

Procesos Estocásticos y Series Temporales

Práctica 2: Descomposición de series temporales

Francisco Javier Mercader Martínez

Problema 1

En el fichero **consumo_leche.txt** se encuentran los datos correspondientes al consumo de leche mensual (en miles de litros) por los habitantes de una determinada ciudad. Se dispone de datos sobre el consumo desde noviembre de 1992 hasta octubre de 2006.

- 1) Representa los datos del consumo de leche en un gráfico temporal y comenta los aspectos más relevantes. ¿La serie presenta Estacionalidad? ¿cómo dirías que es la tendencia?

```
# Cargar los datos
datos_leche <- read.table("consumo_leche.txt", header = TRUE)

# Verificar la estructura de los datos
head(datos_leche)

##      leche
## 1    589
## 2    561
## 3    640
## 4    656
## 5    727
## 6    697

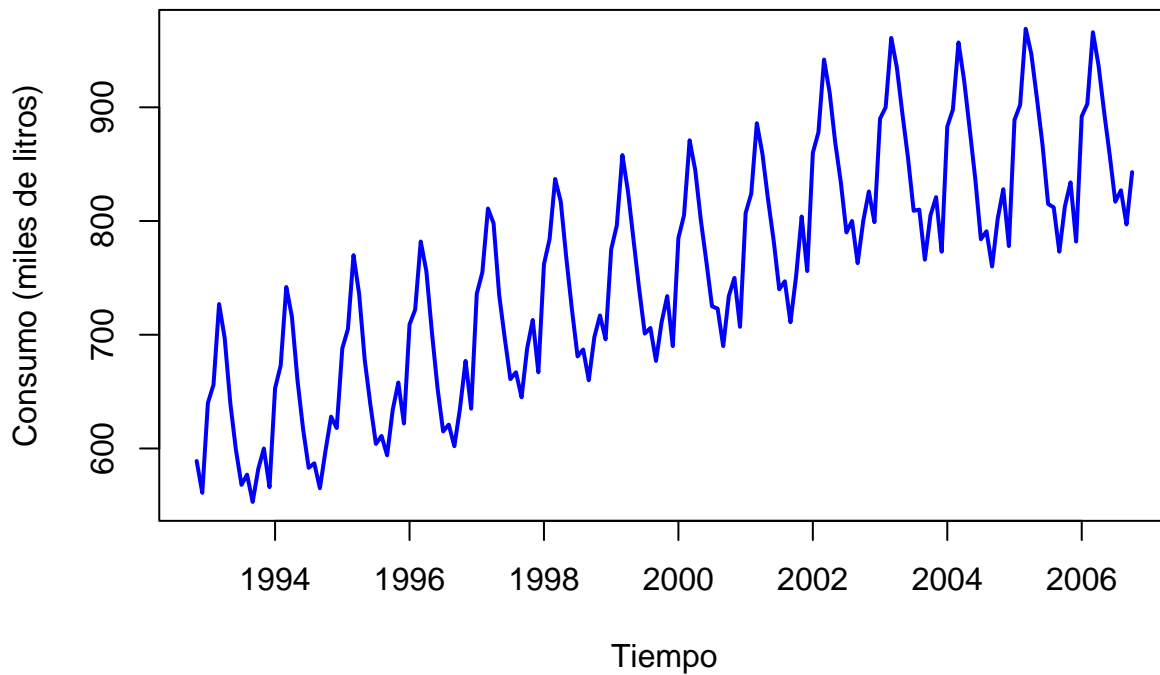
str(datos_leche)

## 'data.frame':    168 obs. of  1 variable:
## $ leche: int  589 561 640 656 727 697 640 599 568 577 ...

# Convertir a serie temporal
# Datos mensuales desde noviembre de 1992 hasta octubre de 2006
# Frecuencia = 12 (datos mensuales)
# start = c(1992, 11) indica que comienza en noviembre de 1992
leche_ts <- ts(datos_leche$leche, frequency = 12, start = c(1992, 11))

# Representación gráfica
plot(leche_ts,
     main = "Consumo mensual de leche (Nov 1992 - Oct 2006)",
     xlab = "Tiempo",
     ylab = "Consumo (miles de litros)",
     col = "blue",
     lwd = 2)
```

