PRÁCTICA 2- Series Temporales (LITUANIA)

María Pallares Diez

2025-02-27

Introducción

En este análisis, se estudia la evolución del **consumo eléctrico en Lituania** desde **enero de 2012 hasta diciembre de 2023**, utilizando datos de **Eurostat**. La serie temporal refleja la demanda energética del país, influenciada por **factores climáticos, económicos y geopolíticos**, y presenta **patrones estacionales** marcados por la variabilidad climática a lo largo del año.

Para modelar esta serie, se ha explorado el uso de **métodos de alisado exponencial (ETS)**, probando diferentes configuraciones y transformaciones (raíz cuadrada, inversa y ajuste por días laborables) con el objetivo de minimizar el error de predicción.

El estudio busca **seleccionar el modelo más adecuado**, evaluar su desempeño a través de métricas de error y validar su capacidad predictiva mediante una **comparación con un método ingenuo con estacionalidad**. Finalmente, se ha realizado una **predicción a tres años vista**, proporcionando una perspectiva sobre la evolución futura del consumo eléctrico en Lituania y la fiabilidad de las previsiones obtenidas.

```r
library(forecast)
library(ggplot2)
library(tseries)
library(timeDate)
datos <- read.csv("ELE_Lituania.csv", header = TRUE)
electricidad <- ts(datos, start = c(2012, 1), frequency = 12)
ndiffs(electricidad, alpha = 0.05, test = "kpss")
```

```
## [1] 0
```

La prueba **KPSS** se ha aplicado a la serie de electricidad para determinar si es necesario diferenciarla para hacerla estacionaria. El resultado obtenido es **0**, lo que indica que la serie **ya es estacionaria** y **no requiere diferenciación**. Esto sugiere que se puede modelar directamente sin aplicar transformaciones adicionales para eliminar tendencias, lo que comprobaremos a continución en la búsqueda del mejor modelo.

Búsqueda del mejor modelo

Para seleccionar el mejor modelo de alisado exponencial para el **consumo eléctrico en Lituania**, se evaluaron **cuatro enfoques**:

1. **ETS(M,N,A)**: Aplicado directamente sobre la serie original.
2. **ETS(A,N,A) con raíz cuadrada**: Para reducir la varianza en series con heterocedasticidad.
3. **ETS(M,A,M) con transformación inversa**: Útil ante efectos de escala.

4. **ETS(A,N,A) ajustado por días laborables**: Considera la variabilidad mensual en días hábiles.

