



Modélisation et Prévision du Taux d'Incidence de la Varicelle en France

INTRODUCTION

LA MODÉLISATION ET LA PRÉVISION DES SÉRIES TEMPORELLES JOUENT UN RÔLE CLÉ DANS L'ANALYSE DES DONNÉES ÉPIDÉMIOLOGIQUES. EN FRANCE, LE RÉSEAU SENTINELLES, COORDONNÉ PAR L'ÉQUIPE SUMO (SURVEILLANCE ET MODÉLISATION DES MALADIES TRANSMISSIBLES) DE L'INSTITUT PIERRE LOUIS D'ÉPIDÉMIOLOGIE ET DE SANTÉ PUBLIQUE, FOURNIT DEPUIS 1984 DES DONNÉES CRUCIALES POUR LA VEILLE SANITAIRE ET LA RECHERCHE. PARMI CES DONNÉES, CELLES CONCERNANT LA VARICELLE, COLLECTÉES DEPUIS 1991, OFFRENT UNE OPPORTUNITÉ UNIQUE D'ÉTUDIER LES FLUCTUATIONS TEMPORELLES ET SAISONNIÈRES DE CETTE MALADIE.

DANS CE PROJET, NOUS NOUS FOCALISONS SUR L'ANALYSE DES TAUX D'INCIDENCE HEBDOMADAIRES DE LA VARICELLE EN FRANCE. L'OBJECTIF PRINCIPAL EST D'EXAMINER LES TENDANCES ET LES SCHÉMAS SAISONNIERS DE CES DONNÉES, PUIS D'EFFECTUER DES PRÉVISIONS POUR LES PÉRIODES FUTURES. UNE ÉVALUATION DES MODÈLES DE PRÉVISION SERA ÉGALEMENT RÉALISÉE EN COMPARANT LEURS PERFORMANCES À L'AIDE DE MÉTRIQUES COMME L'ERREUR QUADRATIQUE MOYENNE (EQM).

CE RAPPORT EST STRUCTURÉ AUTOUR DE PLUSIEURS ÉTAPES : TOUT D'ABORD, UNE PRÉPARATION ET UNE EXPLORATION DES DONNÉES SERONT EFFECTUÉES, SUIVIES PAR UNE DÉCOMPOSITION DES SÉRIES TEMPORELLES POUR IDENTIFIER LEURS COMPOSANTES PRINCIPALES. ENSUITE, DES MODÈLES DE PRÉVISION SERONT APPLIQUÉS, NOTAMMENT LES MODÈLES ARMA, ETS ET UNE RÉGRESSION LINÉAIRE, POUR ESTIMER LES TENDANCES FUTURES. ENFIN, LES PERFORMANCES DE CES MODÈLES SERONT COMPARÉES AFIN DE DÉTERMINER LA MÉTHODE LA PLUS ADAPTÉE À LA PRÉVISION DU TAUX D'INCIDENCE DE LA VARICELLE.

EN RÉSUMÉ, CETTE ÉTUDE ILLUSTRE LES TECHNIQUES DE MODÉLISATION ET DE PRÉVISION APPLIQUÉES À UNE PROBLÉMATIQUE RÉELLE, TOUT EN OFFRANT UNE MEILLEURE COMPRÉHENSION DES DYNAMIQUES ÉPIDÉMIOLOGIQUES DE LA VARICELLE EN FRANCE

PRÉSENTATION DES DONNÉES

LES DONNÉES UTILISÉES DANS CETTE ÉTUDE PROVIENNENT DU RÉSEAU SENTINELLES ET COUVRENT LES SEMAINES DE 1991 À NOS JOURS. ELLES CONTIENNENT DES INFORMATIONS SUR L'INCIDENCE HEBDOMADAIRE DE LA VARICELLE EN FRANCE. VOICI LES PRINCIPALES CARACTÉRISTIQUES DE CES DONNÉES :

- WEEK : IDENTIFIANT DE LA SEMAINE (FORMAT AAAASW), PERMETTANT DE SUIVRE L'ÉVOLUTION TEMPORELLE DES CAS.
- INDICATOR : INDICATEUR SPÉCIFIQUE À LA VARICELLE (VALEUR FIXE : 7).
- INC : NOMBRE TOTAL ESTIMÉ DE CAS HEBDOMADAIRES.
- INC_LOW ET INC_UP : LIMITES INFÉRIEURE ET SUPÉRIEURE DU NOMBRE DE CAS ESTIMÉS, PERMETTANT D'ÉVALUER L'INCERTITUDE DES ESTIMATIONS.
- INC100 : TAUX D'INCIDENCE POUR 100 000 HABITANTS.
- INC100_LOW ET INC100_UP : LIMITES INFÉRIEURE ET SUPÉRIEURE DU TAUX D'INCIDENCE POUR 100 000 HABITANTS.
- GEO_INSEE ET GEO_NAME : INFORMATIONS GÉOGRAPHIQUES (ICI, TOUTES LES DONNÉES CONCERNENT LA FRANCE ENTIÈRE).

CES DONNÉES SONT COMPLÈTES, SANS VALEURS MANQUANTES, ET PERMETTENT UNE ANALYSE APPROFONDIE DES TENDANCES ÉPIDÉMIOLOGIQUES DE LA VARICELLE SUR PLUSIEURS DÉCENNIES.

	week	indicator	inc	inc_low	inc_up	inc100	inc100_low	inc100_up	geo_insee	geo_name
1	202446	7	2474	760	4188	4	1	7	FR	France
2	202445	7	2751	1232	4270	4	2	6	FR	France
3	202444	7	2135	676	3594	3	1	5	FR	France
4	202443	7	2124	641	3607	3	1	5	FR	France
5	202442	7	2621	1246	3996	4	2	6	FR	France
6	202441	7	2035	381	3689	3	1	5	FR	France
7	202440	7	2125	725	3525	3	1	5	FR	France
8	202439	7	2898	1333	4463	4	2	6	FR	France
9	202438	7	751	0	1513	1	0	2	FR	France
10	202437	7	916	28	1804	1	0	2	FR	France
11	202436	7	2235	870	3600	3	1	5	FR	France
12	202435	7	1620	285	2955	2	0	4	FR	France
13	202434	7	2560	622	4498	4	1	7	FR	France
14	202433	7	1971	536	3406	3	1	5	FR	France
15	202432	7	4399	1944	6854	7	3	11	FR	France