Consumo mensual de leche (Nov 1992 – Oct 2006)



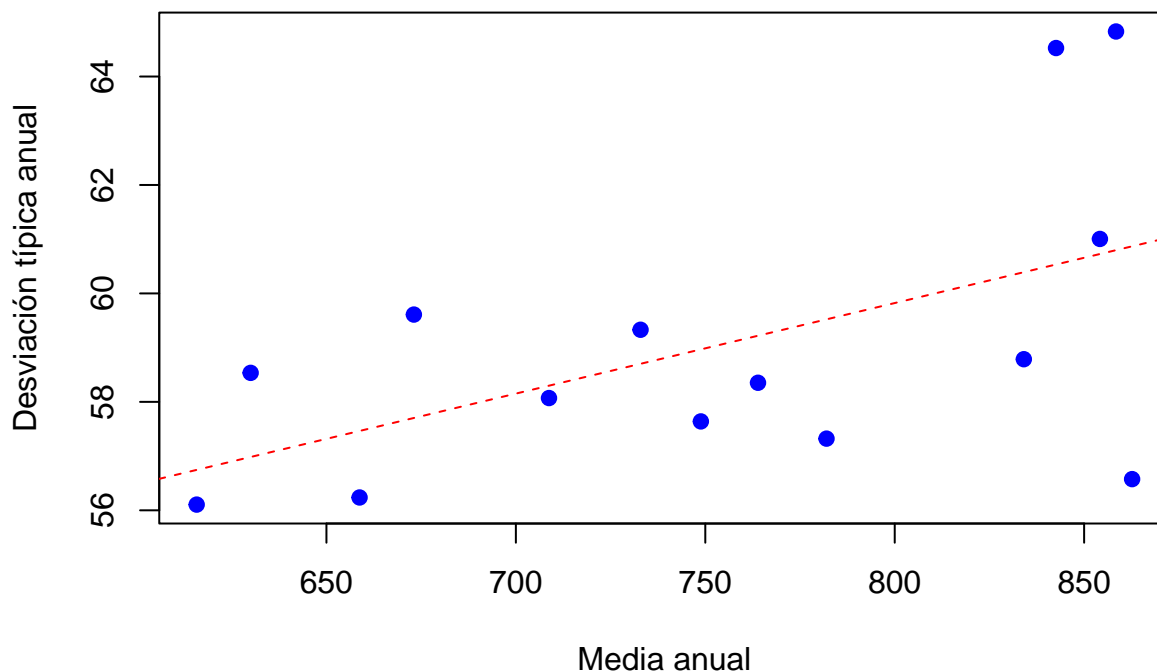
La serie presenta una clara estacionalidad, con picos de consumo en ciertos meses del año. La tendencia parece ser creciente a lo largo del tiempo, indicando un aumento en el consumo de leche.

- 2) Determina si se trata de un modelo aditivo o multiplicativo (realiza un gráfico de desviaciones típicas frente a medias para cada año).

```
# Análisis de la relación entre media y desviación típica por año
medias_anuales <- as.vector(aggregate(leche_ts, FUN = mean))
desviaciones_anuales <- as.vector(aggregate(leche_ts, FUN = sd))

plot(medias_anuales, desviaciones_anuales,
     main = "Relación Media vs Desviación Típica",
     xlab = "Media anual",
     ylab = "Desviación típica anual",
     pch = 19,
     col = "blue")
# Añadimos una línea de regresión para visualizar la relación
abline(lm(desviaciones_anuales ~ medias_anuales), col = "red", lty = 2)
```

