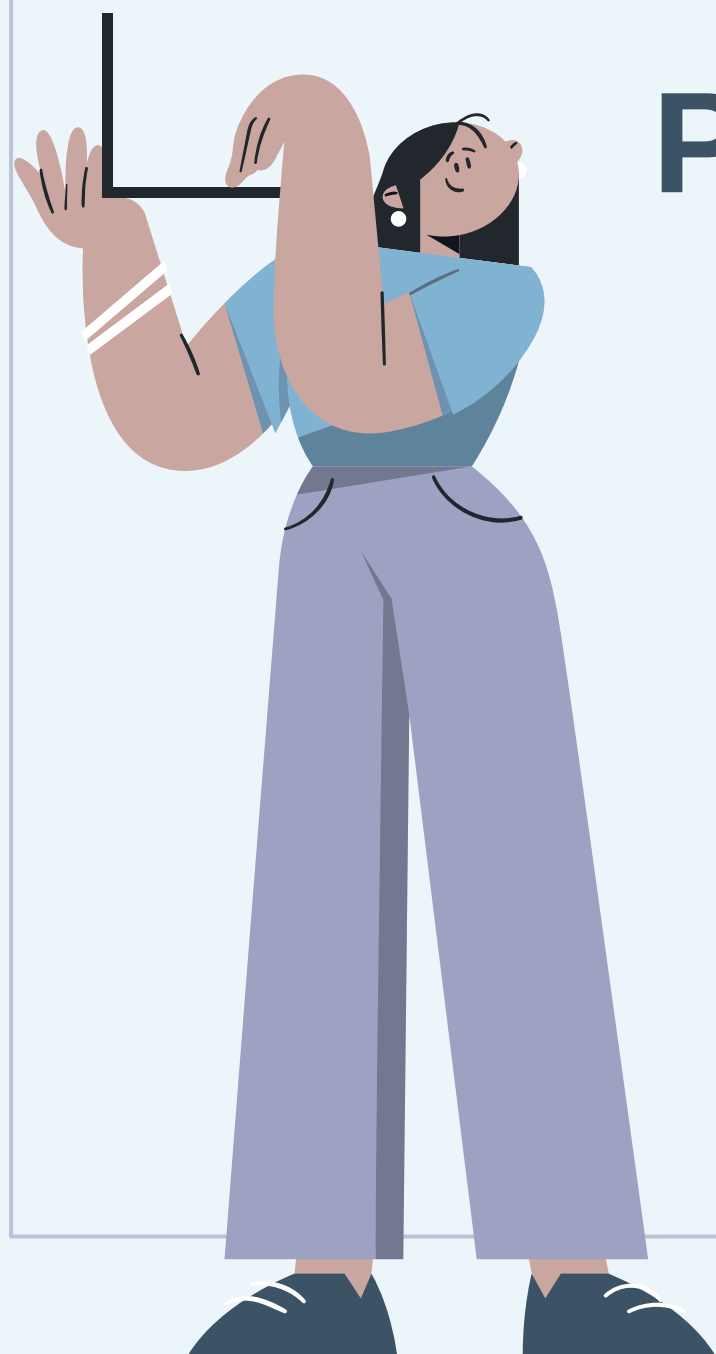
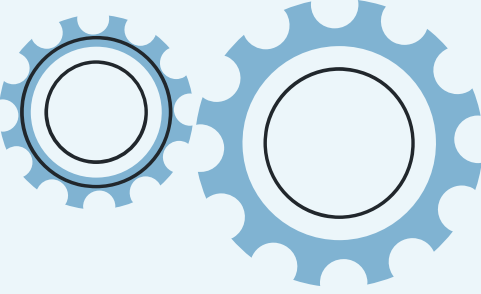


Production industrielle électronique et Dépenses de construction





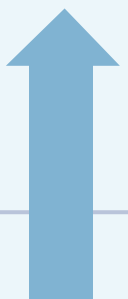
Sommaire

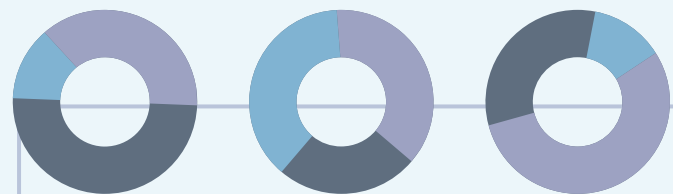


01 Analyse statistique et graphique

02 Ajustement de la série

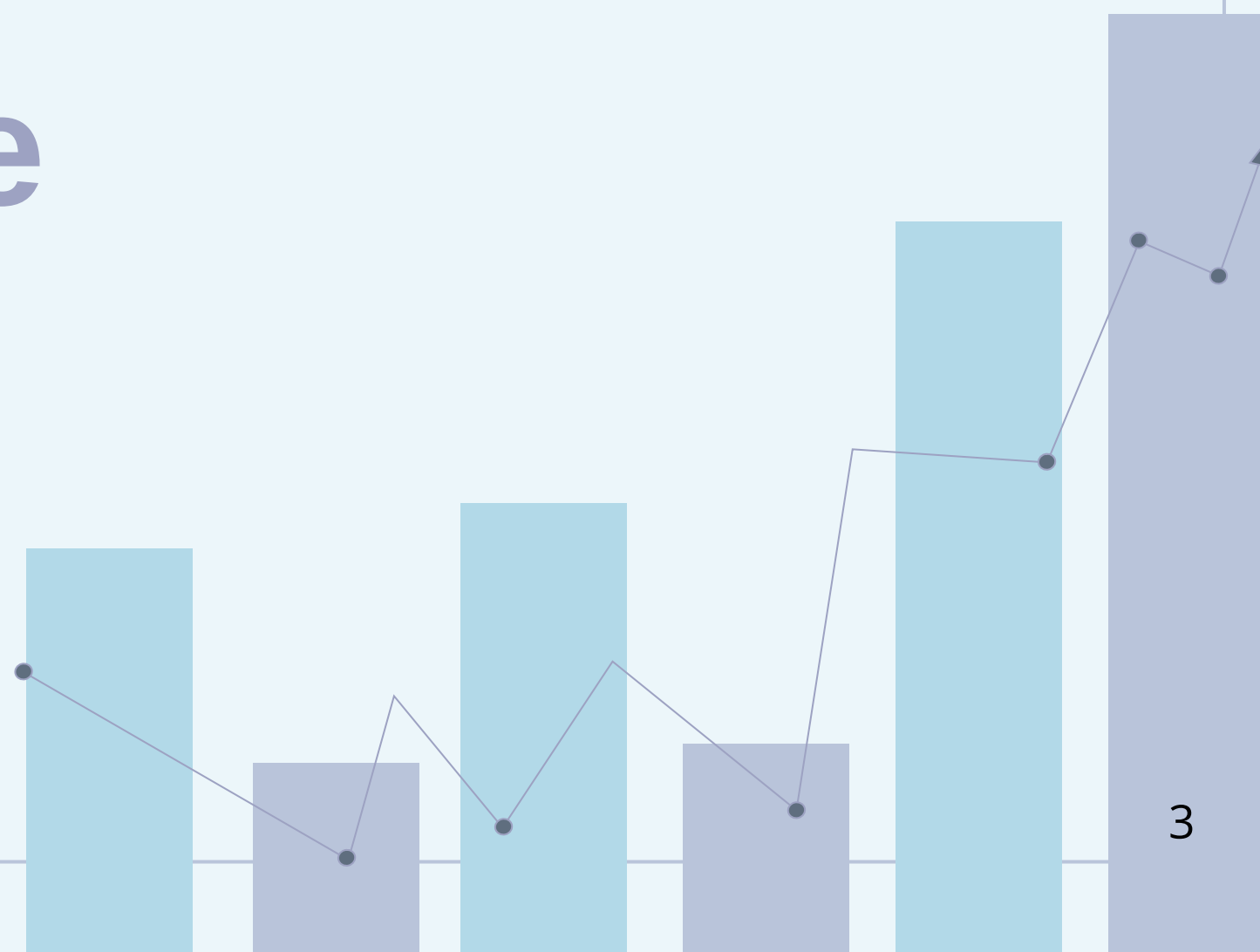
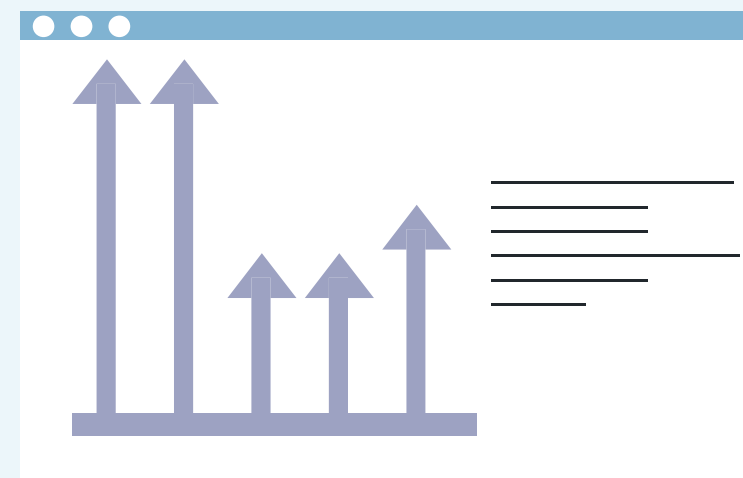
03 Analyse et modélisation des résidus



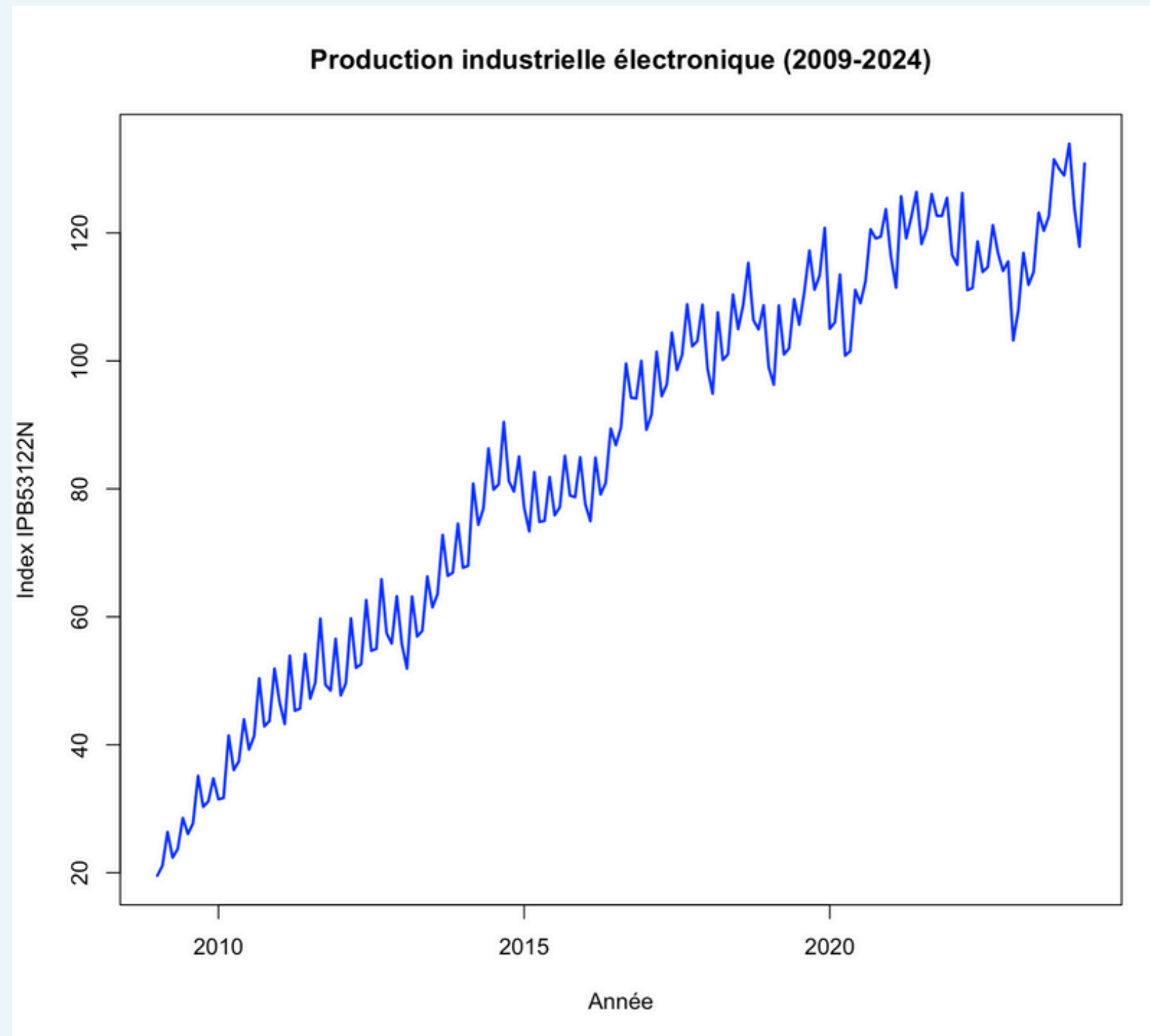


Partie 1

Data Set 1: Production industrielle électronique



Visualisation de la serie



- Le graphique montre l'évolution de la production industrielle électronique entre 2009 et 2024.
- On observe une tendance générale à la hausse, indiquant une forte croissance du secteur sur cette période.
- La courbe présente également une saisonnalité marquée, avec des fluctuations régulières, probablement liées à des cycles de production ou à la demande du marché.
- Quelques périodes de ralentissement, notamment autour de 2015 et 2024