PRÉSENTATION DES DONNÉES

```
{r}
# Charger les bibliothèques nécessaires
#install.packages(c("forecast", "ggplot2", "tseries", "readxl", "dplyr"))
library(forecast)
library(ggplot2)
library(tseries)
library(readxl)
library(dplyr)
```

POUR MENER À BIEN CE PROJET, NOUS AVONS UTILISÉ PLUSIEURS BIBLIOTHÈQUES EN R, INDISPENSABLES POUR L'ANALYSE ET LA MODÉLISATION DES SÉRIES TEMPORELLES. PARMI ELLES, FORECAST A ÉTÉ EMPLOYÉE POUR AJUSTER ET PRÉDIRE LES MODÈLES ARMA ET ETS, GGPLOT2 POUR LA VISUALISATION DES DONNÉES, ET TSERIES POUR EFFECTUER DES TESTS STATISTIQUES. DE PLUS, READXL ET DPLYR NOUS ONT PERMIS DE MANIPULER ET DE NETTOYER LES DONNÉES EFFICACEMENT. CES OUTILS NOUS ONT OFFERT UNE BASE SOLIDE POUR EXPLORER LES TENDANCES, LA SAISONNALITÉ ET LES PRÉVISIONS LIÉES À L'INCIDENCE DE LA VARICELLE.

```
{r charger les données}

varicelle_data <- read.csv("F:/serie/varicelle.csv", sep = ";", header = TRUE)
```

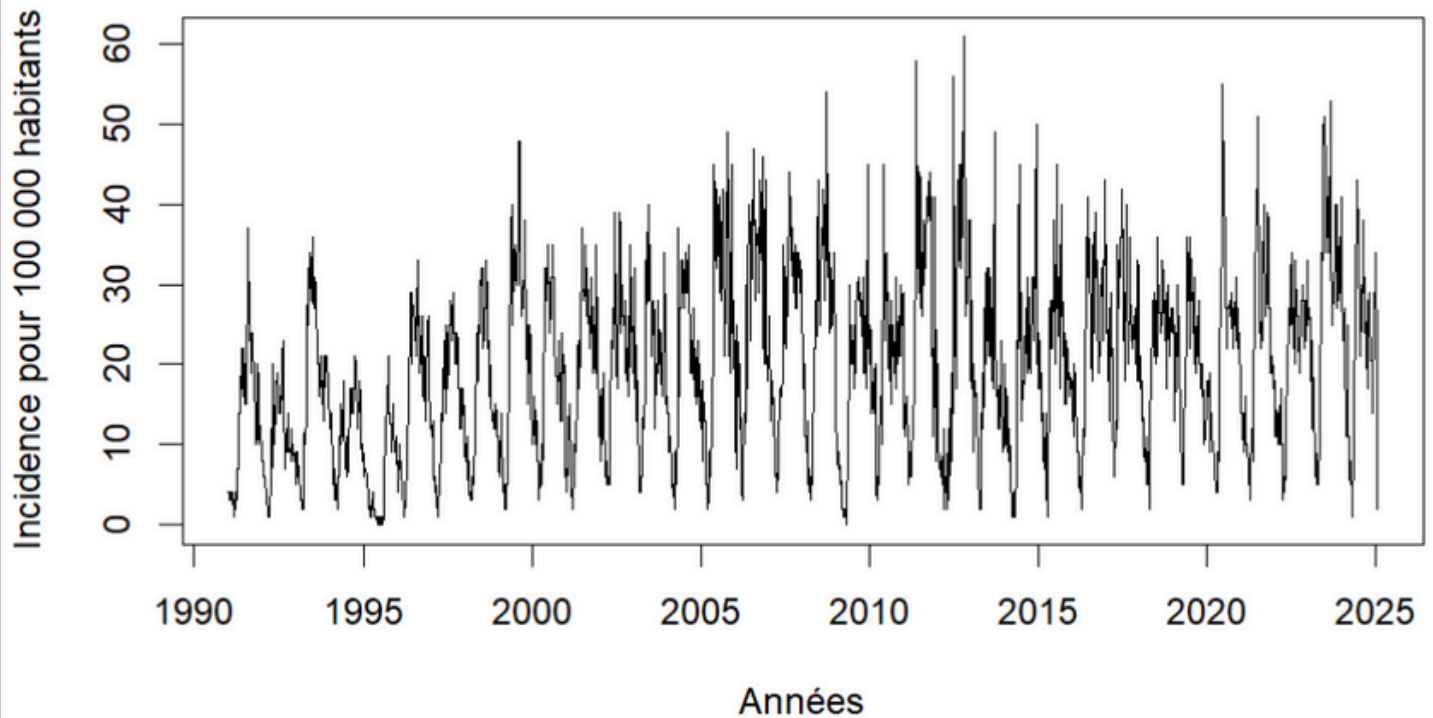
LES DONNÉES CHARGÉES SONT ENSUITE STOCKÉES DANS L'OBJET VARICELLE_DATA, QUI SERA UTILISÉ POUR LES ANALYSES.

ANALYSE DE LA SÉRIE TEMPORELLE DES TAUX D'INCIDENCE DE LA VARICELLE

LES DONNÉES HEBDOMADAIRES DE 1991 À 2025, CONVERTIES EN SÉRIE TEMPORELLE AVEC UNE FRÉQUENCE DE 52 OBSERVATIONS PAR AN, RÉVÈLENT UNE VARIABILITÉ SAISONNIÈRE MARQUÉE, TYPIQUE DES MALADIES INFECTIEUSES COMME LA VARICELLE. ON OBSERVE DES PICS ÉPIDÉMIQUES RÉGULIERS SUIVIS DE CREUX OÙ LES TAUX D'INCIDENCE DIMINUENT FORTEMENT.

UNE BAISSSE PROGRESSIVE DES TAUX D'INCIDENCE EST NOTABLE DEPUIS LES ANNÉES 2000, AVEC DES NIVEAUX FAIBLES VOIRE NULS ENTRE 2020 ET 2025. CETTE TENDANCE POURRAIT S'EXPLIQUER PAR L'INTRODUCTION DE LA VACCINATION, DES MESURES DE PRÉVENTION OU DES CHANGEMENTS SOCIAUX. CES RÉSULTATS ILLUSTRONT L'UTILITÉ DES SÉRIES TEMPORELLES POUR COMPRENDRE ET SUIVRE LES DYNAMIQUES ÉPIDÉMIOLOGIQUES.