Relación Media vs Desviación Típica



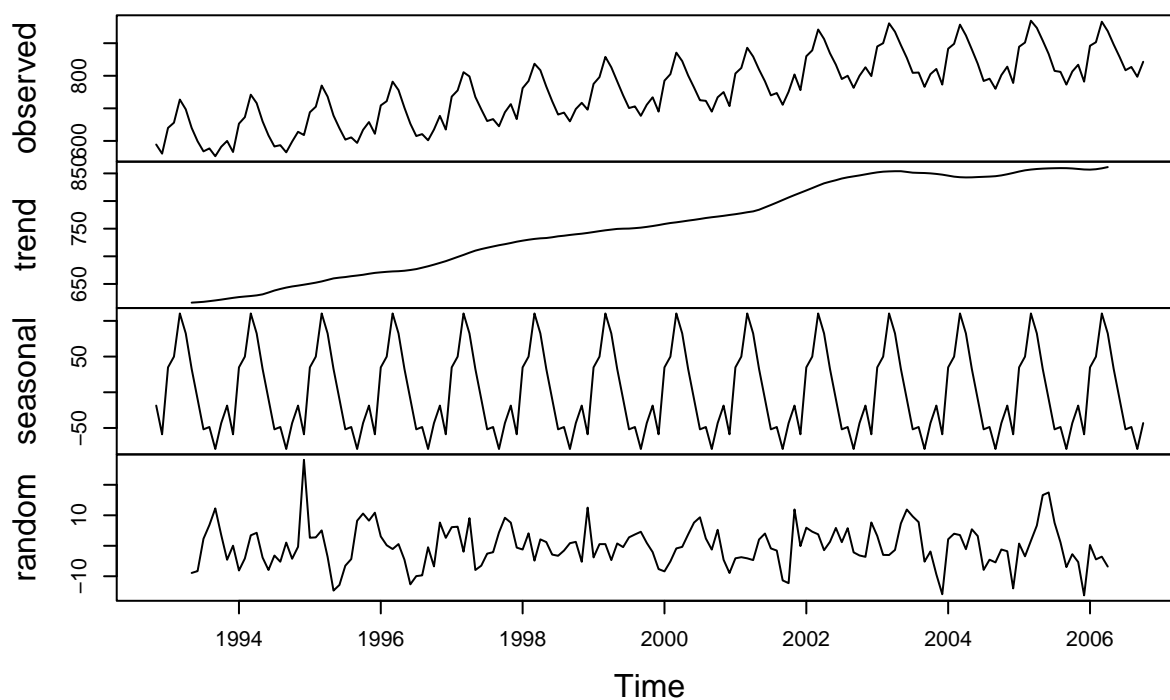
La relación entre la media y la desviación típica sugiere que la serie es aditiva, ya que no se observa un aumento proporcional de la desviación típica con la media.

- 3) **Mediante Análisis Clásico** (método de la diferencia o de la razón a la media móvil), extrae las componentes de la serie (Tendencia-Ciclo, Estacionalidad e Irregular) y comenta los resultados.

```
# Realizar la descomposición clásica aditiva
descomposicion_clasica <- decompose(leche_ts, type = "additive")

# Representación conjunta de todas las componentes
plot(descomposicion_clasica)
```

Decomposition of additive time series



- 4) Obtén un modelo determinista que nos permita realizar predicciones para la serie original.