Previo a la validación cruzada, se seleccionó el **mejor modelo ETS para cada transformación** con la función `ets()`, asegurando el ajuste más adecuado.

```
# Definir los parámetros de validación
k <- 120
h <- 12
TT <- length(electricidad)
s <- TT - k - h
# 1. Modelo normal ETS(M,N,A)
mapeNormal <- matrix(NA, s + 1, h)
for (i in 0:s) {
  train.set <- subset(electricidad, start = i + 1, end = i + k)
  test.set <- subset(electricidad, start = i + k + 1, end = i + k + h)
  fit <- ets(train.set, model = "MNA", damped = FALSE)
  fcast <- forecast(fit, h = h)
  predicted_original <- fcast$mean  # Sin transformación
  mapeNormal[i + 1, ] <- 100 * abs(test.set - predicted_original) / test.set
}
errorNormal <- apply(mapeNormal, MARGIN = 2, FUN = median)
# 2. Modelo raíz cuadrada ETS(A,N,A)
mapeRaiz <- matrix(NA, s + 1, h)
for (i in 0:s) {
  train.set <- subset(sqrt(electricidad), start = i + 1, end = i + k)
  test.set <- subset(electricidad, start = i + k + 1, end = i + k + h)
  fit <- ets(train.set, model = "ANA", damped = FALSE)
  fcast <- forecast(fit, h = h)
  predicted_original <- (fcast$mean)^2  # Deshacer la transformación de raíz cuadrada
  mapeRaiz[i + 1, ] <- 100 * abs(test.set - predicted_original) / test.set
}
errorRaiz <- apply(mapeRaiz, MARGIN = 2, FUN = median)
# 3. Modelo inverso ETS(M,A,M)
mapeInv <- matrix(NA, s + 1, h)
for (i in 0:s) {
  train.set <- subset(1 / (electricidad + 1), start = i + 1, end = i + k)
  test.set <- subset(electricidad, start = i + k + 1, end = i + k + h)
  fit <- ets(train.set, model = "MAM", damped = FALSE)
  fcast <- forecast(fit, h = h)
  predicted_original <- 1 / fcast$mean - 1  # Deshacer la transformación inversa
  mapeInv[i + 1, ] <- 100 * abs(test.set - predicted_original) / test.set
}
errorInv <- apply(mapeInv, MARGIN = 2, FUN = median)
# 4. Modelo ajustado por días laborables ETS(A,N,A)
festivos_lt <- as.Date(c(
  "2012-01-01", "2012-02-16", "2012-03-11", "2012-04-08", "2012-05-01", "2012-06-24", "2012-07-06", "20
  "2013-01-01", "2013-02-16", "2013-03-11", "2013-04-08", "2013-05-01", "2013-06-24", "2013-07-06", "20
  "2014-01-01", "2014-02-16", "2014-03-11", "2014-04-08", "2014-05-01", "2014-06-24", "2014-07-06", "20
  "2015-01-01", "2015-02-16", "2015-03-11", "2015-04-08", "2015-05-01", "2015-06-24", "2015-07-06", "20
  "2016-01-01", "2016-02-16", "2016-03-11", "2016-04-08", "2016-05-01", "2016-06-24", "2016-07-06", "20
  "2017-01-01", "2017-02-16", "2017-03-11", "2017-04-08", "2017-05-01", "2017-06-24", "2017-07-06", "20
  "2018-01-01", "2018-02-16", "2018-03-11", "2018-04-08", "2018-05-01", "2018-06-24", "2018-07-06", "20
  "2019-01-01", "2019-02-16", "2019-03-11", "2019-04-08", "2019-05-01", "2019-06-24", "2019-07-06", "20
  "2020-01-01", "2020-02-16", "2020-03-11", "2020-04-08", "2020-05-01", "2020-06-24", "2020-07-06", "20
```

```r
))
fechas <- seq(as.Date("2012-01-01"), by = "month", length.out = length(electricidad))
dias_laborables_lt <- sapply(fechas, function(x) sum(!weekdays(seq(x, length = 30, by = "day")) %in% c(
mapeLab <- matrix(NA, s + 1, h)
for (i in 0:s) {
  train.set <- subset(electricidad / dias_laborables_lt, start = i + 1, end = i + k)
  test.set <- subset(electricidad, start = i + k + 1, end = i + k + h)
  fit <- ets(train.set, model = "ANA", damped = FALSE)
  fcast <- forecast(fit, h = h)
  predicted_original <- fcast$mean * dias_laborables_lt[i + k + 1]  # Deshacer la transformación
  mapeLab[i + 1, ] <- 100 * abs(test.set - predicted_original) / test.set
}
errorLab <- apply(mapeLab, MARGIN = 2, FUN = median)
# Comparar errores de cada modelo
errores <- data.frame(
  Horizonte = 1:h,
  Normal = errorNormal,
  Raiz_Cuadrada = errorRaiz,
  Inversa = errorInv,
  Laborables = errorLab
)
print(errores)
```

```
##    Horizonte    Normal Raiz_Cuadrada   Inversa Laborables
## 1          1  2.701136      3.099808  3.124371   3.408927
## 2          2  3.834796      3.863402  4.716512   3.982686
## 3          3  3.904960      4.106227  7.055493   5.692220
## 4          4  4.643505      4.220252  7.903511   6.096004
## 5          5  8.190689      8.675442  8.553695   8.262513
## 6          6  8.842868      8.549035  9.532417   9.418650
## 7          7 10.578176      9.522932  9.540811  10.642168
## 8          8 10.492894      9.923654 11.026323  11.293280
## 9          9 10.773330     11.153195 12.176035  11.393935
## 10        10 11.356017     11.165228 12.707756  11.497979
## 11        11 11.667914     10.895326 13.150837  11.717753
## 12        12 12.121960     11.732180 12.965916  12.211458
```

La tabla presenta los errores **MAPE** para cada transformación aplicada a la serie de consumo eléctrico en Lituania. El modelo **ETS(M,N,A)** muestra los errores más bajos en los primeros meses y estabilidad en el largo plazo. Las transformaciones (Raíz Cuadrada, Inversa, Laborables) no mejoran la precisión y, en algunos casos, incrementan el error, mientras que el modelo inverso **ETS(M,A,M)** presenta los errores más altos, descartándolo.