Taux d'incidence de la varicelle



ANALYSE DU GRAPHIQUE SUR L'INCIDENCE DE LA VARICELLE (1990-2025)

TENDANCE GÉNÉRALE :

LE GRAPHIQUE NE MONTRE PAS DE TENDANCE NETTE À LA HAUSSE OU À LA BAISSÉ. LES TAUX D'INCIDENCE OSCILLENT AUTOUR D'UNE MOYENNE RELATIVEMENT STABLE TOUT AU LONG DE LA PÉRIODE ÉTUDIÉE.

SAISONNALITÉ :

UNE COMPOSANTE SAISONNIÈRE EST OBSERVABLE AVEC DES PICS RÉCURRENTS. CES FLUCTUATIONS PÉRIODIQUES POURRAIENT ÊTRE LIÉES À DES FACTEURS ENVIRONNEMENTAUX OU COMPORTEMENTAUX INFLUENÇANT LA PROPAGATION DE LA VARICELLE.

VARIABILITÉ :

L'AMPLITUDE DES PICS D'INCIDENCE VARIE SELON LES ANNÉES. CERTAINES PÉRIODES PRÉSENTENT DES TAUX PLUS ÉLEVÉS, CE QUI PEUT CORRESPONDRE À DES ÉPIDÉMIES LOCALISÉES OU À DES VARIATIONS DANS LA TRANSMISSION DE LA MALADIE.

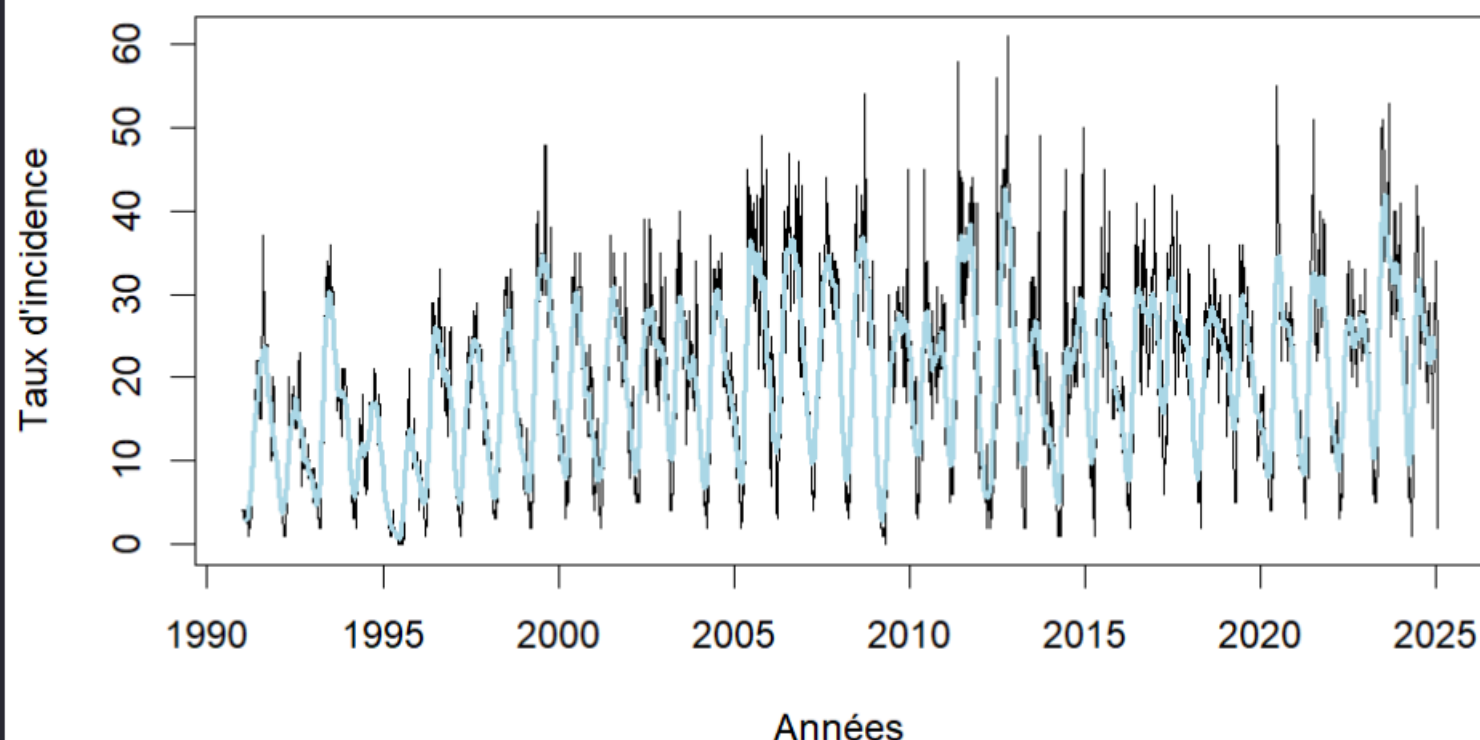
ANOMALIES OU RUPTURES :

LE GRAPHIQUE NE PRÉSENTE PAS DE RUPTURE DE TENDANCE ÉVIDENTE. TOUTEFOIS, CERTAINES FLUCTUATIONS IMPORTANTES POURRAIENT ÊTRE LIÉES À DES ÉVÉNEMENTS SPÉCIFIQUES, TELS QUE L'INTRODUCTION DE CAMPAGNES DE VACCINATION OU DES CHANGEMENTS DANS LA MÉTHODOLOGIE DE COLLECTE DES DONNÉES.

PÉRIODE RÉCENTE (2020-2025) :

LES DONNÉES LES PLUS RÉCENTES MONTRENT UNE VARIABILITÉ SIMILAIRE AUX ANNÉES PRÉCÉDENTES, SUGGÉRANT QUE LA DYNAMIQUE GLOBALE DE LA VARICELLE RESTE INCHANGÉE.

Tendance avec Moyennes Mobiles



ANALYSE DU GRAPHIQUE AVEC MOYENNES MOBILES (1990-2025)

TENDANCE GÉNÉRALE :

LE LISSAGE PAR MOYENNES MOBILES (COURBE BLEUE) PERMET DE MIEUX VISUALISER LA TENDANCE GLOBALE DES DONNÉES. ON OBSERVE UNE STABILITÉ DE LA TENDANCE GÉNÉRALE, SANS AUGMENTATION NI DIMINUTION MARQUÉE DU TAUX D'INCIDENCE SUR LA PÉRIODE ÉTUDIÉE.

SAISONNALITÉ :

LA COMPOSANTE SAISONNIÈRE RESTE VISIBLE MÊME APRÈS L'APPLICATION DES MOYENNES MOBILES. LES FLUCTUATIONS RÉGULIÈRES INDIQUENT DES CYCLES ANNUELS, CARACTÉRISÉS PAR DES PICS PÉRIODIQUES.