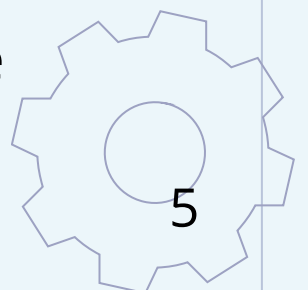
Informations Sur DATA

$\text{corr}(X, Y)$

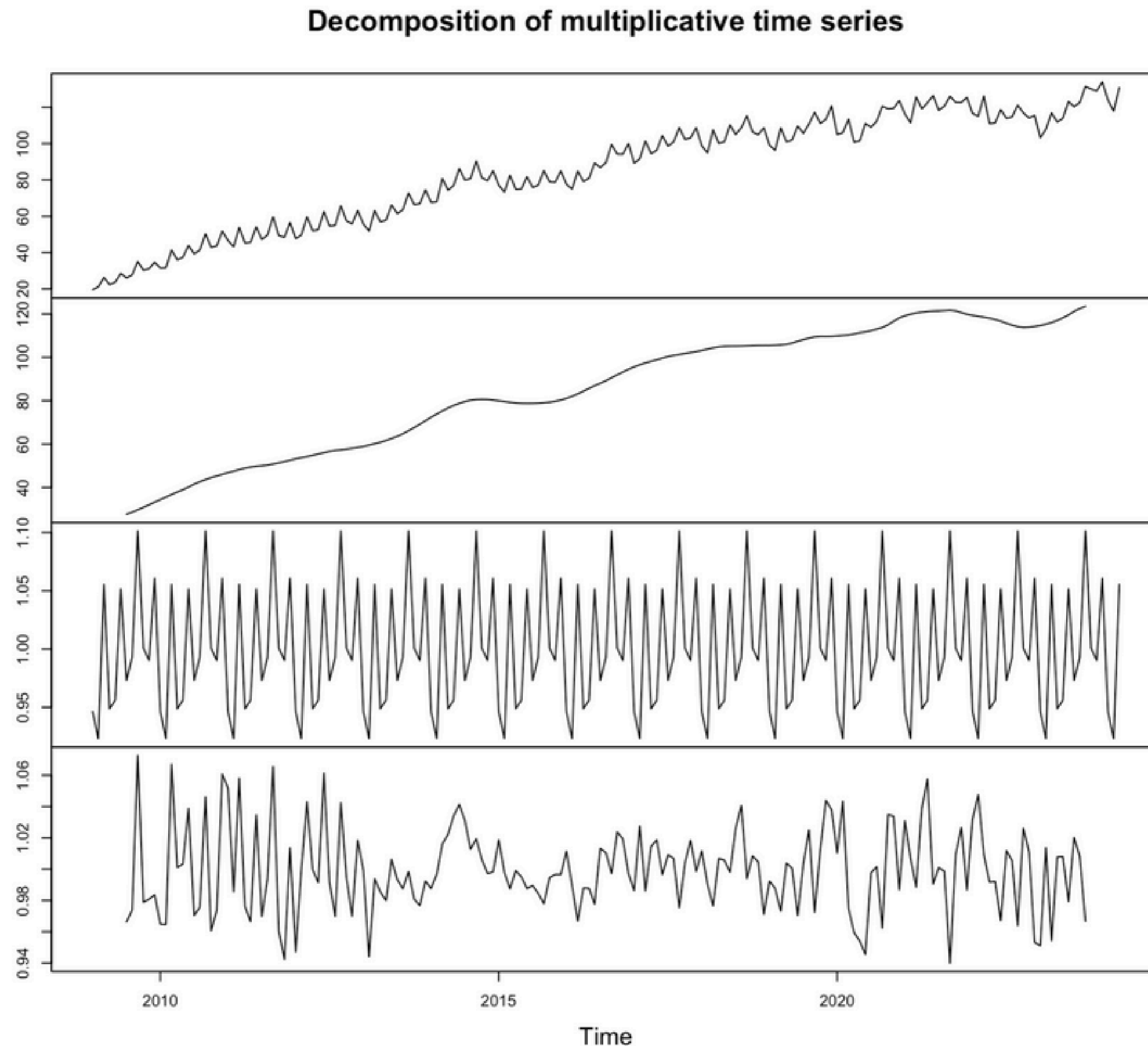
```
> summary(production_electronique)
observation_date      IPB53122N
Min.      :2009-01-01  Min.      : 19.55
1st Qu.   :2012-10-16  1st Qu.   : 57.63
Median    :2016-08-01  Median    : 89.57
Mean      :2016-07-31  Mean      : 85.08
3rd Qu.   :2020-05-16  3rd Qu.   :111.26
Max.      :2024-03-01  Max.      :133.94
```

```
> mean(production_electronique_ts)
[1] 85.07533
> # Variance
> var(production_electronique_ts)
[1] 941.1843
> # Écart-type
> sd(production_electronique_ts)
[1] 30.67873
```

- Les statistiques descriptives de la série temporelle indiquent une forte progression de la production industrielle électronique entre 2009 et 2024.
- L'indice IB53122N est passé d'un minimum de 19,55 en janvier 2009 à un maximum de 133,94 en mars 2024, traduisant une multiplication par plus de 6 en 15 ans.
- La dispersion entre les quartiles (du 1er quartile à 57,63 jusqu'au 3e à 111,26) reflète à la fois une croissance soutenue et une variabilité progressive, probablement liée à la saisonnalité ou aux cycles économiques.



Décomposition de la série temporelle



Interpretation :

✕ Modèle :


Décomposition multiplicative car l'amplitude des variations est proportionnelle à la tendance

 **La tendance :**

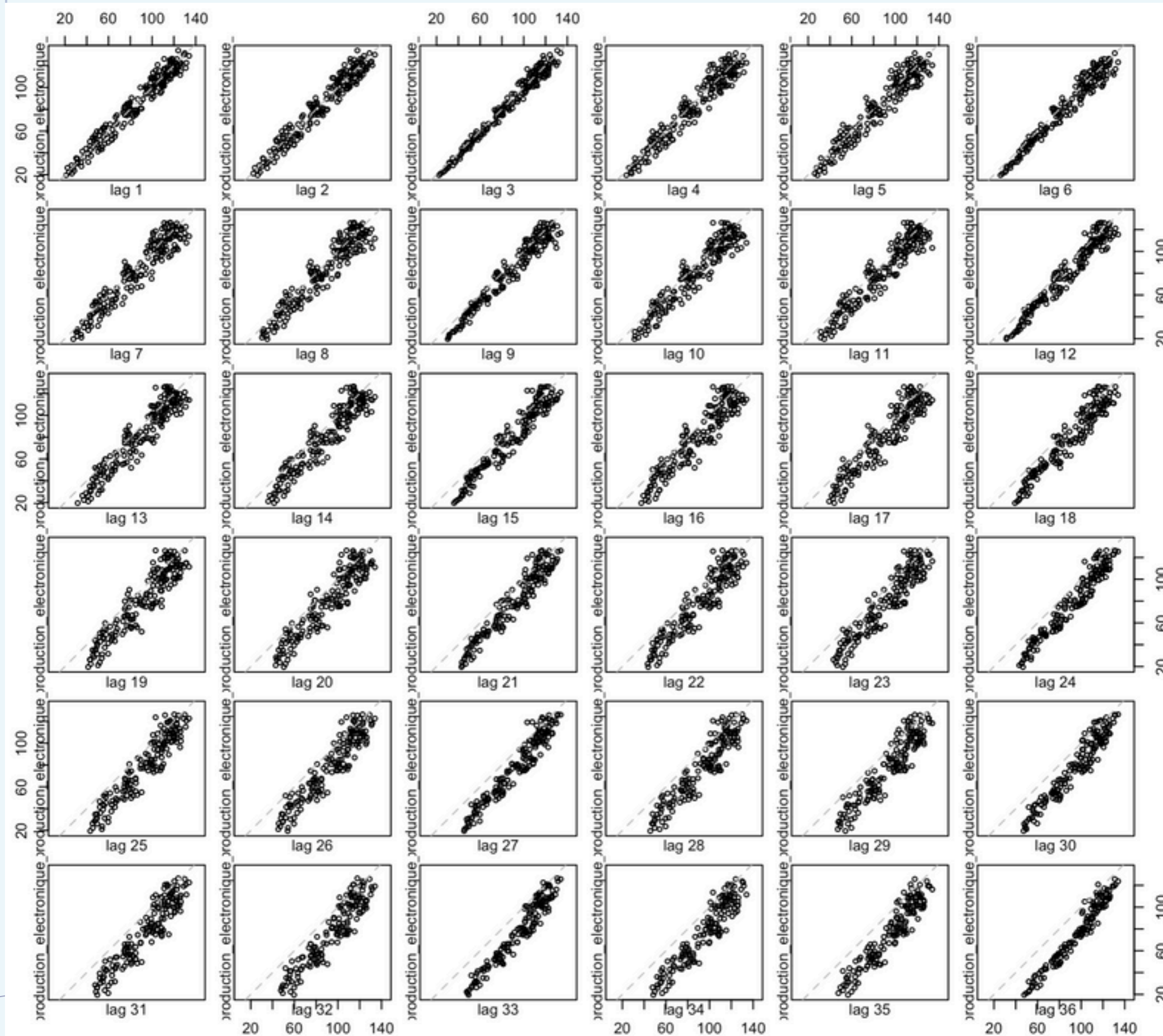
indique une forte croissance au début, suivie d'un ralentissement récent.

 **Saisonnalité :**

Stable et régulière

 Les résidus sont modérés, représentant des événements aléatoires.

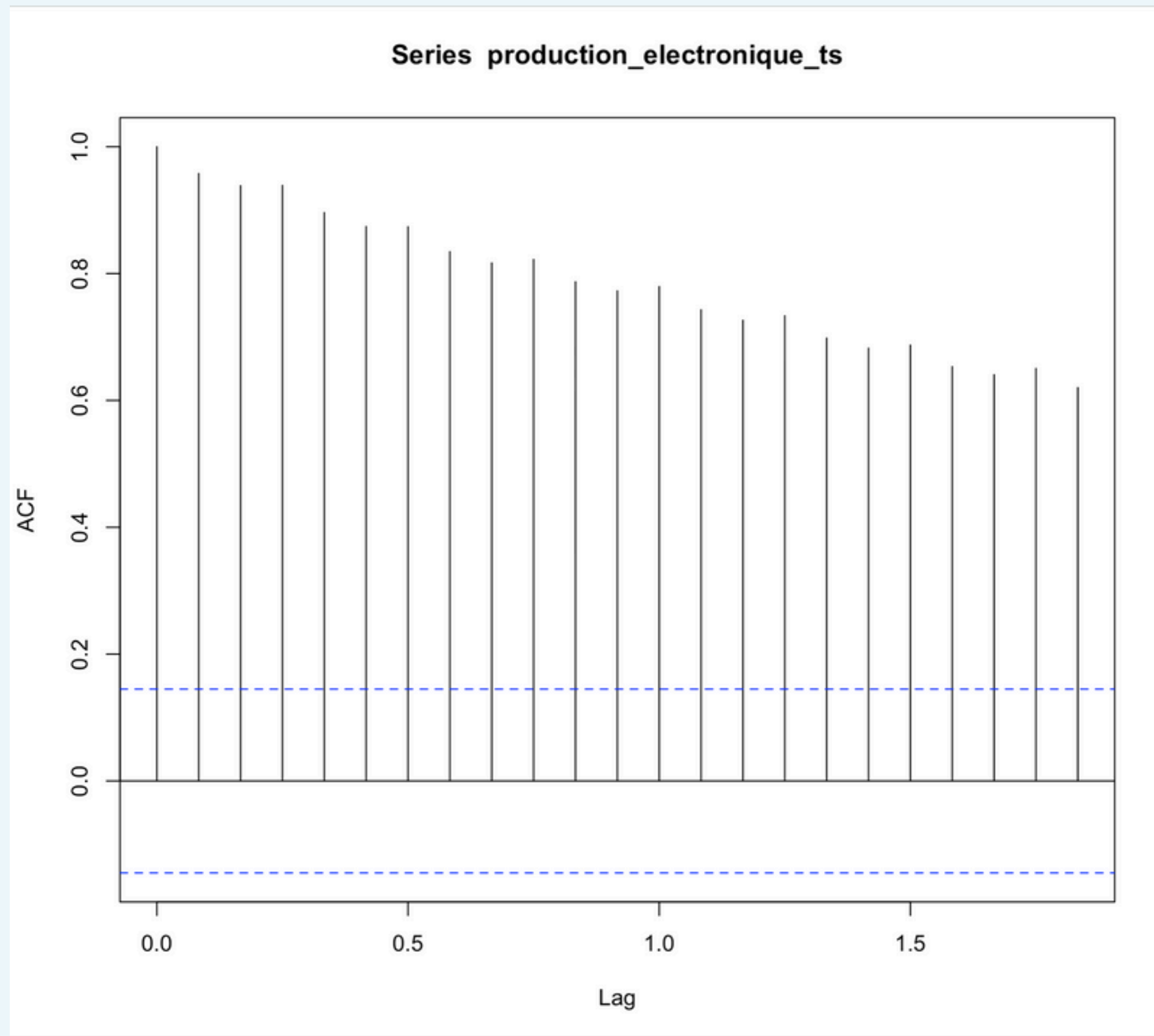
Autocorrelation



Interpretation :

- 📍 Les points suivent une forme proche d'une droite
- ➡ Il existe une forte corrélation entre les valeurs successives de la série temporelle.