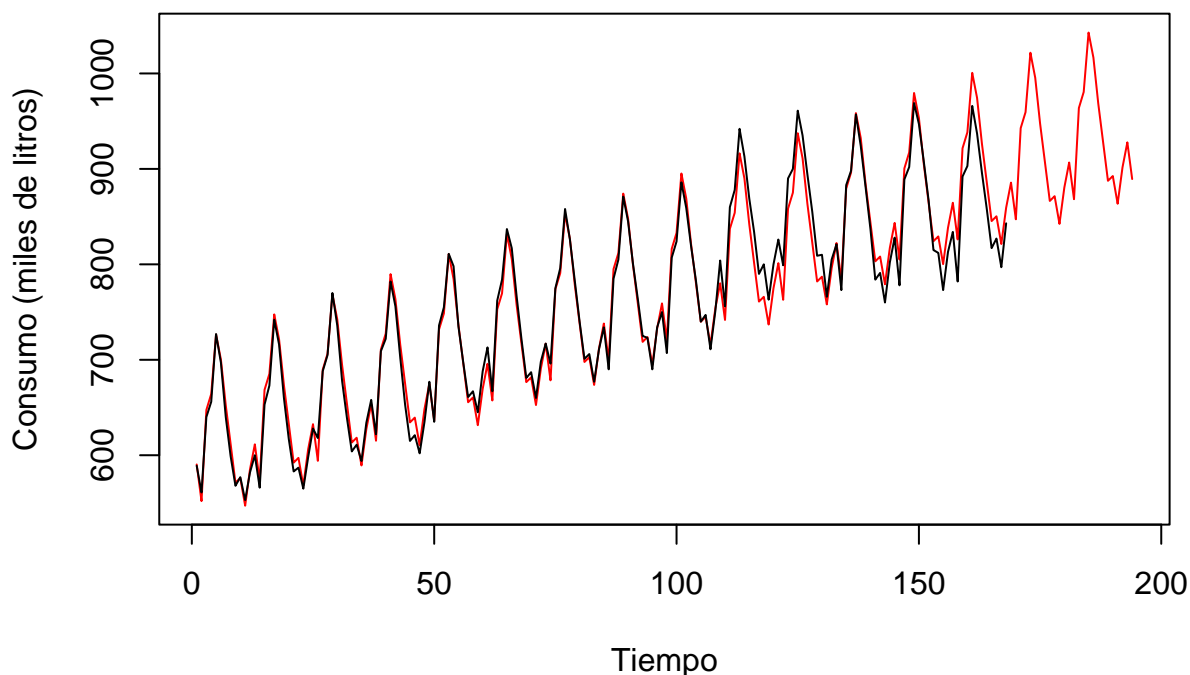
```
tiempo <- 1:length(leche_ts)
tendencia.lm <- lm(descomposicion_clasica$trend ~ tiempo)
tendencia.lm
```

```
##
## Call:
## lm(formula = descomposicion_clasica$trend ~ tiempo)
##
## Coefficients:
## (Intercept)      tiempo
##      607.270       1.758
```

- 5) Con el modelo del apartado anterior, calcula los valores ajustados de la serie para los meses observados (168 meses) y realiza la predicción para los meses restantes de 2006 (noviembre y diciembre) y los dos años siguientes al completo (2007 y 2008). ¿Cuánto será el consumo de leche en noviembre de 2008?

```
tiempo_pred <- 1:(length(leche_ts) + 2 + 24) # 2 meses restantes de 2006 + 24 meses de 2007 y
↪ 2008
estacional_pred <- c(descomposicion_clasica$seasonal,
                     descomposicion_clasica$seasonal[1:26]) # Repetir estacionalidad para los
↪ meses adicionales
tendencia_pred <- predict(tendencia.lm, data.frame(tiempo = tiempo_pred))

datos_pred <- tendencia_pred + estacional_pred
plot(tiempo_pred, datos_pred, type = "l", col = 'red',
     xlab = "Tiempo", ylab = "Consumo (miles de litros)")
lines(tiempo, leche_ts, type = "l")
```



```
consumo_nov_2008 <- datos_pred[length(leche_ts) + 26 - 1]
cat("Consumo predicho en noviembre de 2008:", consumo_nov_2008, "(miles de litros)")
```

```
## Consumo predicho en noviembre de 2008: 927.7984 (miles de litros)
```