Dado que la serie es **estacionaria**, según la prueba **KPSS**, las transformaciones no aportan mejoras significativas. Por ello, se selecciona **ETS(M,N,A)** como el mejor modelo sin necesidad de ajustes adicionales.

```r
mejor_modelo<- ets(electricidad)
summary(mejor_modelo)
```

```
## ETS(M,N,A)
##
## Call:
## ets(y = electricidad)
```

```
##
##   Smoothing parameters:
##     alpha = 0.5564
##     gamma = 1e-04
##
##   Initial states:
##     l = 882.9695
##     s = 116.8095 44.3362 21.7039 -51.2182 -40.9389 -65.9988
##            -73.9594 -57.6515 -53.9676 35.8369 19.4219 105.6259
##
##   sigma:  0.0385
##
##      AIC     AICc      BIC
## 1751.192 1754.942 1795.740
##
## Training set error measures:
##                       ME     RMSE      MAE         MPE     MAPE      MASE
## Training set -0.04847888 33.12552 25.17995 -0.09890442 2.806865 0.5760404
##                      ACF1
## Training set -0.008019885
```

El modelo seleccionado es **ETS(M,N,A)**, lo que indica que la serie presenta un **nivel multiplicativo**, **sin tendencia**, y **con estacionalidad aditiva**. Esto sugiere que la serie mantiene una estructura estacional constante, sin crecimiento ni decrecimiento en la tendencia.

Los **parámetros de suavizado** obtenidos son:

- $\alpha = 0.5564$: Se asigna un peso moderado a las observaciones recientes, lo que indica que el modelo responde a los cambios en la serie sin ser demasiado reactivo.
- $\gamma = 1 \times 10^{-4}$: La estacionalidad varía mínimamente en el tiempo, lo que indica estabilidad en los patrones estacionales.

El estado inicial muestra un **nivel base de 882.97**, con variaciones estacionales significativas.

El modelo presenta una **varianza del error (sigma) de 0.0385**, lo que sugiere que la dispersión en los residuos es baja y el modelo es estable.

En términos de criterios de información:

- **AIC** = 1751.19, **AICc** = 1754.94, **BIC** = 1795.74, lo que indica un **buen equilibrio entre ajuste y complejidad del modelo**.

Indicadores de Calidad del Modelo

Para evaluar la precisión del modelo **ETS(M,N,A)** en la serie de **consumo eléctrico en Lituania**, se analizan las siguientes métricas de error:

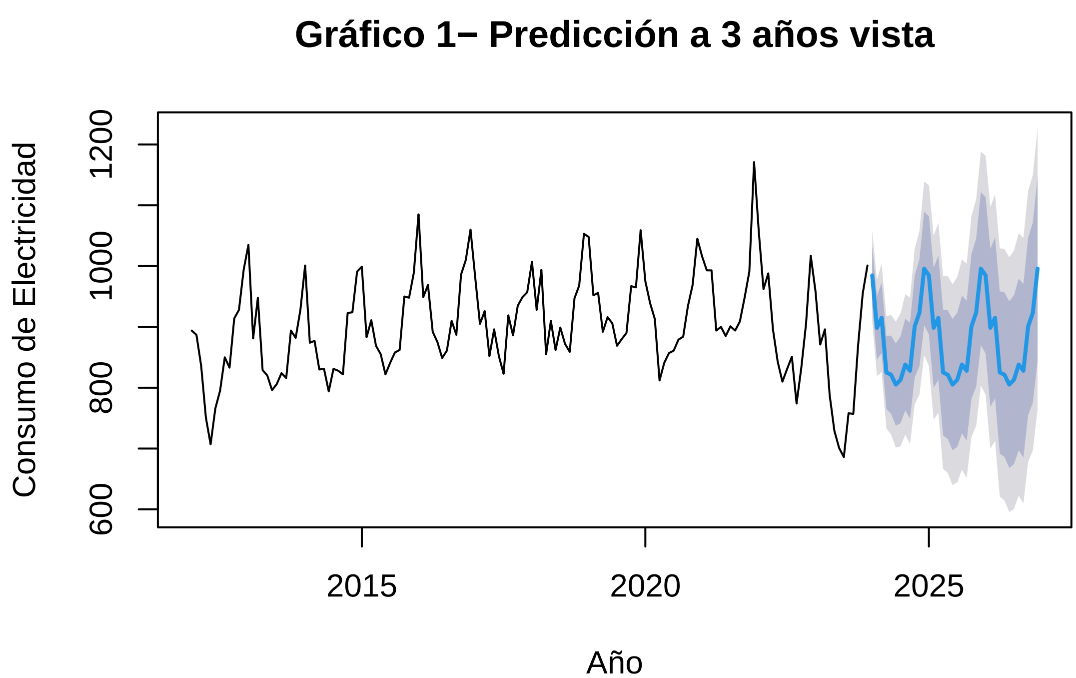
- **El error medio (ME = -0.0484)** confirma que el modelo **no tiene un sesgo notable**, ya que no tiende a sobrestimar ni subestimar el consumo eléctrico.
- **La raíz del error cuadrático medio (RMSE = 33.12)** refleja la variabilidad de los errores, mostrando que las desviaciones entre las predicciones y los valores reales son moderadas.