Calcul de l'autocorrélation (ACF)

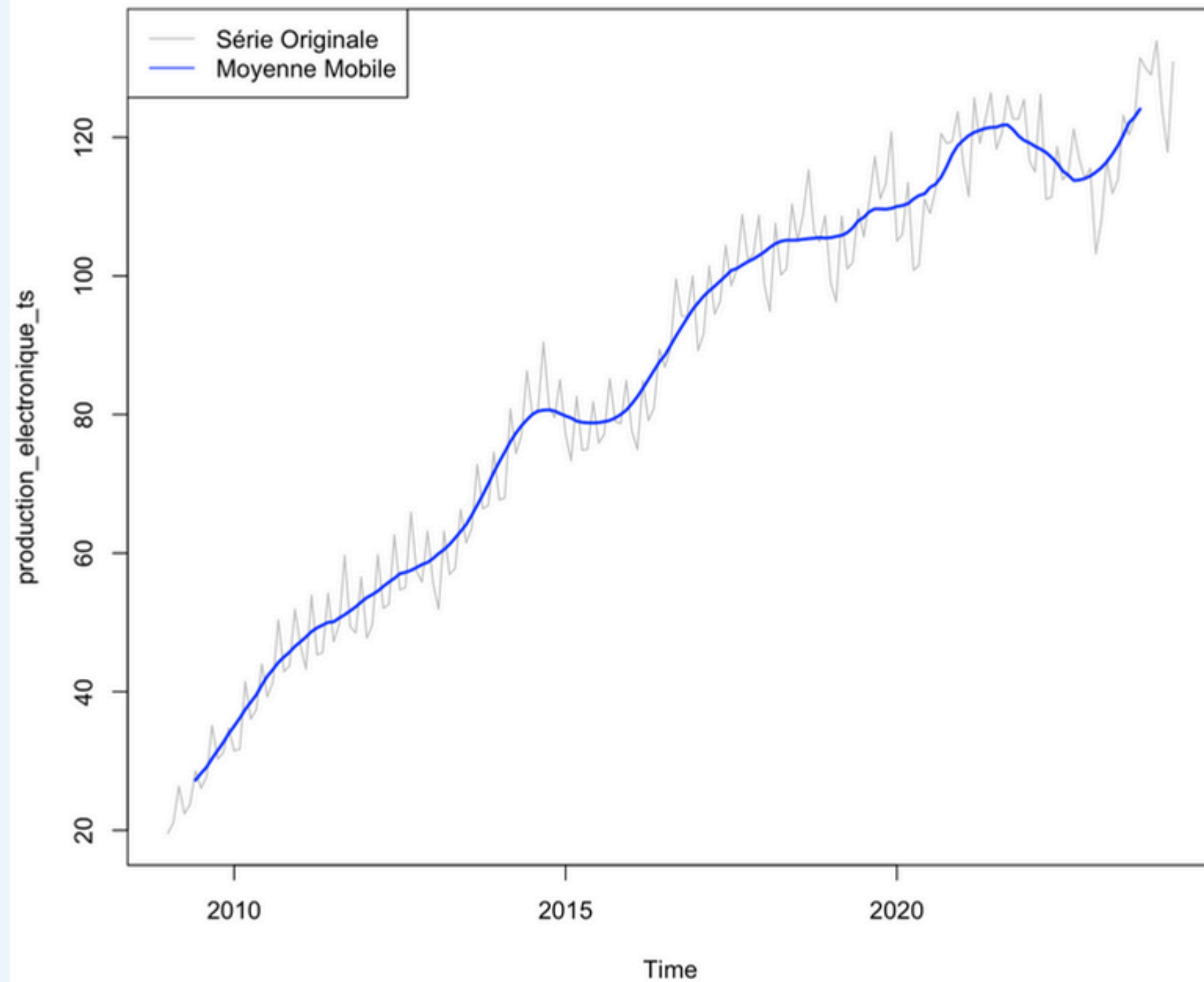


Interpretation :

▼ La série production_electronique_ts montre une autocorrélation qui diminue progressivement à et significativement. Cela suggère qu'il pourrait y avoir une forte influence des valeurs récentes sur les valeurs passées

Ajustement de la tendance par moyenne mobile

Ajustement par moyenne mobile



Call:

```
lm(formula = production_electronique_ts ~ temps)
```

Residuals:

Min	1Q	Median	3Q	Max
-24.9493	-4.9350	-0.2682	5.7103	18.2411

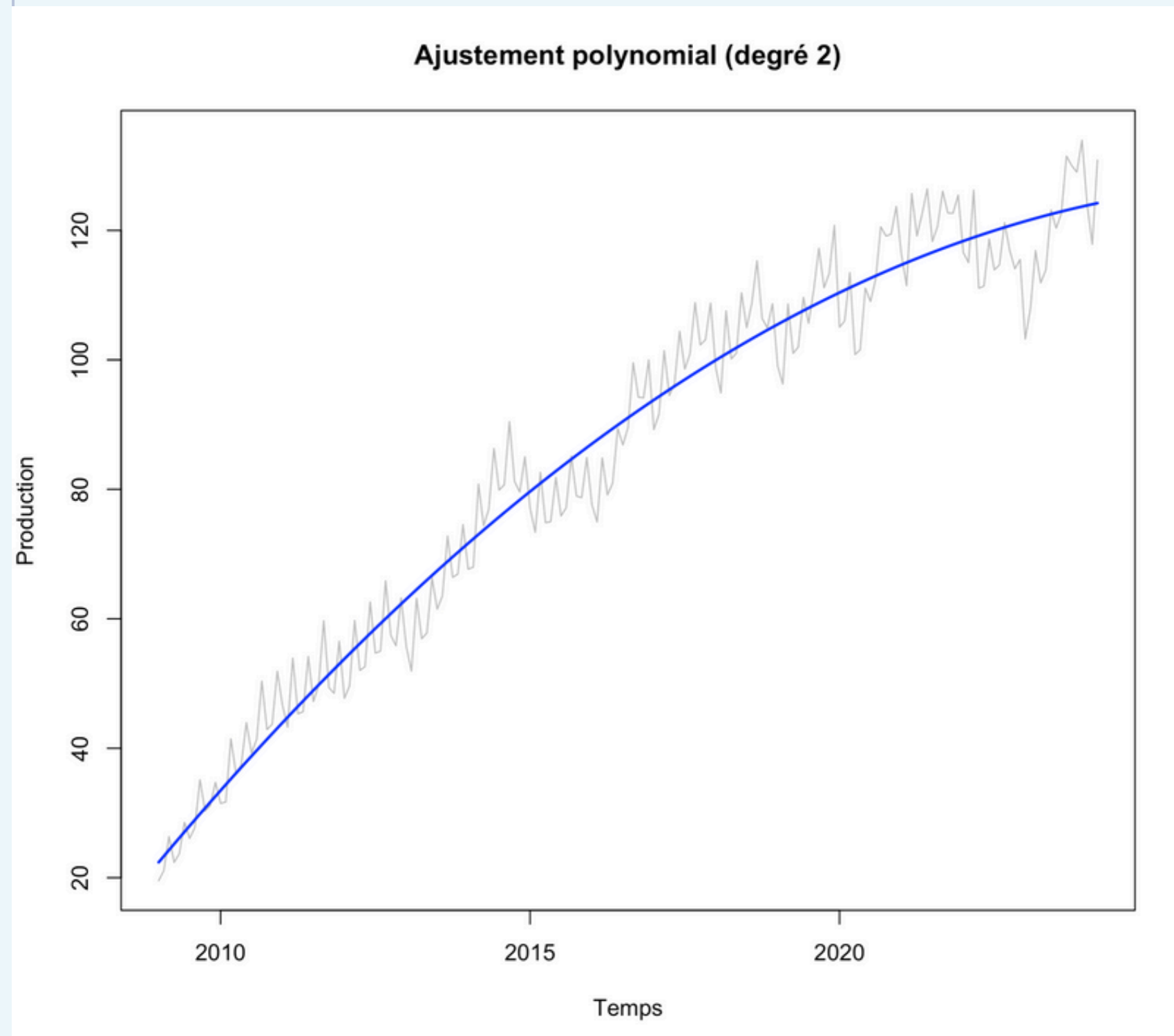
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.345e+04	2.694e+02	-49.95	<2e-16 ***
temps	6.713e+00	1.336e-01	50.26	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.955 on 181 degrees of freedom
Multiple R-squared: 0.9331, Adjusted R-squared: 0.9328
F-statistic: 2526 on 1 and 181 DF, p-value: < 2.2e-16

Ajustement de la tendance par polynome de deg 2



```
> summary(modele_poly)
```

Call:

```
lm(formula = production_electronique_ts ~ poly(temps, 2, raw = TRUE))
```

Residuals:

Min	1Q	Median	3Q	Max
-18.2153	-4.0683	-0.3013	4.1396	13.3882

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-1.270e+06	1.018e+05	-12.47	<2e-16	***
poly(temps, 2, raw = TRUE)1	1.253e+03	1.010e+02	12.41	<2e-16	***
poly(temps, 2, raw = TRUE)2	-3.090e-01	2.504e-02	-12.34	<2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.871 on 180 degrees of freedom

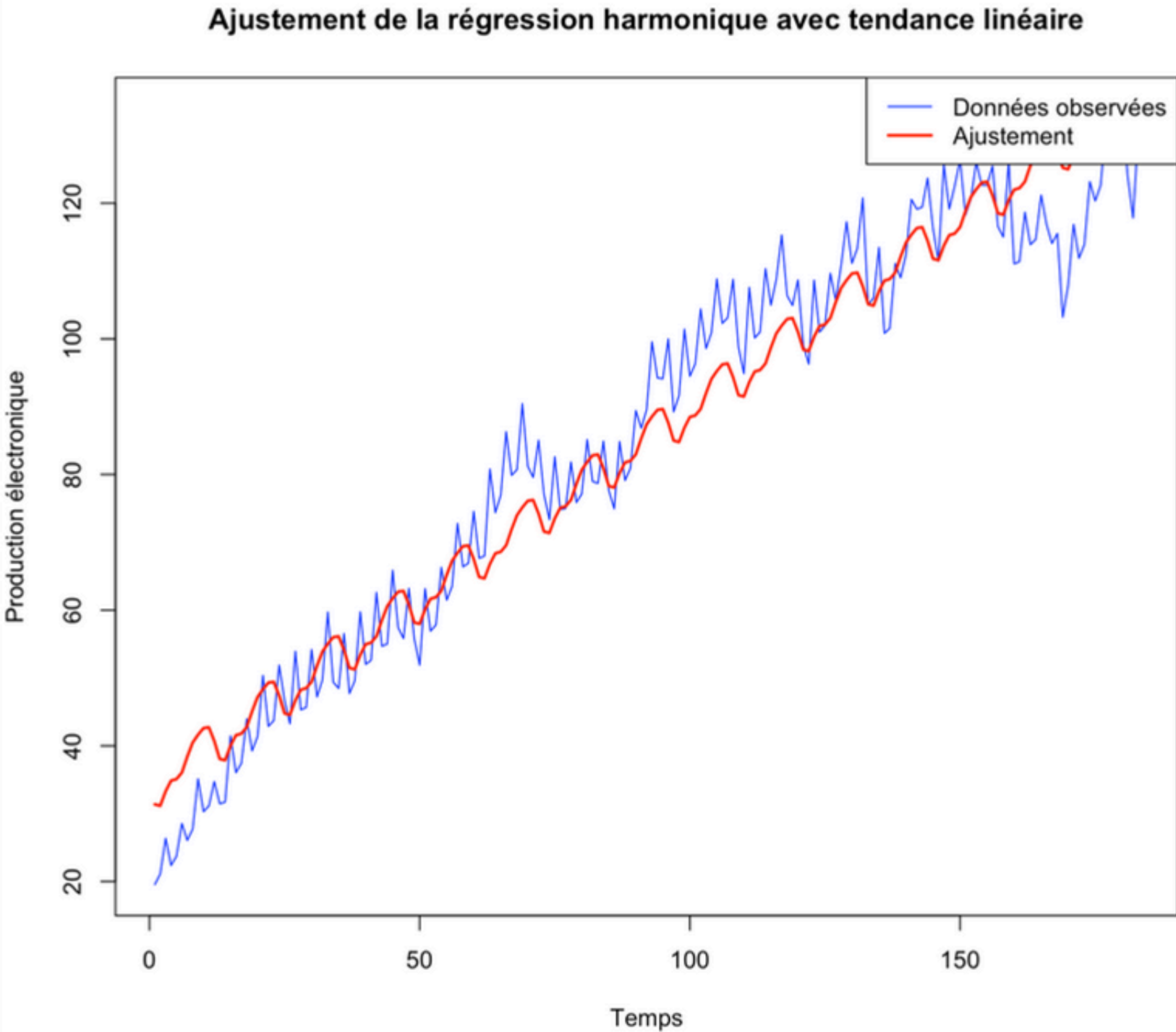
Multiple R-squared: 0.9638, Adjusted R-squared: 0.9634