VARIABILITÉ RÉDUITE :

PAR RAPPORT AU GRAPHIQUE BRUT, LES MOYENNES MOBILES ATTÉNUENT LES FLUCTUATIONS À COURT TERME ET METTENT EN ÉVIDENCE LES VARIATIONS IMPORTANTES SUR DES PÉRIODES PLUS LONGUES. CELA FACILITE L'OBSERVATION DES TENDANCES STRUCTURELLES.

RÉGULARITÉ DES PICS :

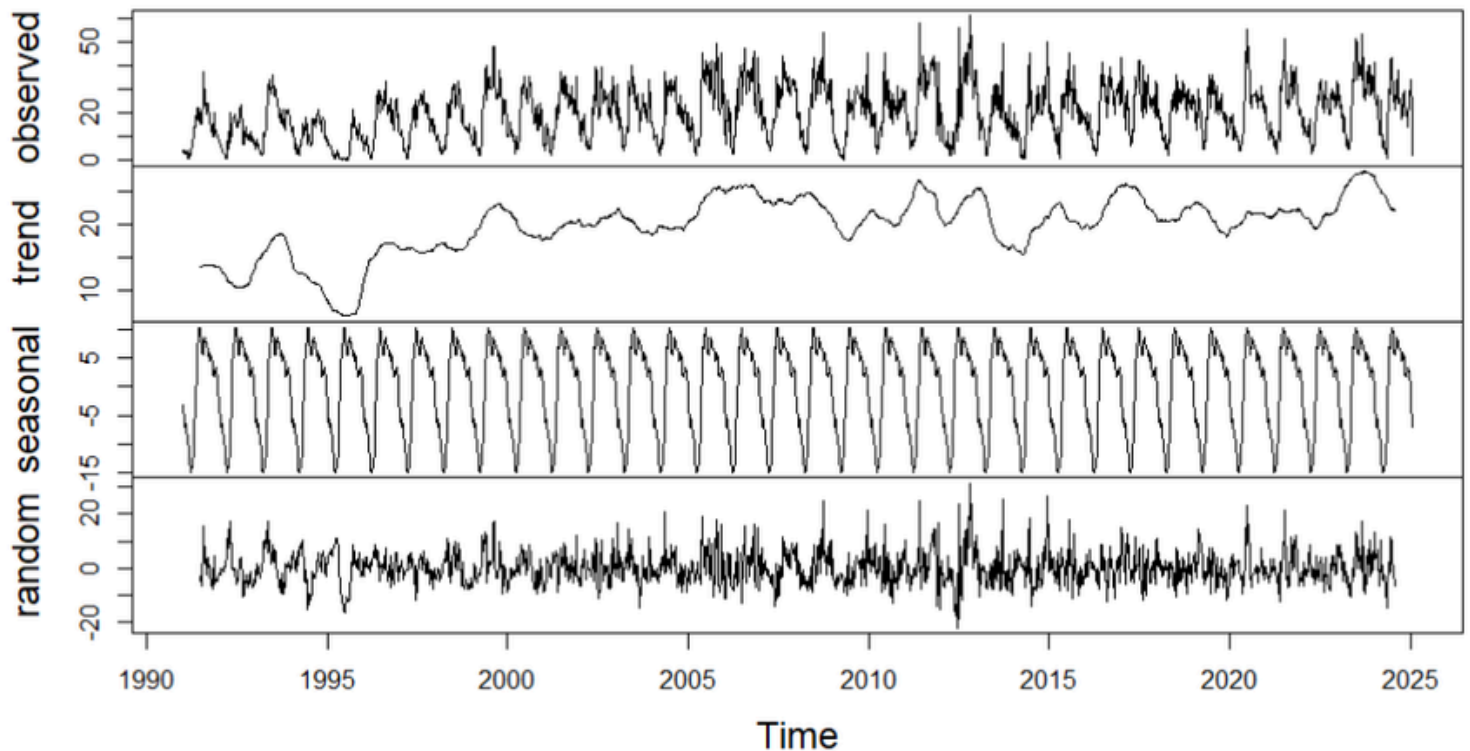
LES PICS DE TAUX D'INCIDENCE SONT RÉGULIÈREMENT ESPACÉS, CONFIRMANT L'HYPOTHÈSE D'UNE SAISONNALITÉ BIEN DÉFINIE LIÉE À DES FACTEURS ENVIRONNEMENTAUX OU ÉPIDÉMIOLOGIQUES.

PÉRIODE RÉCENTE (2020-2025) :

LE COMPORTEMENT DES MOYENNES MOBILES SUR LA PÉRIODE RÉCENTE EST COHÉRENT AVEC CELUI DES ANNÉES PRÉCÉDENTES. AUCUNE RUPTURE DE TENDANCE OU CHANGEMENT NOTABLE N'EST VISIBLE.

L'AJOUT DES MOYENNES MOBILES ENRICHIT L'ANALYSE EN FILTRANT LE BRUIT

Decomposition of additive time series



DÉCOMPOSITION D'UNE SÉRIE TEMPORELLE ADDITIVE (1990-2025)

GRAPHIQUE OBSERVÉ (OBSERVED) :

LA COURBE INITIALE REPRÉSENTE LES DONNÉES BRUTES. ON Y OBSERVE UNE FORTE VARIABILITÉ COMBINANT UNE TENDANCE, UNE SAISONNALITÉ ET DES COMPOSANTES ALÉATOIRES.

COMPOSANTE TENDANCE (TREND) :

CE GRAPHIQUE MET EN ÉVIDENCE LA TENDANCE GLOBALE SUR LA PÉRIODE ÉTUDIÉE. APRÈS UNE LÉGÈRE BAISSÉ INITIALE, UNE AUGMENTATION PROGRESSIVE EST PERCEPTIBLE JUSQU'EN 2020, SUIVIE D'UNE LÉGÈRE DIMINUTION SUR LA FIN DE LA PÉRIODE.

COMPOSANTE SAISONNALITÉ (SEASONAL) :

LA DÉCOMPOSITION RÉVÈLE UNE SAISONNALITÉ RÉGULIÈRE ET STABLE TOUT AU LONG DE LA PÉRIODE. LES FLUCTUATIONS PÉRIODIQUES SEMBLANT UNIFORMES, INDICANT DES CYCLES RÉPÉTITIFS COHÉRENTS.