- **El error absoluto medio (MAE = 25.18)** indica que, en promedio, las predicciones difieren del valor real en aproximadamente 25 unidades de consumo eléctrico.
- **El error porcentual medio (MPE = -0.0989%)** señala que los errores no muestran una tendencia sistemática en términos relativos.
- **El error absoluto porcentual medio (MAPE = 2.81%)** evidencia que el modelo realiza **predicciones precisas**, con un margen de error bajo.
- **El error absoluto escalado medio (MASE = 0.576)** sugiere que el modelo **supera a un método ingenuo con estacionalidad**, lo que refuerza su calidad predictiva.
- **La autocorrelación de los residuos en el primer retardo (ACF1 = -0.0080)** indica que **los errores no siguen patrones sistemáticos**, lo que sugiere que la estructura de la serie ha sido bien capturada.

Estos resultados muestran que el modelo **ETS(M,N,A) ofrece predicciones precisas y errores bien distribuidos**, validando su adecuación para modelar la evolución del consumo eléctrico en Lituania.

Predicción a 3 años vista

```r
pronostico <- forecast(mejor_modelo, h = 36)
plot(pronostico, main = "Gráfico 1- Predicción a 3 años vista", ylab = "Consumo de Electricidad",
    xlab = "Año")
```



En la **Gráfica 1** se presenta la predicción del consumo eléctrico en Lituania para los próximos **3 años** utilizando el modelo **ETS(M,N,A)**, el cual ha demostrado ser el más adecuado para la serie. La serie original se muestra en **negro**, mientras que la predicción se representa en **azul**, con bandas de confianza en distintos tonos de **gris**.

Se observa que la predicción mantiene **un patrón estacional estable**, con aumentos y descensos regulares a lo largo del año, lo que indica que el modelo captura correctamente la **variabilidad cíclica del consumo**. La tendencia general del consumo parece **estable**, sin un crecimiento ni decrecimiento significativo, lo que sugiere que el comportamiento histórico se mantiene en las estimaciones futuras.

A medida que avanza el tiempo, las bandas de confianza se **ensanchan**, lo que refleja **mayor incertidumbre en las predicciones a largo plazo**. Esto es un comportamiento esperado en modelos de series temporales, ya que las pequeñas variaciones acumuladas generan mayor dispersión en las proyecciones futuras.

El modelo **ETS(M,N,A)** proporciona una estimación confiable en el corto y mediano plazo. Sin embargo, dada la creciente incertidumbre más allá del **primer año de predicción**, es recomendable actualizar y reevaluar el modelo periódicamente, especialmente considerando factores externos que puedan afectar el consumo energético en el país.