F-statistic: 2395 on 2 and 180 DF, p-value: < 2.2e-16

➔ Interpretation :

Le modèle polynomial (degré 2) est le meilleur ajustement car Il présente le coefficient de détermination R^2 le plus élevé, avec une valeur de 0.9638

Ajustement de la regression harmonique avec tendance linéaire



	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	33.73408	1.14894	29.361	< 2e-16	***
t	0.55857	0.01083	51.572	< 2e-16	***
Mc1	0.17958	0.81138	0.221	0.825096	
Mc2	-0.38169	0.80878	-0.472	0.637558	
Mc3	0.47971	0.81090	0.592	0.554900	
Ms1	-2.98713	0.80684	-3.702	0.000286	***
Ms2	-1.03534	0.80884	-1.280	0.202230	
Ms3	-0.49500	0.80663	-0.614	0.540233	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Residual standard error: 7.734 on 175 degrees of freedom
Multiple R-squared: 0.9389, Adjusted R-squared: 0.9364
F-statistic: 384.1 on 7 and 175 DF, p-value: < 2.2e-16

Le modèle montre que la production électronique suit une tendance haussière régulière avec une saisonnalité annuelle significative, bien captée par les termes sinusoidaux.

stationnarité(ADF)

```
> adf.test(production_electronique_ts) # Test de Dickey-Fuller (ADF)
```

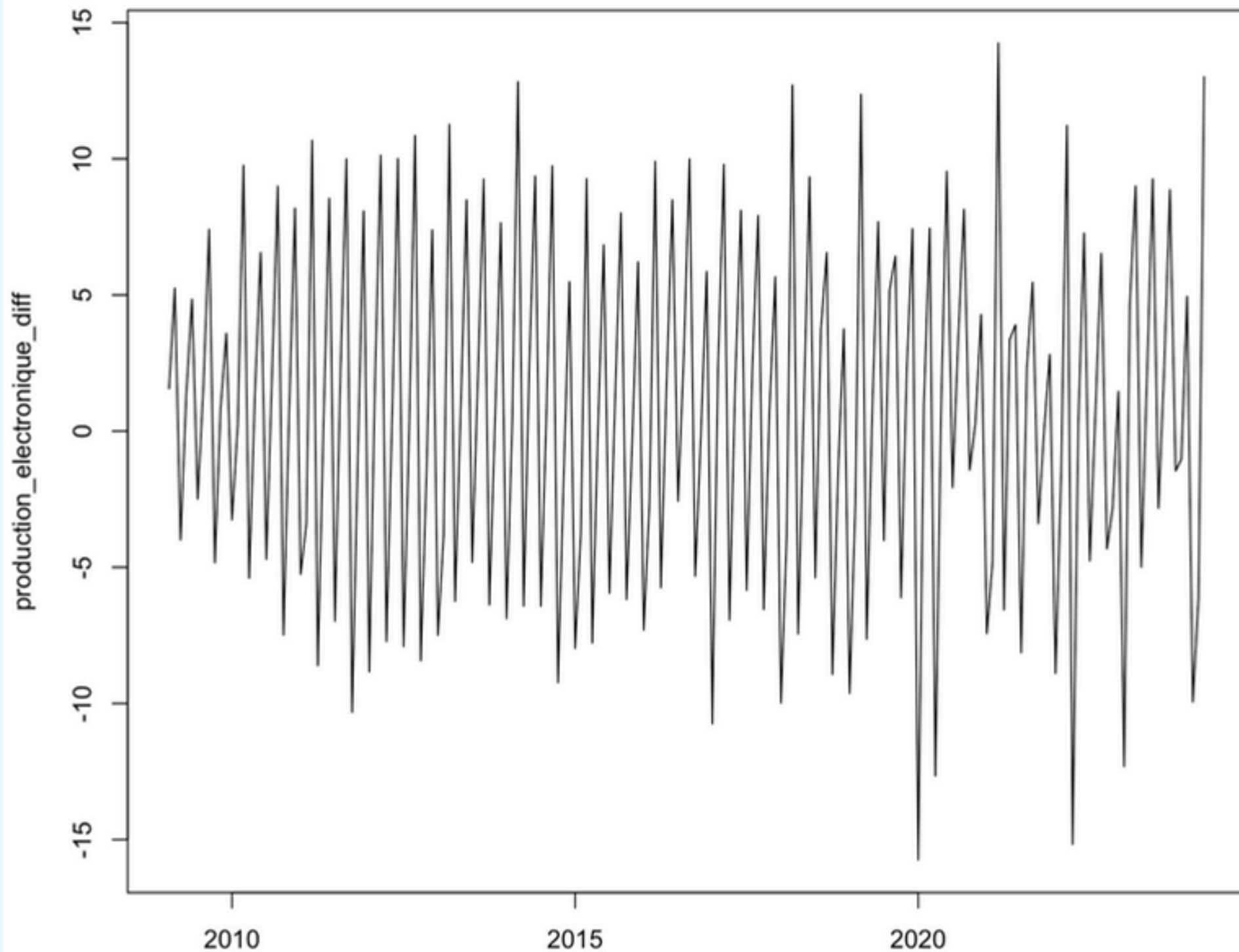
Augmented Dickey-Fuller Test

```
data: production_electronique_ts  
Dickey-Fuller = -2.3466, Lag order = 5, p-value = 0.4313  
alternative hypothesis: stationary
```

- Hypothèse nulle (H_0) : la série n'est pas stationnaire (elle a une racine unitaire).
- Hypothèse alternative (H_1) : la série est stationnaire.
- $p\text{-value} = 0.4313 > 0.05$, on accepte $H_0 \rightarrow$ La série n'est pas stationnaire.

différenciation

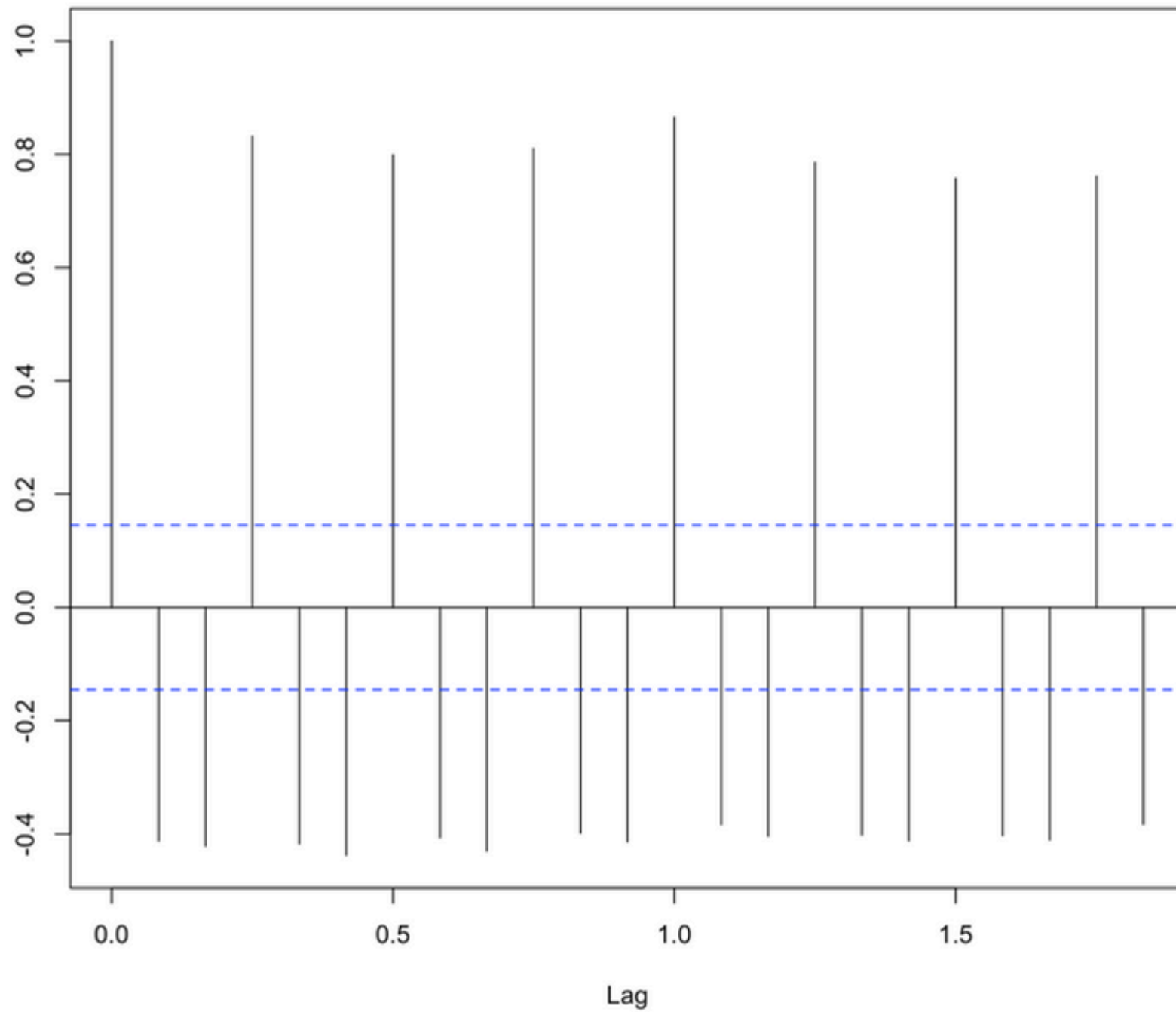
Série différenciée d'ordre 1



- Le graphique montre que la série différenciée d'ordre 1 oscille autour de zéro, ce qui indique que la tendance a été éliminée.
- Cela suggère que la série est désormais plus stationnaire, une condition nécessaire pour la modélisation ARIMA.

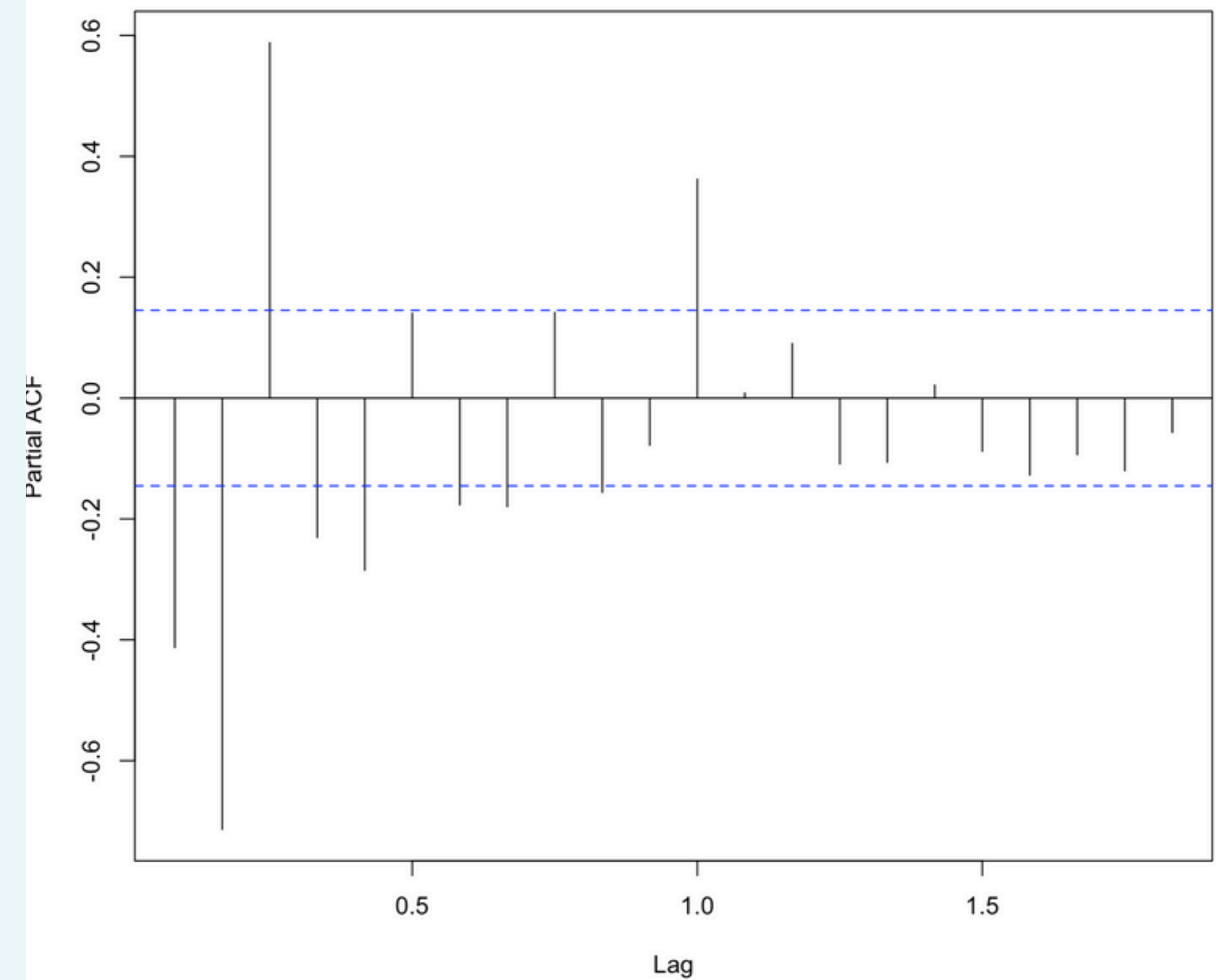
différenciation

ACF de la série différenciée



AR (autorégressif)

PACF de la série différenciée



“Grâce à ces deux courbes, on peut choisir les bons paramètres pour notre modèle ARIMA : le p (ordre AR), le d (ordre de différenciation) et le q (ordre MA).”