COMPOSANTE ALÉATOIRE (RANDOM) :

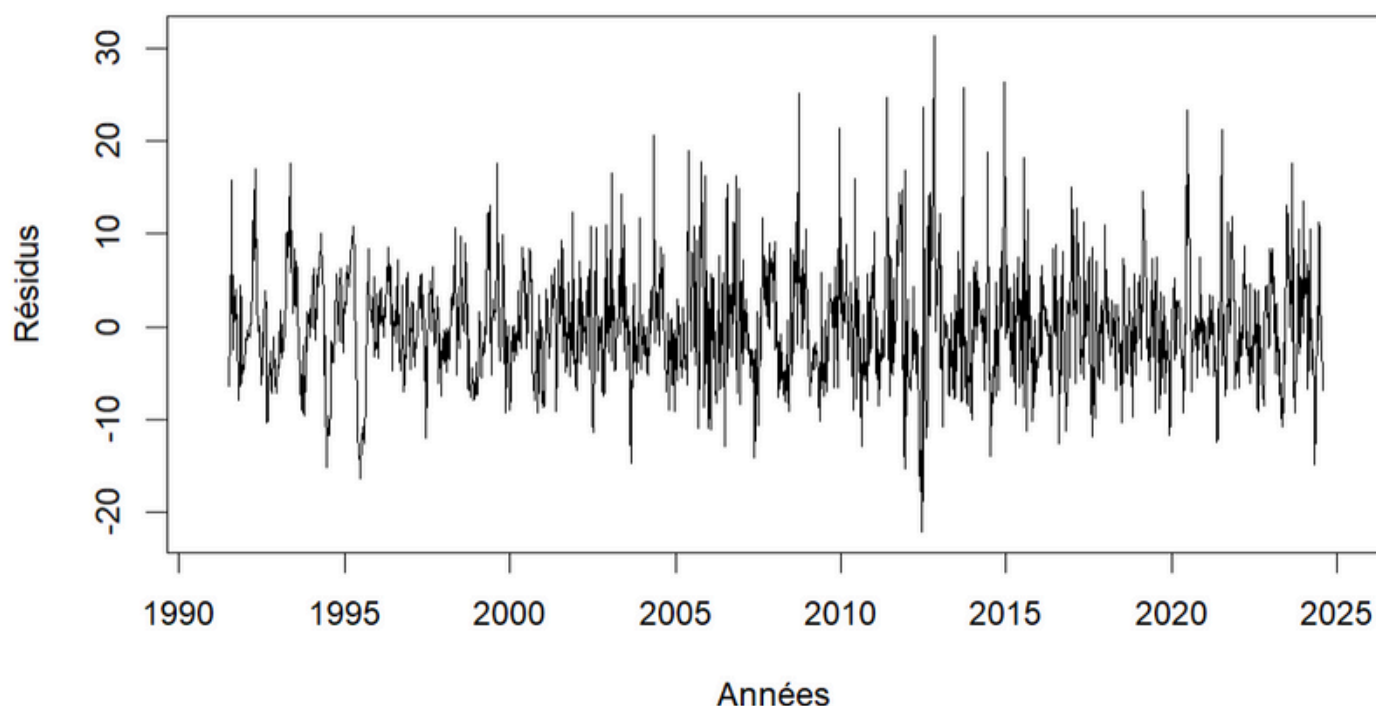
LES VARIATIONS RÉSIDUELLES REPRÉSENTENT LES IRRÉGULARITÉS ET LE BRUIT DES DONNÉES. CES FLUCTUATIONS NE SUIVENT PAS DE TENDANCE OU DE SAISONNALITÉ SPÉCIFIQUES ET PEUVENT ÊTRE ATTRIBUÉES À DES FACTEURS IMPRÉVISIBLES.

UTILITÉ DE LA DÉCOMPOSITION :

CETTE DÉCOMPOSITION EN TROIS COMPOSANTES (TENDANCE, SAISONNALITÉ ET RÉSIDU) PERMET UNE ANALYSE APPROFONDIE DES SÉRIES TEMPORELLES. ELLE AIDE À ISOLER LES PHÉNOMÈNES CYCLIQUES DES VARIATIONS À LONG TERME ET DES VARIATIONS NON EXPLIQUÉES.

LA DÉCOMPOSITION RÉVÈLE UNE TENDANCE EN HAUSSE JUSQU'EN 2020, SUIVIE D'UNE LÉGÈRE BAISSÉ, UNE SAISONNALITÉ STABLE AVEC DES CYCLES RÉGULIERS, ET DES RÉSIDUS ALÉATOIRES IMPRÉVISIBLES. CELA MET EN