Calidad de las previsiones con origen de previsión móvil

Para evaluar la precisión del modelo **ETS(M,N,A)**, se ha utilizado la metodología de **origen de previsión móvil**, donde se recalcula el error en distintos horizontes de predicción.

```r
k <- 120
h <- 12
s <- length(electricidad) - k - h
mapeMovil <- matrix(NA, s + 1, h)
for (i in 0:s) {
    train.set <- subset(electricidad, start = i + 1, end = i + k)
    test.set <- subset(electricidad, start = i + k + 1, end = i + k + h)
    fit <- ets(train.set, model = "MNA", damped = FALSE)
    fcast <- forecast(fit, h = h)
    predicted_original <- fcast$mean
    mapeMovil[i + 1, ] <- 100 * abs(test.set - predicted_original)/test.set
}
errorMovil <- apply(mapeMovil, MARGIN = 2, FUN = median)
print(errorMovil)
```

```
##  [1]  2.701136  3.834796  3.904960  4.643505  8.190689  8.842868 10.578176
##  [8] 10.492894 10.773330 11.356017 11.667914 12.121960
```

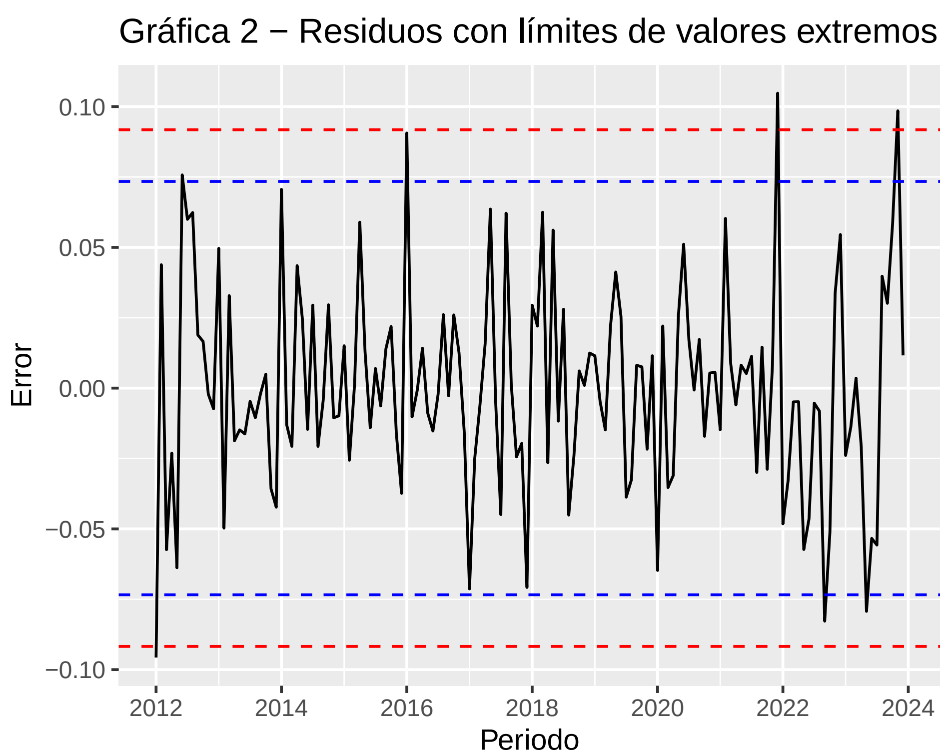
El análisis de los errores MAPE obtenidos revela que **la precisión del modelo ETS(M,N,A) es alta en el corto plazo, pero se deteriora conforme aumenta el horizonte de predicción**.

- Se observa que **el error de predicción aumenta con el horizonte de previsión**, lo que es un comportamiento esperado en modelos de series temporales.
- En los **primeros 3 meses**, el **MAPE se mantiene por debajo del 5%**, indicando una alta precisión en predicciones a corto plazo.
- A partir del **quinto mes**, el error comienza a **incrementarse de forma más pronunciada**, superando el **10% después del séptimo mes**.
- En los últimos meses de previsión, el error **se estabiliza en torno al 12%**, lo que sugiere una disminución en la confiabilidad del modelo a largo plazo.

Estos resultados indican que el modelo **ETS(M,N,A) ofrece buenas predicciones en el corto plazo**, pero su precisión disminuye a medida que se extiende el horizonte de previsión, debido a la acumulación de incertidumbre en la tendencia y la estacionalidad.