Model ARIMA

```
series: production_electronique_ts  
ARIMA(2,1,2)(2,1,1)[12]
```

Coefficients:

	ar1	ar2	ma1	ma2	sar1	sar2	sma1
	-0.9587	-0.8915	0.8878	0.7158	-0.6613	-0.4237	-0.3008
s.e.	0.1407	0.0934	0.2024	0.1200	0.1340	0.1144	0.1389

```
sigma^2 = 4.496: log likelihood = -371.84  
AIC=759.69 AICc=760.58 BIC=784.77
```

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-0.1586519	2.001074	1.452874	-0.2555057	1.759116	0.1771049	0.01130244

```
> stats::Box.test(residus, lag = 20, type = "Ljung-Box")
```

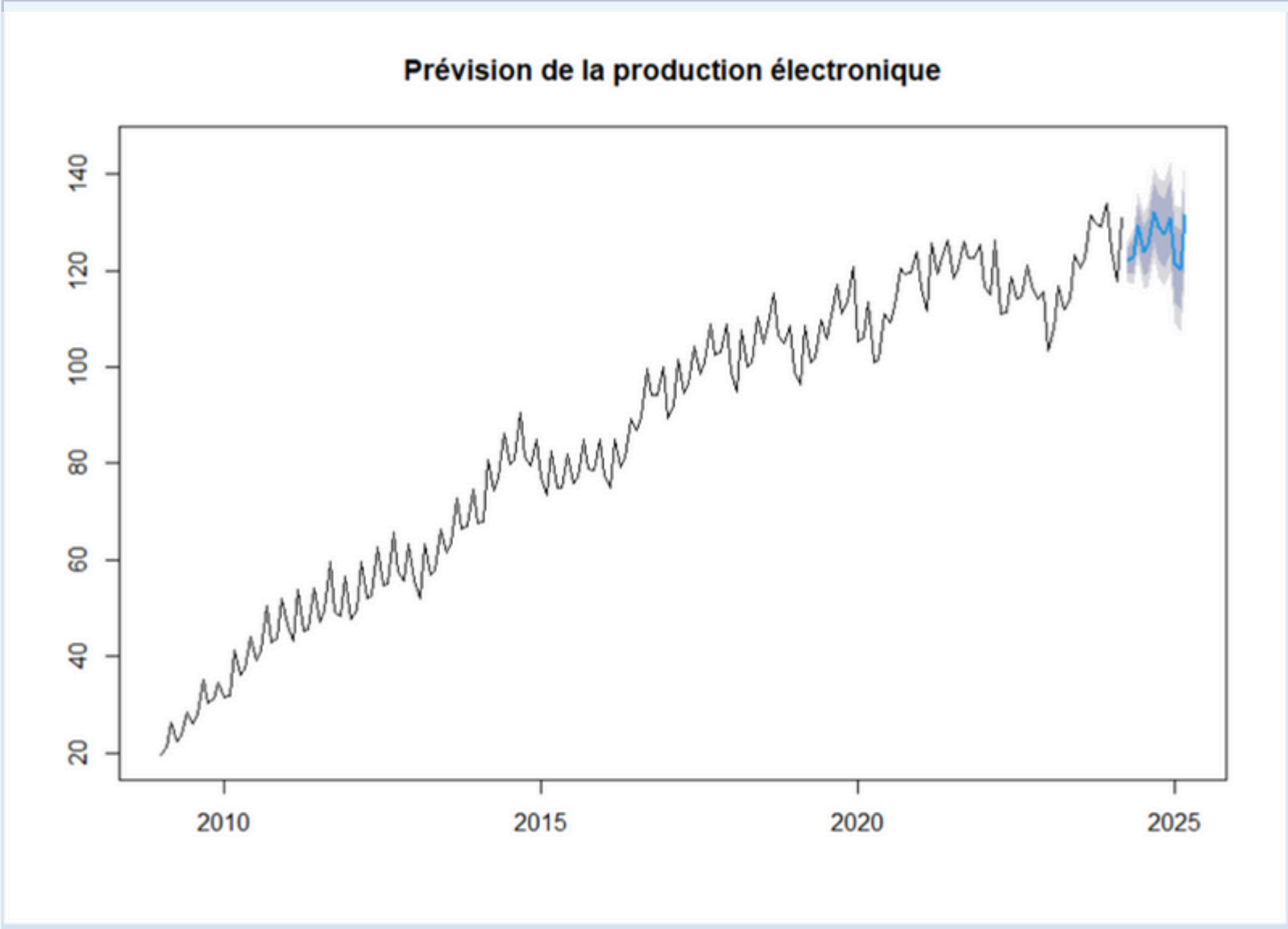
Box-Ljung test

```
data: residus  
X-squared = 26.725, df = 20, p-value = 0.1432
```

p-value = 0.1432 → plus grande que 0.05

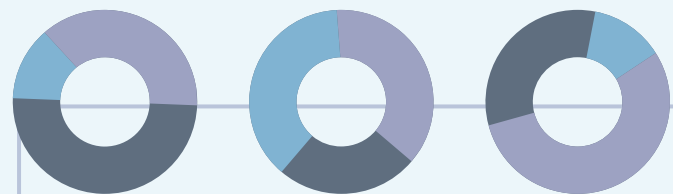
- ➡ Donc, on accepte H_0 « Les résidus ne sont pas autocorrélés => ressemblent à du bruit blanc)»
- ➡ le model capture la structure de la série. => Les résidus sont aléatoires
- ➡ Il ya rien a modelisé pour la composante aléatoire

Prévision



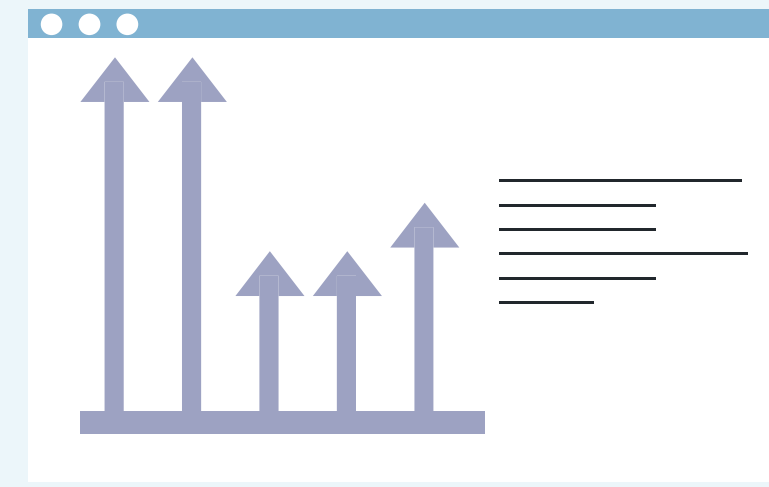
	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Apr 2024		121.8963	119.1790	124.6136	117.7406	126.0520
May 2024		122.8675	119.1584	126.5766	117.1949	128.5401
Jun 2024		129.3273	124.9985	133.6561	122.7069	135.9477
Jul 2024		123.9229	118.8294	129.0165	116.1330	131.7128
Aug 2024		125.2975	119.6184	130.9767	116.6121	133.9830
Sep 2024		131.9876	125.8702	138.1050	122.6318	141.3433
Oct 2024		128.9557	122.2867	135.6246	118.7563	139.1550
Nov 2024		127.5936	120.4709	134.7164	116.7003	138.4869
Dec 2024		130.9247	123.4371	138.4123	119.4734	142.3760
Jan 2025		121.3873	113.4509	129.3238	109.2497	133.5250
Feb 2025		120.1916	111.8725	128.5106	107.4687	132.9145
Mar 2025		131.5114	122.8700	140.1528	118.2956	144.7273

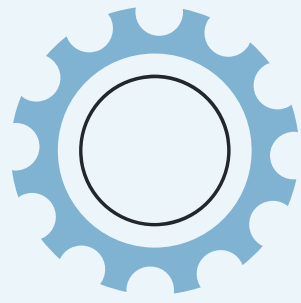
- La production électronique est prévue à la hausse avec fluctuations modérées, autour de 120 à 132 unités.
- Exemple d'interprétation concrète :
==>En septembre 2024, la production est estimée à 131,99, avec 95 % de chances qu'elle soit entre 122,63 et 141,34.



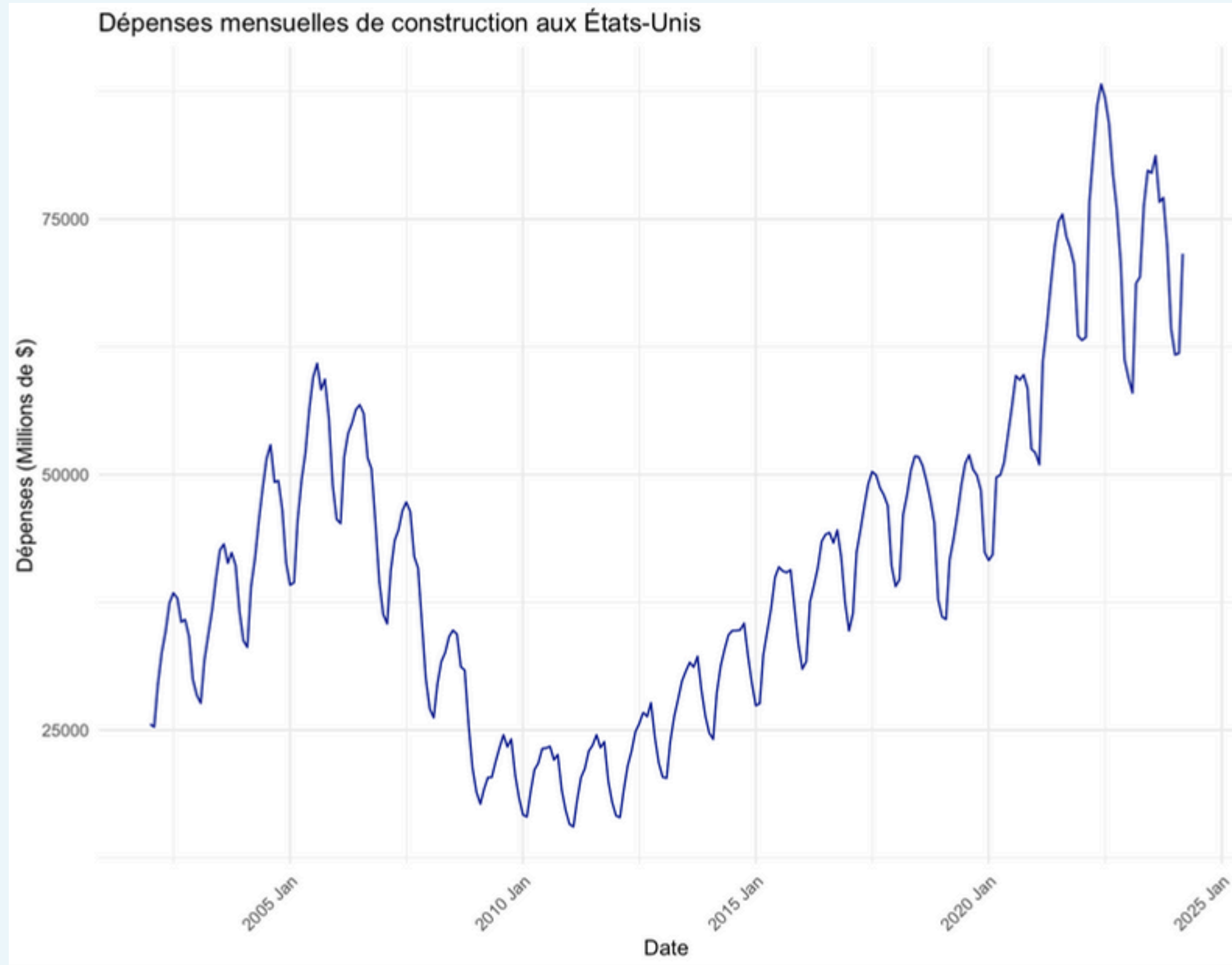
Partie 2

Data Set 2: dépenses de construction





Visualisation de la serie



Ce graphique représente les dépenses mensuelles de construction aux États-Unis sur une période de 25 ans, de janvier 2000 à janvier 2025.

◆ **Tendance générale** : On observe une augmentation progressive des dépenses jusqu'en 2007, suivie d'une baisse notable autour de 2008. Une reprise s'amorce ensuite, avec une croissance régulière à partir de 2012, qui devient plus marquée vers 2020, atteignant un pic autour de 2022

◆ **Saisonnalité et fluctuations** :

Des variations périodiques, ce qui pourrait indiquer une saisonnalité dans les investissements de construction .