- 6) Representa en un mismo gráfico la secuencia de la serie observada (en negro), la serie ajustada (en azul) y de la serie predicha (en rojo) con el modelo determinista.

```
# Gráfico con serie observada, ajustada y predicha
plot(tiempo_pred, datos_pred, type = "l", col = "red", lwd = 2,
     main = "Serie observada, ajustada y predicha",
     xlab = "Tiempo",
     ylab = "Consumo (miles de litros)",
     xlim = c(1, length(tiempo_pred)),
     ylim = range(c(leche_ts, datos_pred), na.rm = TRUE))
```

```

# Serie observada (negro)
lines(tiempo, leche_ts, type = "l", col = "black", lwd = 2)

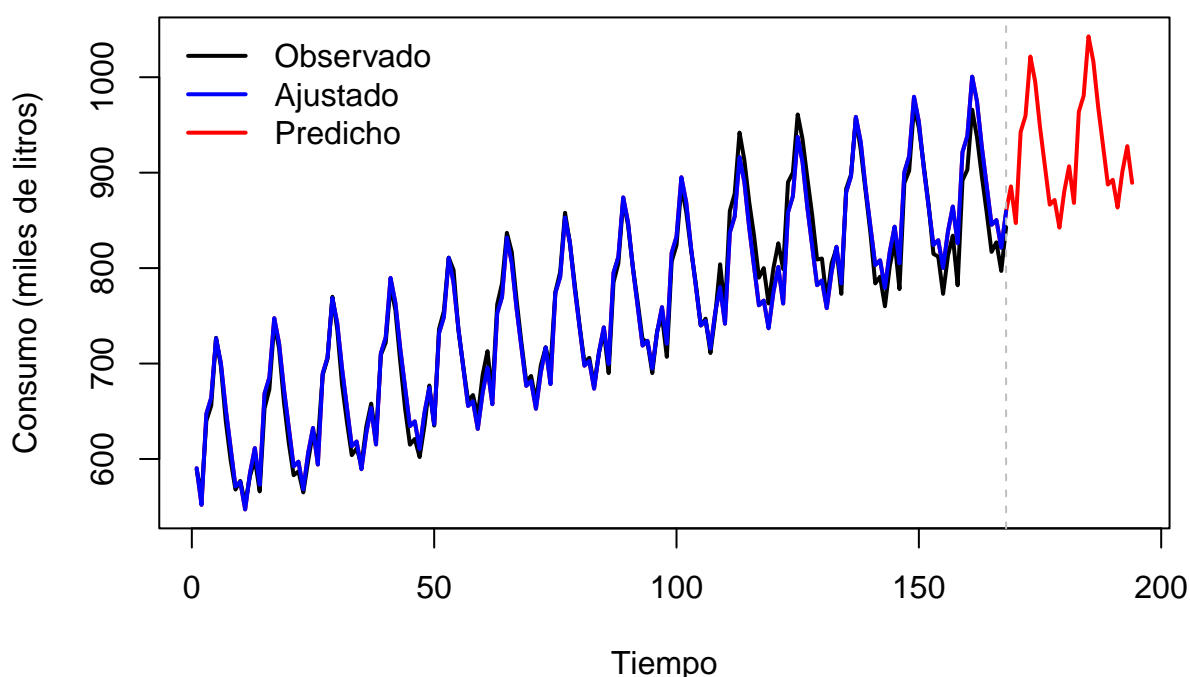
# Serie ajustada (azul) - solo para los datos observados
lines(tiempo, datos_pred[1:length(leche_ts)], type = "l", col = "blue", lwd = 2)

# Añadir línea vertical para separar observado de predicho
abline(v = length(leche_ts), lty = 2, col = "gray")

# Leyenda
legend("topleft",
      legend = c("Observado", "Ajustado", "Predicho"),
      col = c("black", "blue", "red"),
      lwd = 2,
      bty = "n")

```

Serie observada, ajustada y predicha



Problema 2

Con los datos de **consumo_leche.txt**, responder a los apartados del problema anterior pero usando la **descomposición STL**. Selecciona un ancho de banda para estimar la componente estacional de 5 años. Comparar las medidas de error MAE y RMSE en el tramo observado para ambos métodos (Análisis Clásico y STL). Comentar los resultados.

