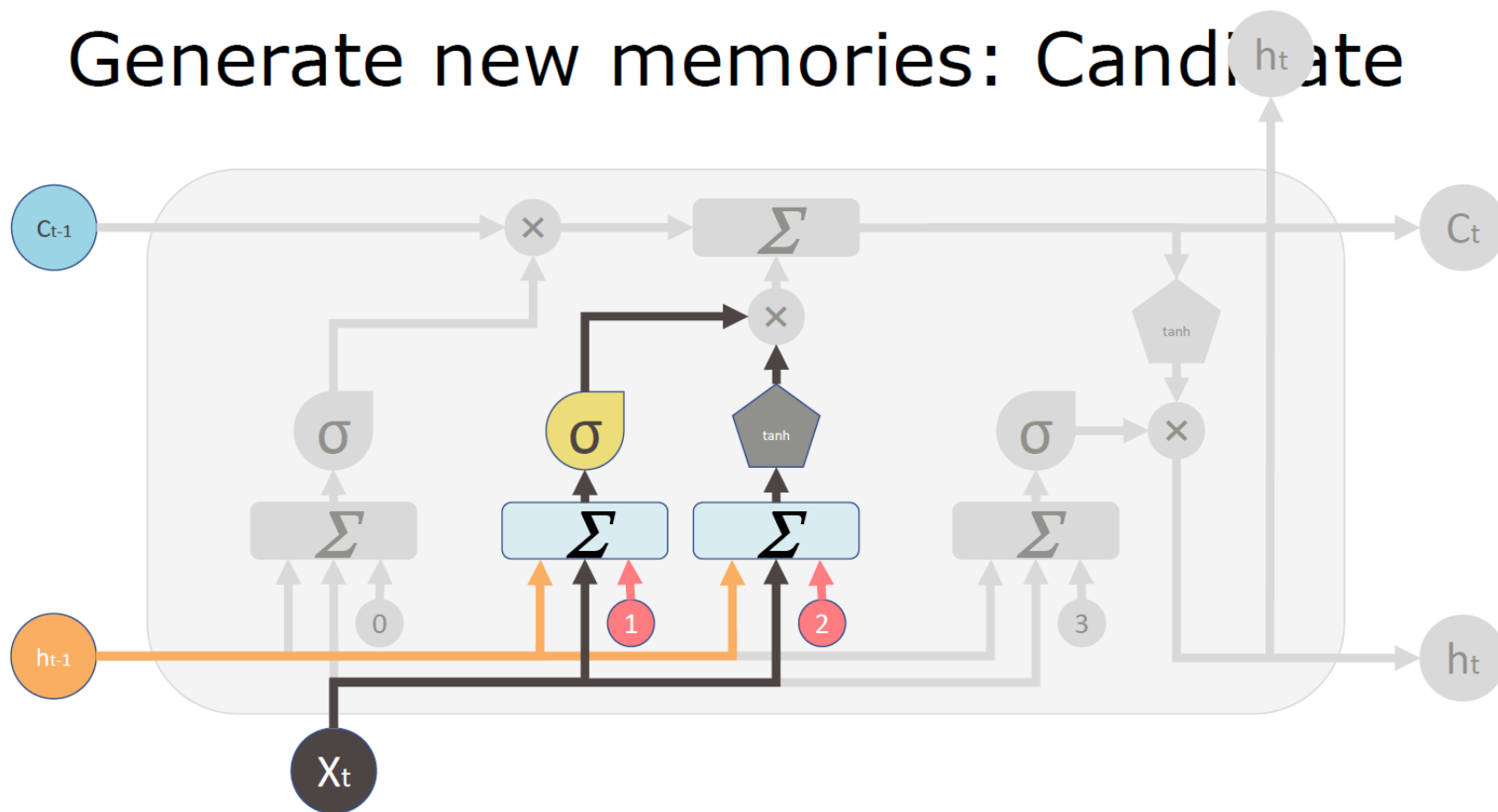
LSTMs

Generate new memories: Input



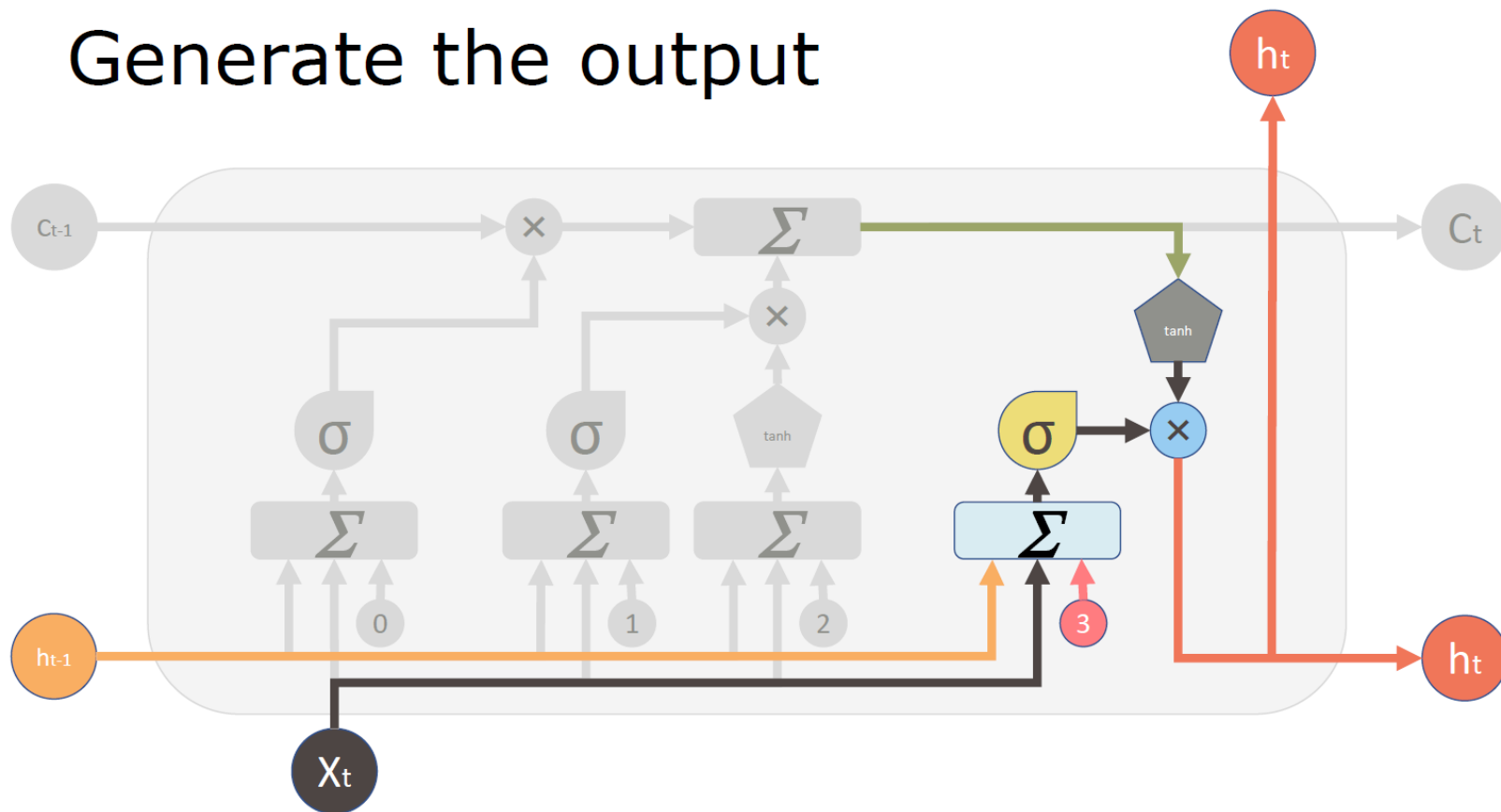
LSTMs

Generate new memories: Candidate



LSTMs

Generate the output

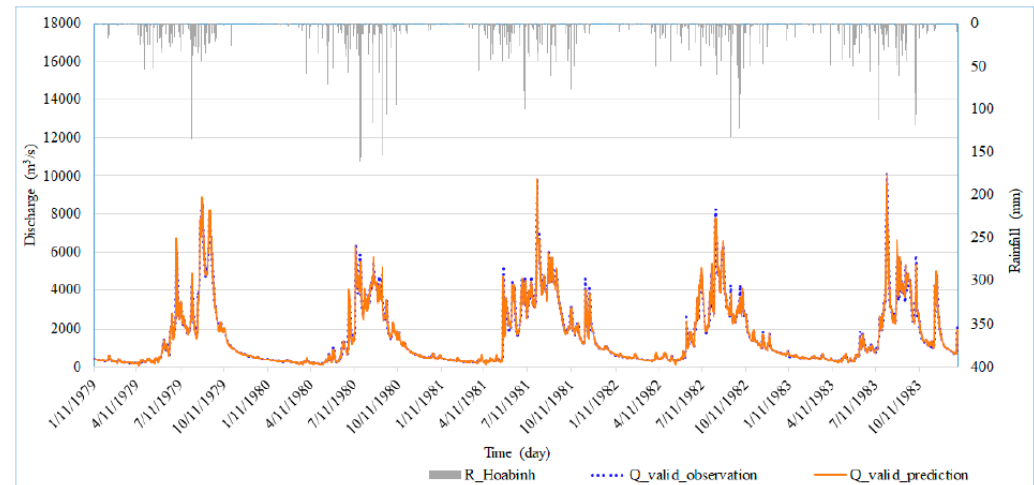
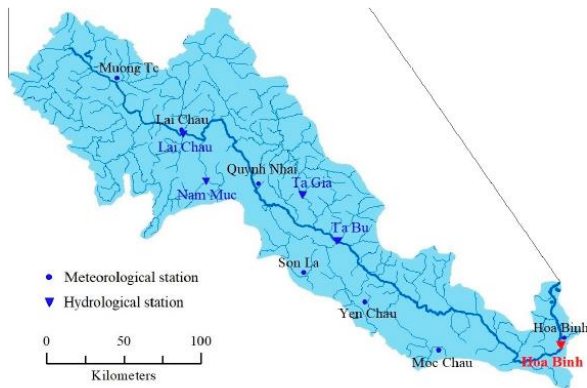
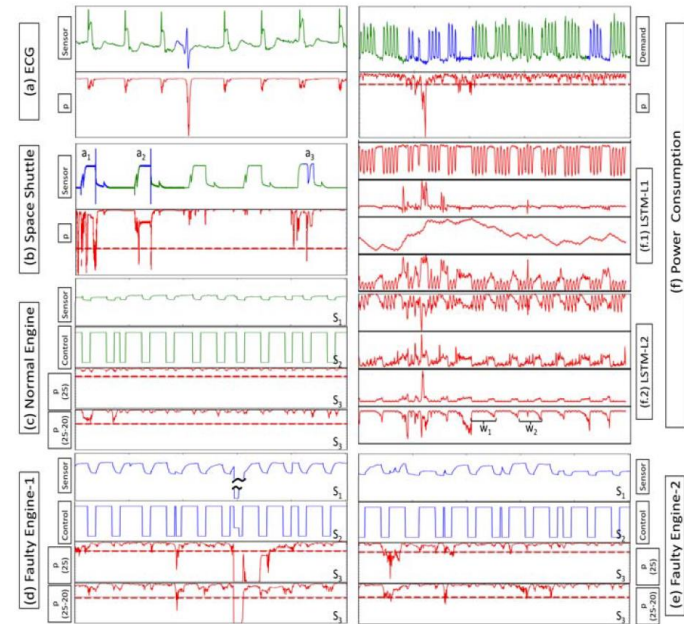


LSTMs - ejemplos

The FBI is chasing a criminal on the run .
 The FBI is chasing a criminal on the run .
 The FBI is chasing a criminal on the run .
 The FBI is chasing a criminal on the run .
 The FBI is chasing a criminal on the run .
 The FBI is chasing a criminal on the run .