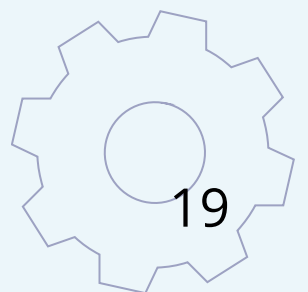
Informations Sur DATA

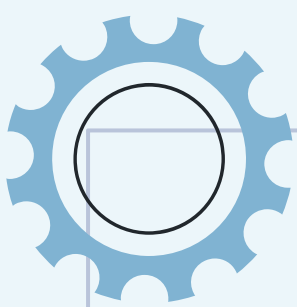
Corr(X,Y)

```
> head(construction_ts)
      Jan  Feb  Mar  Apr  May  Jun
2002 25572 25298 29407 32578 34621 37440
> tail(construction_ts)
      Jan  Feb  Mar Apr May Jun Jul Aug Sep  Oct  Nov  Dec
2023
2024 61716 61866 71611
> str(construction_ts)
Time-Series [1:267] from 2002 to 2024: 25572 25298 29407 32578 34621 ...
> summary(construction_ts)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 15541  28760  40416  41746  50923  88196
> class(construction_ts)
[1] "ts"
```

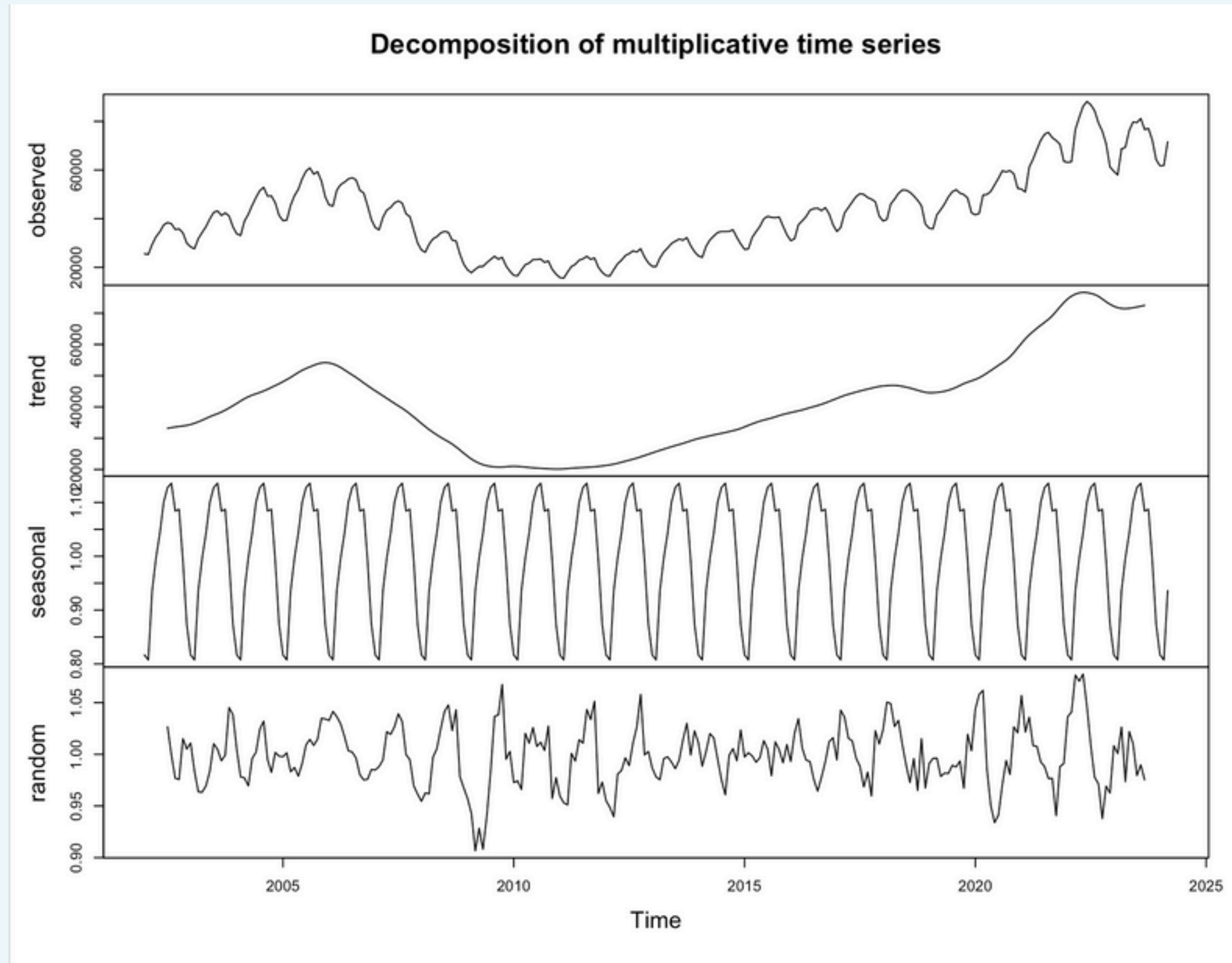
```
> cat("Moyenne:", mean_value, "\n")
Moyenne: 41745.97
> cat("Variance:", var_value, "\n")
Variance: 277439667
> cat("Écart-type:", std_value, "\n")
Écart-type: 16656.52
```

La série construction_ts a une moyenne élevée (~41 746) et une forte dispersion (écart-type ~16 657), ce qui indique une grande variabilité autour de la tendance générale.





Décomposition



1 Observé :

On observe des fluctuations régulières avec une tendance générale à la hausse.

2 Tendance :

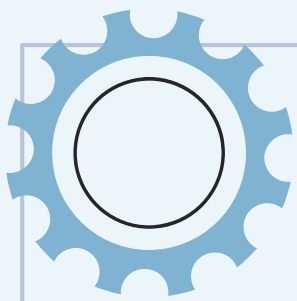
On remarque une augmentation entre 2000 et 2015, suivie d'un léger déclin, puis d'une nouvelle hausse.

3 Saisonnalité :

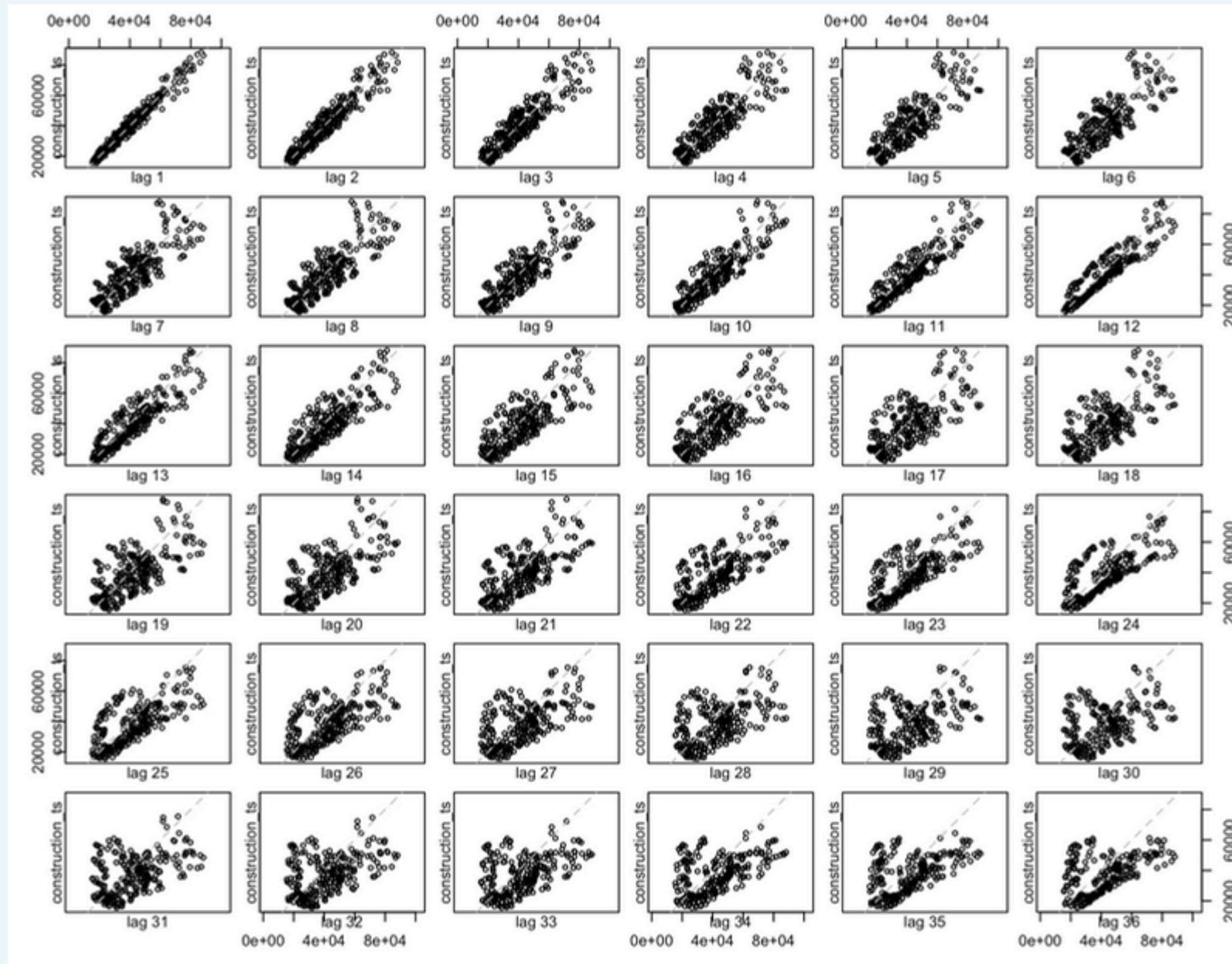
on distingue des tendances périodiques qui se répètent à intervalles fixes.

4 Aléatoire :

Ce dernier graphique représente montre des fluctuations de la serie temporelle

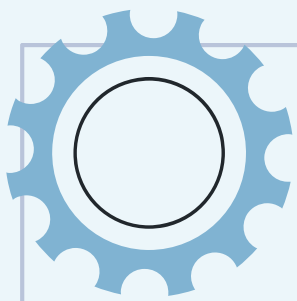


Autocorrélation

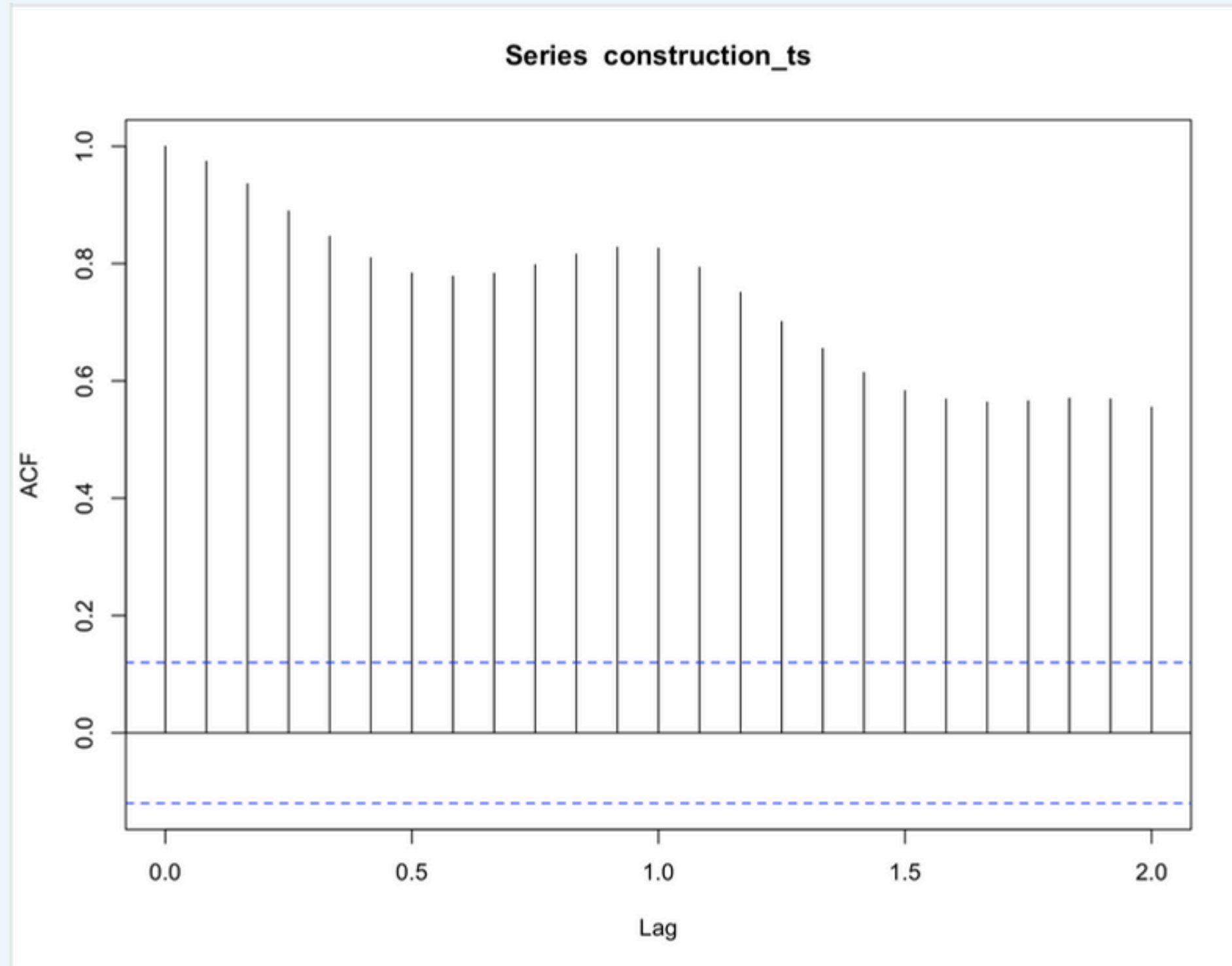


Interpretation:

- ◆ Corrélation temporelle :
 - les points forment un schéma linéaire, cela peut indiquer une forte corrélation. pour les lag 1 ,2,3



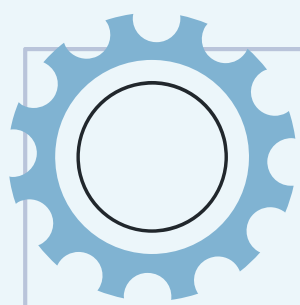
Autocorrélation(acf)



Interpretation:

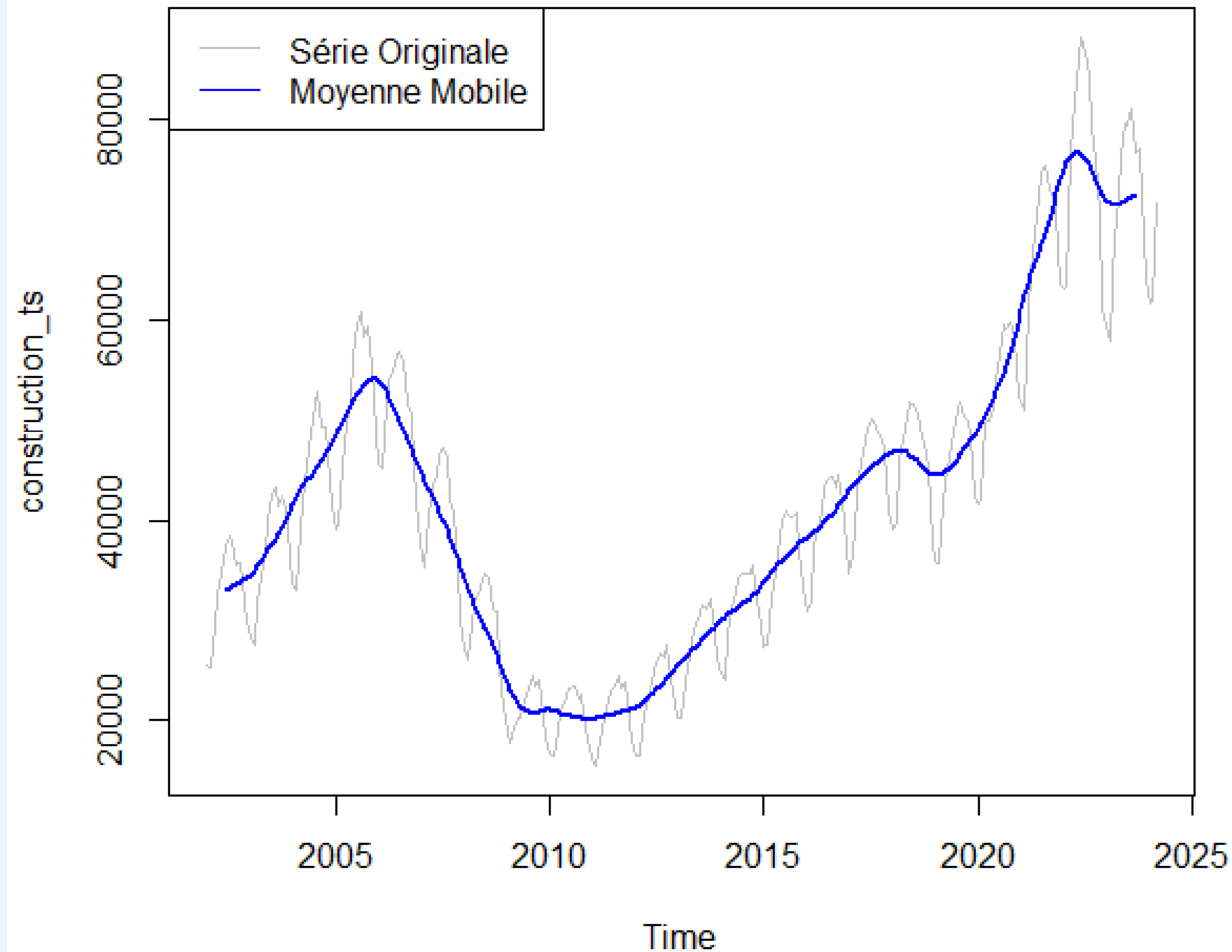
l'autocorrélation diminue lentement avec les lags, cela suggère une tendance dans les données.

- ✓ Si des pics apparaissent à des intervalles réguliers, cela peut indiquer une saisonnalité dans la série.
- ✓ Si l'autocorrélation chute brusquement après quelques lags, cela peut suggérer un processus stationnaire.



Autocorrélation(acf)

Ajustement par moyenne mobile



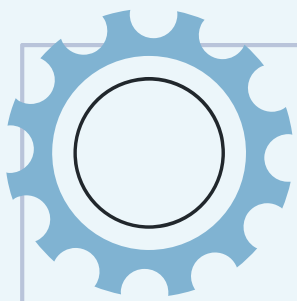
Interprétation :

◆ Lissage des fluctuations :

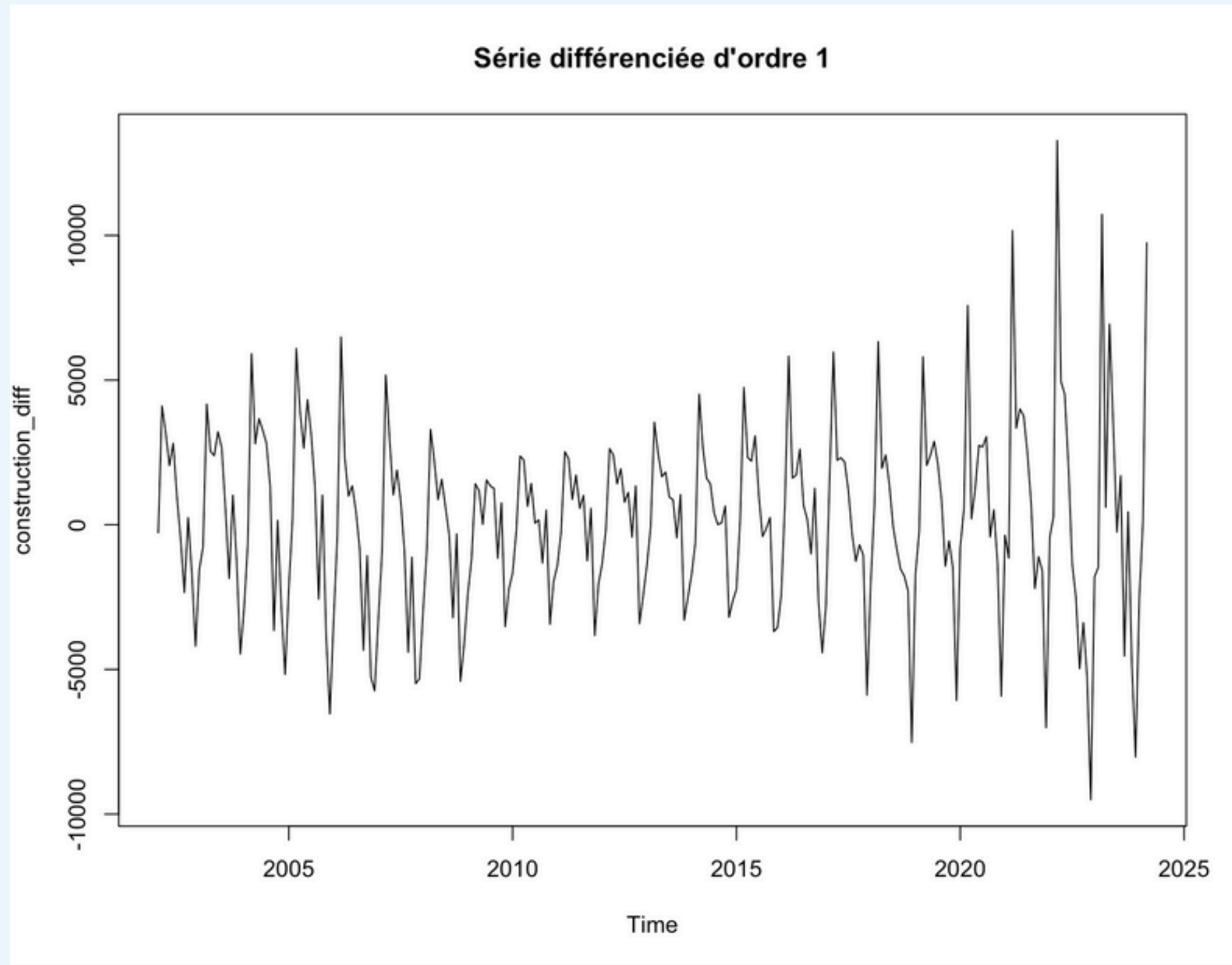
La courbe bleue (moyenne mobile) lisse la série originale (courbe grise), ce qui permet d'atténuer les variations saisonnières et les bruits à court terme.

◆ Tendance générale visible :

On observe clairement une tendance générale à la hausse entre 2010 et 2022, précédée par une chute entre 2006 et 2009. La moyenne mobile met en évidence ces évolutions longues.



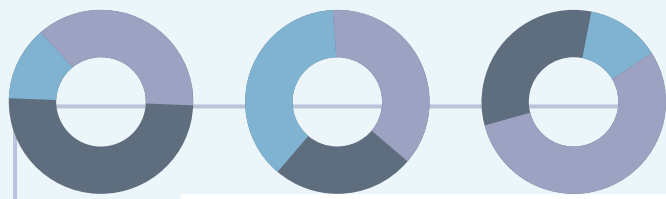
Différenciation



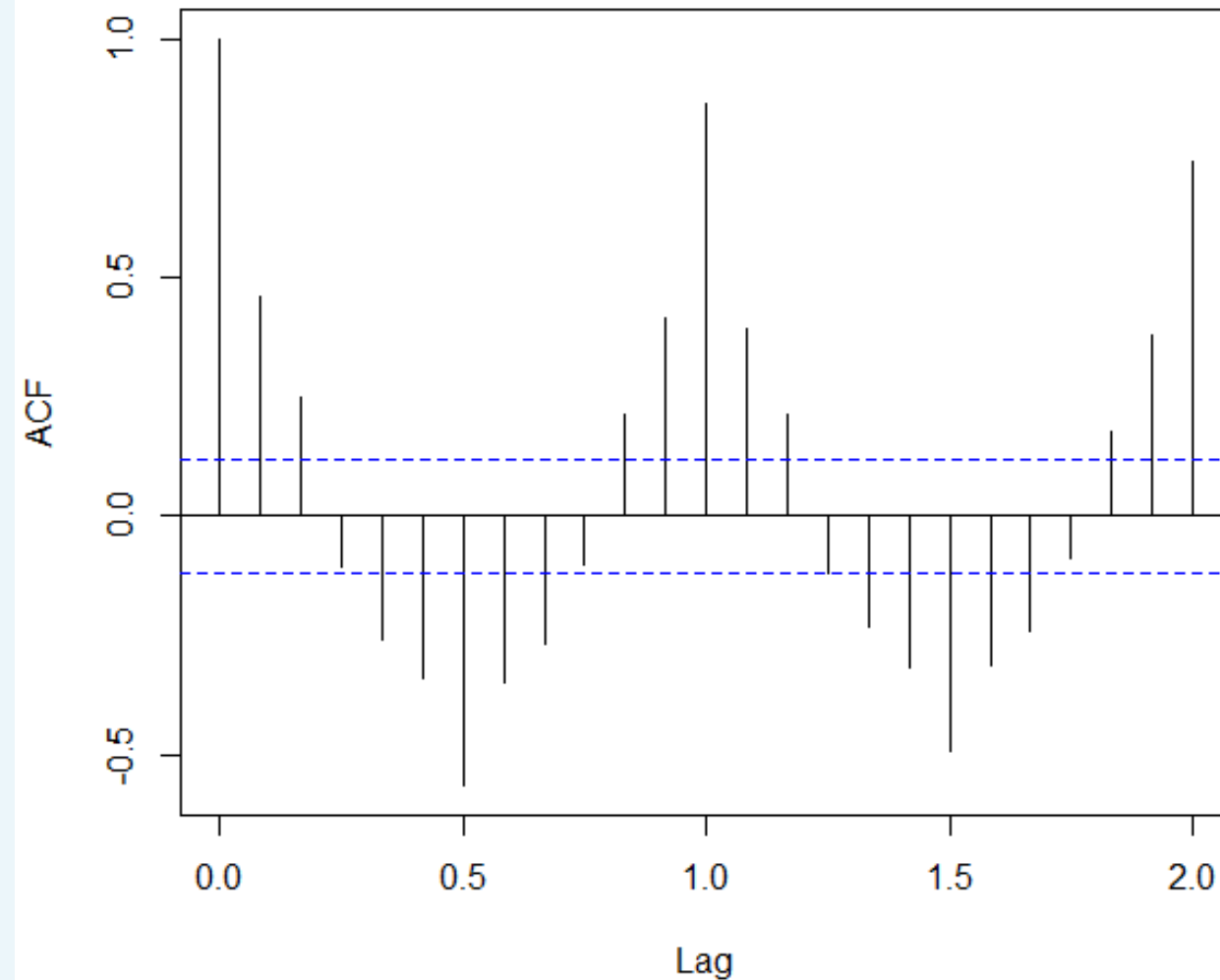
Interprétation :

La série temporelle différenciée d'ordre 1 permet d'éliminer la tendance pour mieux analyser les fluctuations de la variable "construction".