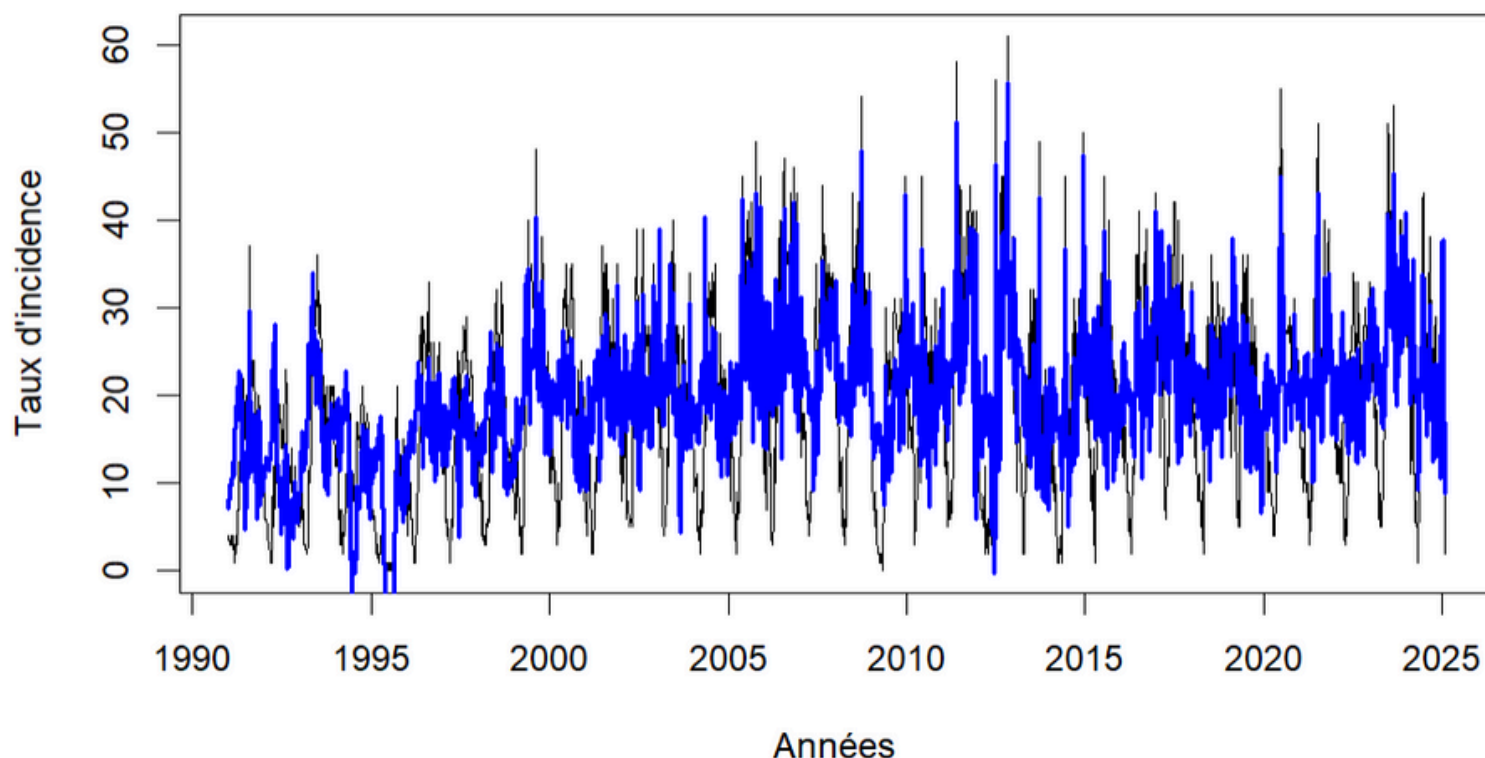
Résidus de la Série



LE GRAPHIQUE DES RÉSIDUS, DANS LE CONTEXTE DE L'ANALYSE DES TAUX D'INCIDENCE POUR 100 000 HABITANTS ASSOCIÉ À L'INDICATEUR, FOURNIT UNE VISION PRÉCISE DES DIFFÉRENCES ENTRE LES VALEURS OBSERVÉES ET LES VALEURS PRÉVUES PAR LE MODÈLE STATISTIQUE. LES RÉSIDUS OSCILLENT AUTOUR DE LA LIGNE ZÉRO SUR L'AXE DES X, COUVRANT LES ANNÉES DE 1990 À 2025, ET VARIENT ENTRE -20 ET 30 SUR L'AXE DES Y. CES FLUCTUATIONS AUTOUR DE ZÉRO INDIQUENT QUE LE MODÈLE PARVIENT GÉNÉRALEMENT À CAPTURER LES TENDANCES GLOBALES DES DONNÉES, BIEN QUE CERTAINES VARIATIONS NOTABLES SOIENT PRÉSENTES. CES ÉCARTS PEUVENT SIGNALER DES MOMENTS OÙ LE MODÈLE EST MOINS PRÉCIS, PEUT-ÊTRE EN RAISON D'ÉVÉNEMENTS EXCEPTIONNELS OU D'ANOMALIES DANS LES DONNÉES.

UN EXAMEN ATTENTIF DES RÉSIDUS EST CRUCIAL POUR IDENTIFIER LES BIAIS OU ERREURS SYSTÉMATIQUES DANS LE MODÈLE. UN MODÈLE CORRECTEMENT AJUSTÉ DEVRAIT MONTRER DES RÉSIDUS RÉPARTIS DE MANIÈRE ALÉATOIRE AUTOUR DE ZÉRO SANS TENDANCE APPARENTE. LES RÉSIDUS IMPORTANTS PEUVENT INDiquer DES ASPECTS DU MODÈLE NÉCESSITANT DES AJUSTEMENTS. L'ANALYSE DES RÉSIDUS PERMET DONC DE VÉRIFIER LA PERFORMANCE DU MODÈLE ET DE DÉTERMINER LES POINTS À AMÉLIORER POUR RENFORCER LA PRÉCISION ET LA FIABILITÉ DES PRÉVISIONS DU TAUX D'INCIDENCE.

Série Désaisonnalisée



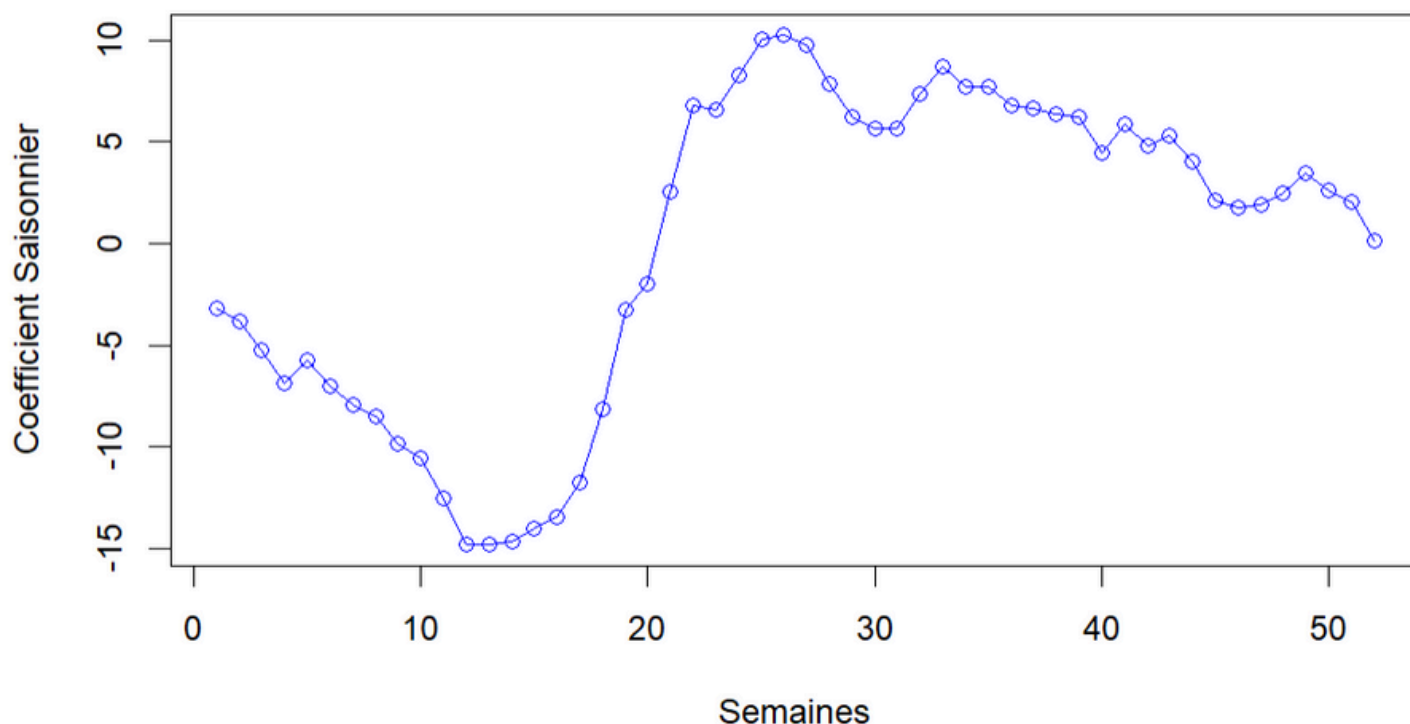
LE GRAPHIQUE ILLUSTRE UNE SÉRIE TEMPORELLE DÉSAISONNALISÉE REPRÉSENTANT LE TAUX D'INCIDENCE POUR 100 000 HABITANTS. LA DÉSAISONNALISATION CONSISTE À RETIRER LES VARIATIONS SAISONNIÈRES DE LA SÉRIE ORIGINALE POUR METTRE EN ÉVIDENCE LES TENDANCES GÉNÉRALES ET LES FLUCTUATIONS INHABITUELLES. DANS CE GRAPHIQUE, LA COURBE BLEUE REPRÉSENTE LA SÉRIE DÉSAISONNALISÉE, C'EST-À-DIRE LES DONNÉES CORRIGÉES DES EFFETS SAISONNIERS, TANDIS QUE LA COURBE NOIRE CORRESPOND À LA SÉRIE ORIGINALE AVANT CORRECTION.