Valores extremos

```r
# Generar las fechas reales de la serie temporal
fechas <- seq(as.Date("2012-01-01"), by = "month", length.out = length(electricidad))
# Obtención de los residuos
residuos <- residuals(mejor_modelo)
# Cálculo de la desviación estándar de los residuos
sd_residuos <- sd(residuos)
limite_superior <- 2.5 * sd_residuos
limite_inferior <- -2.5 * sd_residuos
autoplot(residuos, series = "Error", colour = "black", xlab = "Periodo", ylab = "Error",
    main = "Gráfica 2 - Residuos con límites de valores extremos") + geom_hline(yintercept = c(-2.5,
    -2, 2, 2.5) * sd_residuos, colour = c("red", "blue", "blue", "red"), lty = 2) +
    scale_x_continuous(breaks = seq(2012, 2024, 2))
```




```r
valores_extremos <- fechas[abs(residuos) > limite_superior]
valores_extremos
```

```
## [1] "2012-01-01" "2021-12-01" "2023-11-01"
```

La **Gráfica 2** muestra los residuos del modelo **ETS(M,N,A)** con límites de **±2.5 desviaciones estándar**, en lugar del umbral de 3 que se había considerado previamente. Esta reducción ha permitido detectar **tres valores extremos** en la serie: **enero de 2012, diciembre de 2021 y noviembre de 2023**.

Valores Extremos Detectados

Se han identificado tres valores extremos en la serie. **Enero de 2012** coincide con una **fuerte ola de frío en Europa**, lo que incrementó la demanda de calefacción y provocó un aumento inusual en el consumo eléctrico.

Diciembre de 2021 marca el inicio de la **crisis energética en Europa**, cuando la incertidumbre en el suministro de gas disparó los precios y generó fluctuaciones en la demanda. **Noviembre de 2023** podría estar relacionado con **ajustes en el mercado energético post-crisis**, reflejando cambios en la oferta y demanda.

Estos eventos externos han afectado el comportamiento del consumo eléctrico, generando anomalías en los residuos del modelo. Aunque el **ETS(M,N,A)** captura bien la estructura de la serie, la inclusión de **variables exógenas** podría mejorar su capacidad predictiva en contextos de alta volatilidad.

Comparación del modelo ajustado por Alisado Exponencial con el método ingenuo con estacionalidad

Se ha comparado el modelo **ETS(M,N,A)** con el **método ingenuo con estacionalidad**, evaluando la calidad del ajuste mediante diversas métricas.

```
metodo_ingenuo <- snaive(electricidad, h = 36)
mase_alisado <- accuracy(mejor_modelo)[, "MASE"]
mase_ingenuo <- accuracy(metodo_ingenuo)[, "MASE"]
cat("MASE del modelo de alisado:", mase_alisado, "\n")
```

```
## MASE del modelo de alisado: 0.5760404
```

```
cat("MASE del método ingenuo:", mase_ingenuo, "\n")
```

```
## MASE del método ingenuo: 1
```

El modelo **ETS(M,N,A)** obtiene un **MASE de 0.576**, mientras que el método ingenuo tiene un **MASE de 1**. Esto indica que el modelo de alisado exponencial logra reducir el error en aproximadamente **42%** en comparación con la predicción basada en estacionalidad simple.

Por lo que, podemos concluir que el modelo **ETS(M,N,A)** es más preciso que el método ingenuo, mostrando un mejor ajuste a la serie y proporcionando predicciones más fiables.

Conclusiones

El análisis del **consumo eléctrico en Lituania** ha permitido modelar su evolución con **alisado exponencial (ETS)**, determinando que el modelo **ETS(M,N,A)** es el más adecuado. Su **MAPE del 2.79%** indica una **buena precisión en la predicción**, y su **MASE de 0.57** confirma que supera al método ingenuo con estacionalidad.

La **predicción a tres años** muestra que el consumo eléctrico **mantendrá su patrón estacional**, aunque con **mayor incertidumbre a largo plazo**, reflejada en el ensanchamiento de las bandas de confianza.

El análisis de **residuos** ha identificado **tres valores extremos** (enero 2012, diciembre 2021 y noviembre 2023), asociados a eventos como **olas de frío, la crisis energética y ajustes en la demanda post-crisis**. Estos eventos reflejan la influencia de factores externos en la serie.