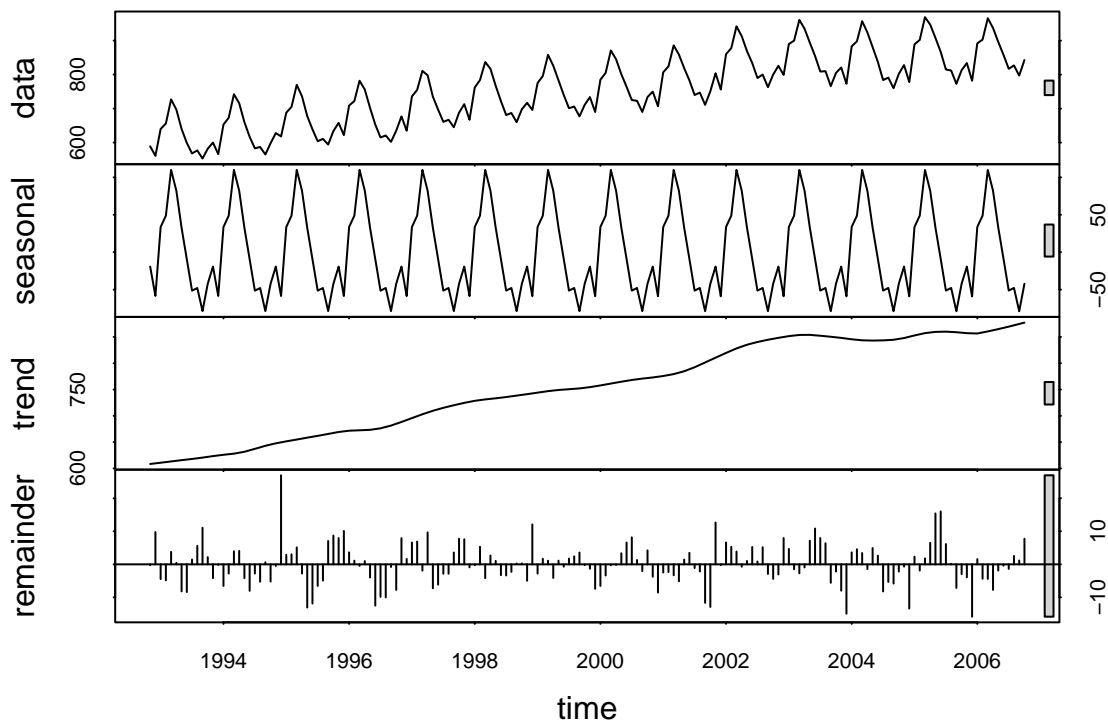
```

# 1) Descomposición STL con ancho de banda de 5 años
# s.window debe ser impar: 5 años * 12 meses = 60, usamos 61
STL_leche <- stl(leche_ts, s.window = 61)

# Representación de las componentes
plot(STL_leche, main = "Descomposición STL del consumo de leche")

```

Descomposición STL del consumo de leche



2) Extraer componentes

```
estacional_STL <- STL_leche$time.series[, 1]
tendencia_STL <- STL_leche$time.series[, 2]
irregular_STL <- STL_leche$time.series[, 3]
```

3) Modelo determinista con STL

```
tiempo_STL <- 1:length(leche_ts)
tendencia_STL.lm <- lm(tendencia_STL ~ tiempo_STL)
```

4) Predicciones con STL

```
tiempo_pred_STL <- 1:(length(leche_ts) + 26)
estacional_STL_pred <- c(estacional_STL, estacional_STL[1:26])
tendencia_STL_pred <- predict(tendencia_STL.lm, data.frame(tiempo_STL = tiempo_pred_STL))

datos_STL_pred <- tendencia_STL_pred + estacional_STL_pred
```

Consumo en noviembre de 2008

```
consumo_nov_2008_STL <- datos_STL_pred[length(leche_ts) + 26 - 1]
cat("Consumo predicho en noviembre de 2008 (STL):", consumo_nov_2008_STL, "(miles de litros)")
```

```
## Consumo predicho en noviembre de 2008 (STL): 922.9582 (miles de litros)
```