ENTRE 1990 ET 2010, LA COURBE BLEUE MET EN ÉVIDENCE UNE AUGMENTATION GLOBALE DU TAUX D'INCIDENCE, BIEN QUE DES VARIATIONS IRRÉGULIÈRES SOIENT TOUJOURS PRÉSENTES. APRÈS 2010, LA SÉRIE MONTRE UNE STABILISATION RELATIVE, AVEC DES FLUCTUATIONS IMPORTANTES QUI TRADUISENT DES ÉVÉNEMENTS EXCEPTIONNELS COMME DES ÉPIDÉMIES PONCTUELLES OU DES CHANGEMENTS DANS LES POLITIQUES DE SANTÉ PUBLIQUE.

IL EST ÉGALEMENT IMPORTANT DE NOTER QU'À PARTIR DE 2020, LES SYNDROMES GRIPPAUX SONT ESTIMÉS À PARTIR DES DONNÉES DES INFECTIONS RESPIRATOIRES AIGÜES (IRA). CE CHANGEMENT MÉTHODOLOGIQUE POURRAIT INFLUENCER LES VARIATIONS DE LA COURBE BLEUE PAR RAPPORT AUX PÉRIODES PRÉCÉDENTES.

EN CONCLUSION, LA COURBE BLEUE PERMET DE MIEUX COMPRENDRE LES TENDANCES DE FOND EN ÉLIMINANT LES VARIATIONS SAISONNIÈRES. CELA FACILITE L'INTERPRÉTATION DES DONNÉES ET CONSTITUE UN OUTIL ESSENTIEL POUR ANALYSER LES SÉRIES TEMPORELLES ET ORIENTER LES ACTIONS DE SANTÉ PUBLIQUE.

Coefficients Saisonniers sur une Année

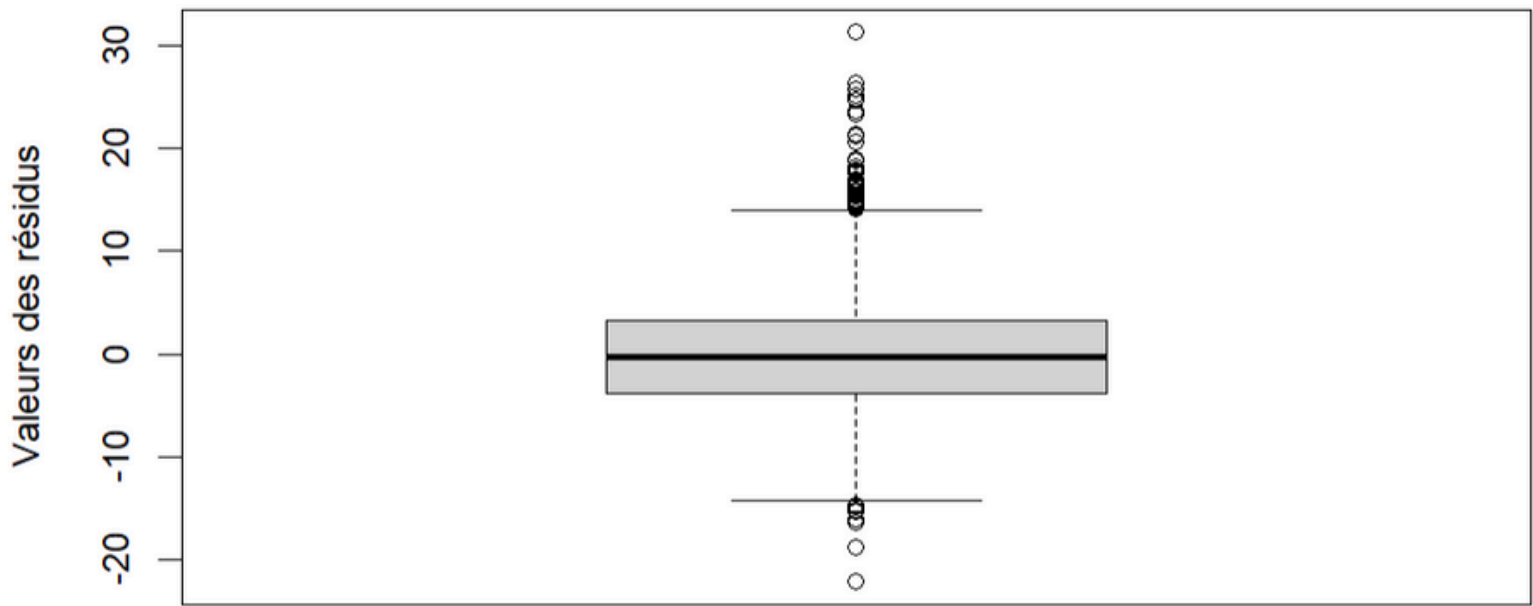


CE GRAPHIQUE REPRÉSENTE LES COEFFICIENTS SAISONNIERS CALCULÉS SUR UNE ANNÉE, EN FONCTION DES SEMAINES. LES COEFFICIENTS SAISONNIERS ILLUSTRONT LES VARIATIONS PÉRIODIQUES DU TAUX D'INCIDENCE, DUES AUX EFFETS SAISONNIERS RÉCURRENTS. CES COEFFICIENTS SONT OBTENUS EN DÉCOMPOSANT LA SÉRIE TEMPORELLE INITIALE, ET ILS PERMETTENT DE COMPRENDRE LES FLUCTUATIONS SAISONNIÈRES.