Les variations observées, avec des pics et des creux marqués, indiquent une volatilité significative, utile pour la modélisation avec ARIMA afin d'améliorer la prévision.

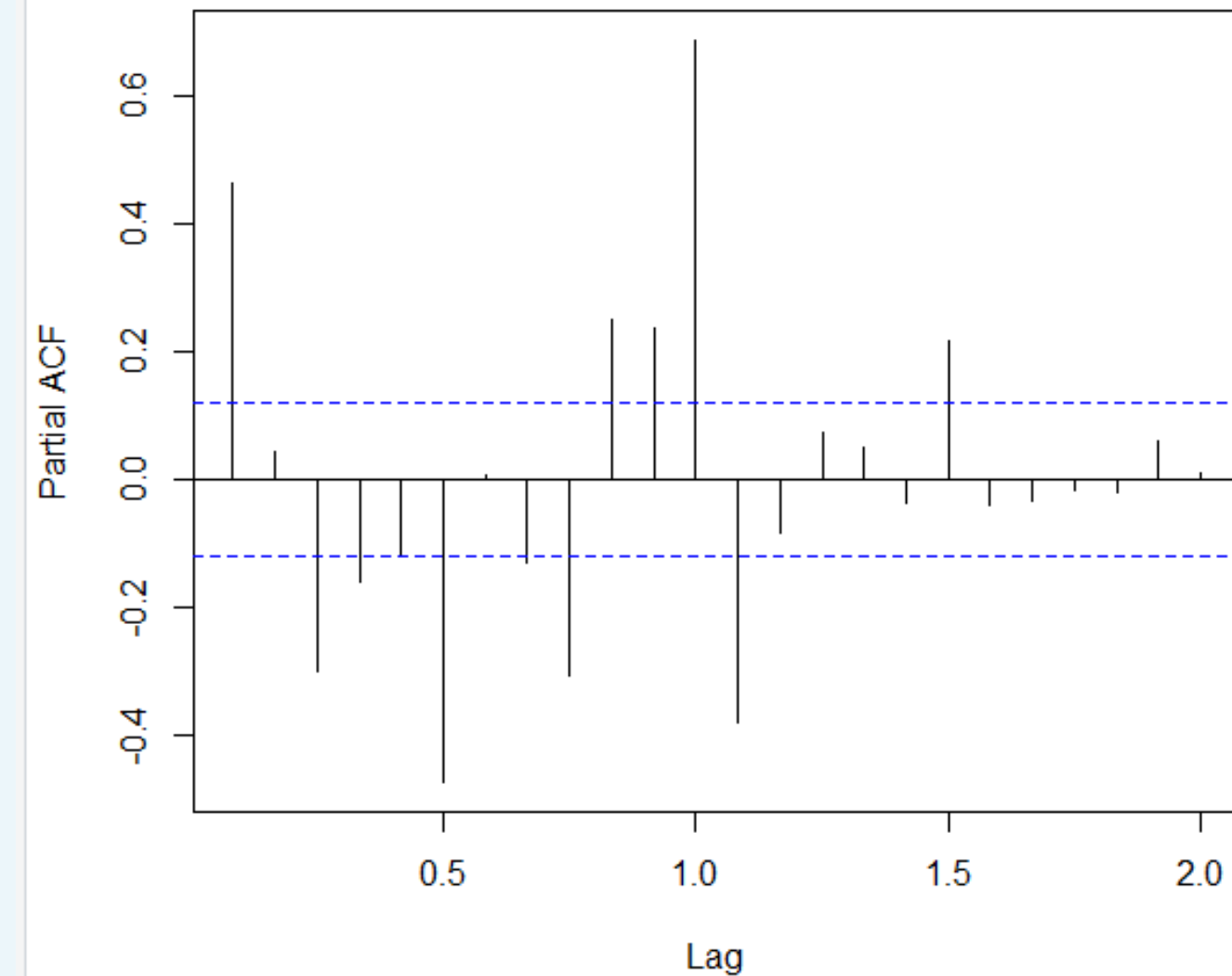


ACF de la série différenciée



- La première ou deuxième barre est significative, ensuite les valeurs chutent rapidement vers zéro.
- 🖱️ Interprétation : cela suggère que la composante MA (Moving Average) est probablement de faible ordre, par exemple $q = 1$ ou $q = 2$.

PACF de la série différenciée



- Ici aussi, une ou deux premières barres significatives, ensuite les valeurs deviennent faibles ou non significatives.
- 🖱️ Interprétation : cela indique que la composante AR (Auto-Regressive) est probablement de faible ordre également, donc $p = 1$ ou $p = 2$.



```
> summary(modele_auto)
Series: construction_ts
ARIMA(1,1,2)(0,1,0)[12]

Coefficients:
      ar1      ma1      ma2
    0.4204  0.0127  0.2431
s.e.  0.1809  0.1886  0.0910

sigma^2 = 1052071:  log likelihood = -2120.12
AIC=4248.23  AICC=4248.4  BIC=4262.38

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -0.7137775 994.4974 700.3269 0.07286256 1.656752 0.1203619 -0.005600652
```

```
> Box.test(residus, lag = 20, type = "Ljung-Box")
```

Box-Ljung test

```
data:  residus
X-squared = 11.783, df = 20, p-value = 0.9233
```

◆ Qualité du modèle :

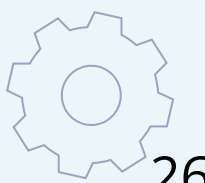
AIC = 4248.23, BIC = 4262.38 : Des valeurs plus faibles indiquent un meilleur ajustement du modèle.

◆ Erreurs du modèle :

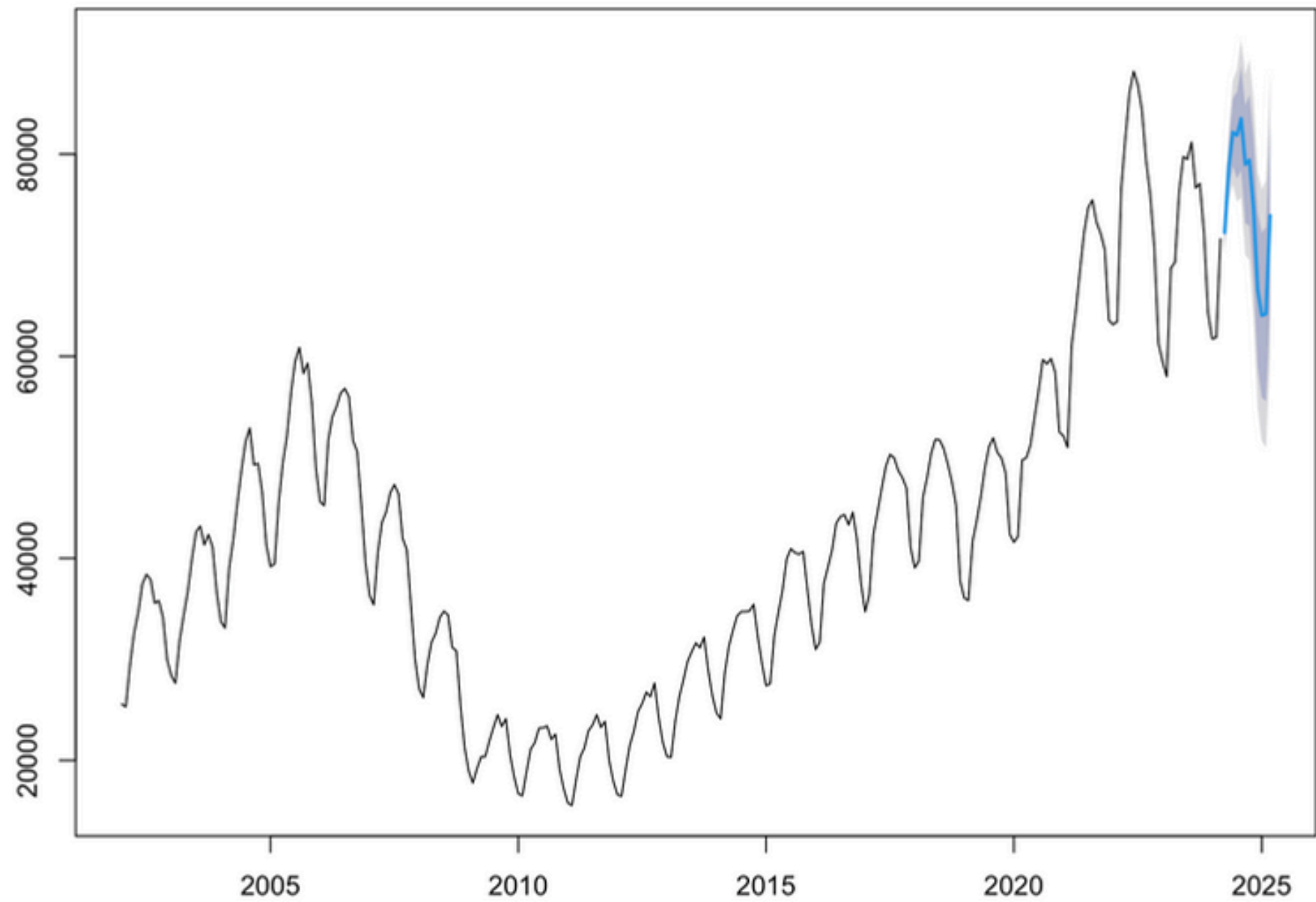
- RMSE = 994.49 et MAE = 700.32 : Indiquent l'écart moyen entre la prévision et la réalité.
- MAPE = 1.65% : Erreur moyenne en pourcentage, ce qui montre un bon niveau de précision.

◆ Interprétation :

- p-value > 0.05 => n'y a pas d'autocorrélation significative dans les résidus.
- Autrement dit, le modèle ARIMA utilisé capture bien la structure des données et les erreurs sont aléatoires. 🚀📊



Prévision des dépenses de construction

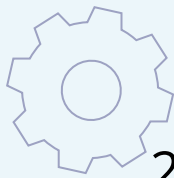


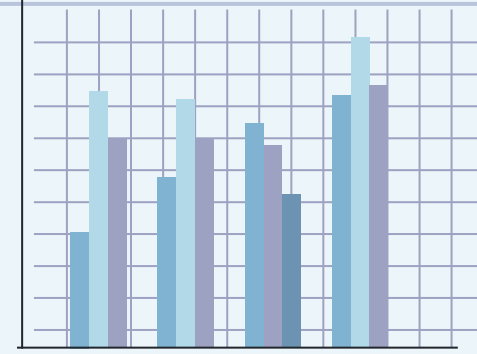
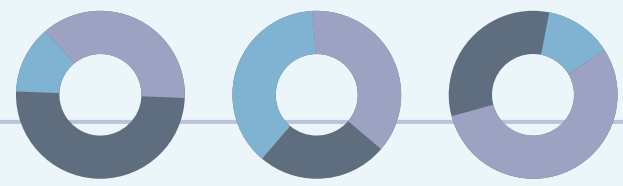
```
> prevision
```

	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Apr 2024		72228.71	70914.22	73543.20	70218.37	74239.06
May 2024		78823.95	76526.93	81120.98	75310.95	82336.95
Jun 2024		82175.13	78822.19	85528.07	77047.25	87303.01
Jul 2024		81858.09	77567.32	86148.87	75295.92	88420.27
Aug 2024		83529.86	78419.31	88640.41	75713.95	91345.77
Sep 2024		78979.25	73143.42	84815.08	70054.12	87904.38
Oct 2024		79420.79	72932.75	85908.83	69498.18	89343.39
Nov 2024		74600.91	67517.54	81684.29	63767.83	85433.99
Dec 2024		66569.13	58935.57	74202.68	54894.62	78243.64
Jan 2025		64035.80	55888.68	72182.91	51575.86	76495.73
Feb 2025		64185.66	55555.31	72816.01	50986.68	77384.64
Mar 2025		73930.60	64842.60	83018.59	60031.71	87829.49

```
> |
```

La prévision montre une augmentation globale des dépenses de construction, mais avec une incertitude croissante en 2025. La zone bleue représente les intervalles de confiance, indiquant la fiabilité des prévisions. Cette analyse peut aider à anticiper les investissements et planifier les budgets





Merci de votre attention