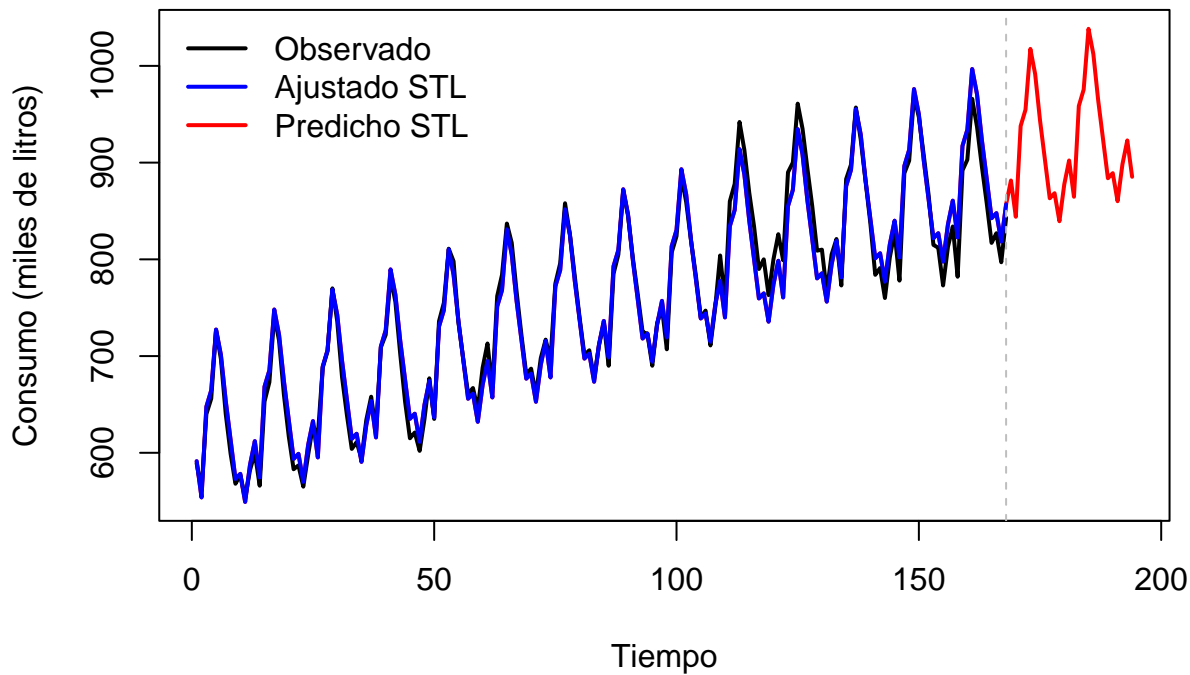
5) Gráfico con serie observada, ajustada y predicha (STL)

```
plot(tiempo_pred_STL, datos_STL_pred, type = "l", col = "red", lwd = 2,
     main = "Serie observada, ajustada y predicha (STL)",
     xlab = "Tiempo", ylab = "Consumo (miles de litros)",
     xlim = c(1, length(tiempo_pred_STL)),
     ylim = range(c(leche_ts, datos_STL_pred), na.rm = TRUE))

lines(tiempo, leche_ts, type = "l", col = "black", lwd = 2)
lines(tiempo, datos_STL_pred[1:length(leche_ts)], type = "l", col = "blue", lwd = 2)
abline(v = length(leche_ts), lty = 2, col = "gray")

legend("topleft",
      legend = c("Observado", "Ajustado STL", "Predicho STL"),
      col = c("black", "blue", "red"),
      lwd = 2, bty = "n")
```

Serie observada, ajustada y predicha (STL)



```
# 6) Comparación de errores: Análisis Clásico vs STL
# Residuos Análisis Clásico
residuos_clasico <- leche_ts - datos_pred[1:length(leche_ts)]
MAE_clasico <- mean(abs(residuos_clasico), na.rm = TRUE)
RMSE_clasico <- sqrt(mean(residuos_clasico^2, na.rm = TRUE))

# Residuos STL
residuos_STL <- leche_ts - datos_STL_pred[1:length(leche_ts)]
MAE_STL <- mean(abs(residuos_STL), na.rm = TRUE)
RMSE_STL <- sqrt(mean(residuos_STL^2, na.rm = TRUE))

# Mostrar resultados
cat("\n=== COMPARACIÓN DE MÉTODOS ===\n")
```

```
##
## === COMPARACIÓN DE MÉTODOS ===
```

```
cat("Análisis Clásico:\n")
```

```
## Análisis Clásico:
```

```
cat("  MAE:", round(MAE_clasico, 4), "\n")
```

```
##    MAE: 11.9645
```

```
cat("  RMSE:", round(RMSE_clasico, 4), "\n\n")
```

```
##    RMSE: 15.6179
```

```
cat("Descomposición STL:\n")
```

```
## Descomposición STL:
```

```
cat("  MAE:", round(MAE_STL, 4), "\n")
```

```
##    MAE: 11.8879
```

```
cat("  RMSE:", round(RMSE_STL, 4), "\n")
```

```
##    RMSE: 15.5117
```

La diferencia entre los métodos puede observarse en las medidas de error MAE y RMSE. Debido a que numéricamente

hablando obtenemos resultados muy similares, podemos concluir que ambos métodos son adecuados para este conjunto de datos, aunque la descomposición STL puede ofrecer una mayor flexibilidad en la modelización de la estacionalidad.