 The FBI is chasing a criminal on the run .
 The FBI is chasing a criminal on the run .

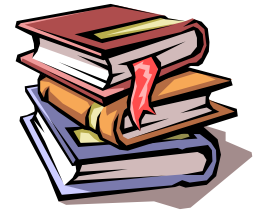
Cheng, Jianpeng, Li Dong, and Mirella Lapata. "Long short-term memory-networks for machine reading." arXiv preprint arXiv:1601.06733 (2016).

Malhotra, Pankaj, et al. "Long short term memory networks for anomaly detection in time series." *Proceedings*. Vol. 89. Presses universitaires de Louvain, 2015.

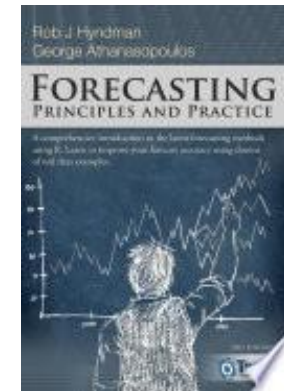


Le, Xuan-Hien, et al. "Application of long short-term memory (LSTM) neural network for flood forecasting." *Water* 11.7 (2019): 1387.

Bibliografía



Hyndman, Rob J., and George Athanasopoulos. Forecasting: principles and practice. OTexts, 2018.



Shumway, Robert H., and David S. Stoffer. Time series analysis and its applications: with R examples. Springer, 2017.



Resumen

1. Regresión

2. Series Temporales

3. Comentarios Finales

Comentarios Finales

Casi todos los problemas de predicción de variables numéricas tienen una componente temporal en los problemas reales

Nuevas tendencias:

- LSTMs

- Ensamblaje de Modelos

Problemas típicos

- Detección de anomalías

- Predicciones encadenadas