ON OBSERVE ICI QUE LES COEFFICIENTS SAISONNIERS SUIVENT UNE COURBE OSCILLANTE. LES VALEURS SONT LES PLUS BASSES (AUTOUR DE -15) ENTRE LES SEMAINES 5 ET 20, INDICANT QUE CETTE PÉRIODE CONNAÎT UNE DIMINUTION SIGNIFICATIVE DU TAUX D'INCIDENCE DUE AUX EFFETS SAISONNIERS. ENSUITE, LES COEFFICIENTS AUGMENTENT RAPIDEMENT POUR ATTEINDRE UN PIC POSITIF (ENVIRON 10) VERS LA SEMAINE 25. CETTE AUGMENTATION REFLÈTE UNE PÉRIODE OÙ LES EFFETS SAISONNIERS AMPLIFIENT LE TAUX D'INCIDENCE. APRÈS CE PIC, LES COEFFICIENTS DIMINUENT PROGRESSIVEMENT, SUGGÉRANT UN RETOUR À DES VALEURS PLUS MODÉRÉES JUSQU'À LA FIN DE L'ANNÉE.

CE GRAPHIQUE PERMET DE VISUALISER LES PÉRIODES OÙ LES EFFETS SAISONNIERS INFLUENCENT LE PLUS FORTEMENT LES DONNÉES. IL EST PARTICULIÈREMENT UTILE POUR AJUSTER LES SÉRIES TEMPORELLES ET POUR EFFECTUER DES PRÉVISIONS PLUS PRÉCISES EN ÉLIMINANT CES VARIATIONS SAISONNIÈRES.

Résidus

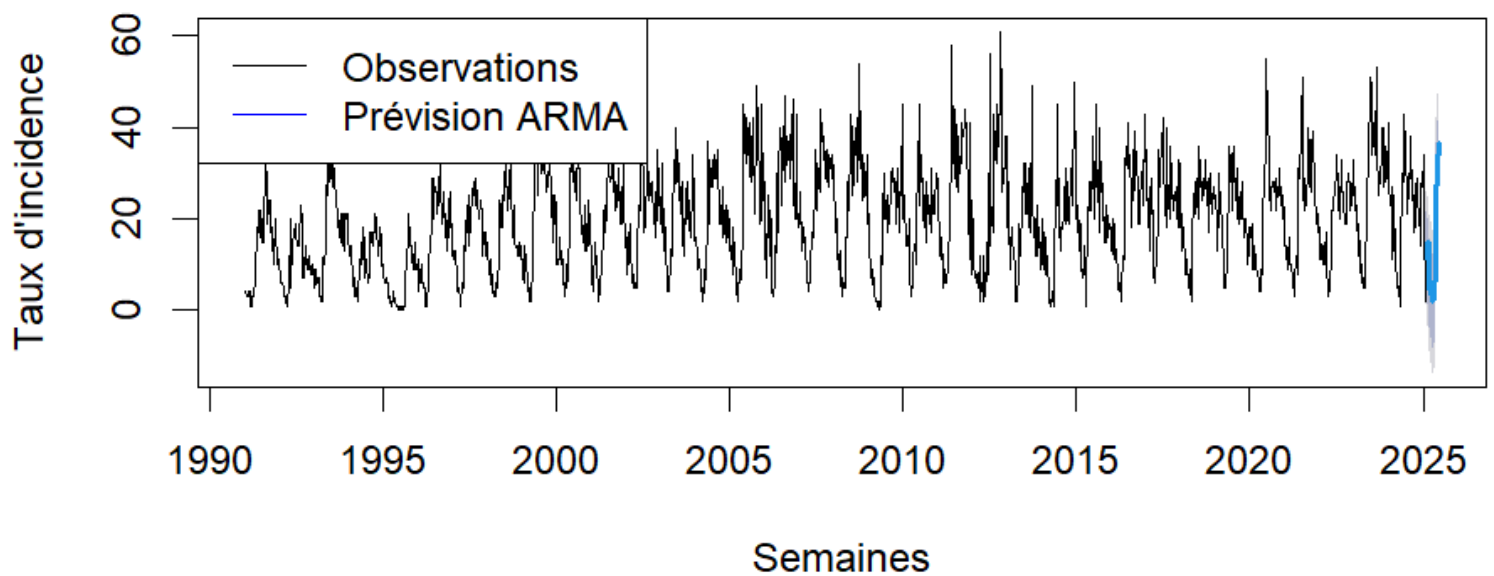


CE GRAPHIQUE EN BOÎTE, OU BOXPLOT, REPRÉSENTE LES RÉSIDUS OBTENUS APRÈS SOUSTRACTION DES COMPOSANTES DE TENDANCE ET SAISONNIÈRE DE LA SÉRIE TEMPORELLE INITIALE. LES RÉSIDUS, QUI SONT LA DIFFÉRENCE ENTRE LES VALEURS OBSERVÉES ET LES VALEURS PRÉVUES PAR LE MODÈLE, PERMETTENT D'ÉVALUER LA QUALITÉ DE L'AJUSTEMENT DU MODÈLE SUR LES DONNÉES.

SUR CE GRAPHIQUE, LES RÉSIDUS SONT PRINCIPALEMENT CONCENTRÉS AUTOUR DE ZÉRO, CE QUI INDIQUE QUE LE MODÈLE CAPTURE BIEN LES COMPOSANTES PRINCIPALES DE LA SÉRIE TEMPORELLE (TENDANCE ET SAISONNALITÉ). LA BOÎTE GRISE CENTRALE MONTRE LA MÉDIANE DES RÉSIDUS (LIGNE NOIRE AU CENTRE) ET L'ÉTENDUE DES RÉSIDUS POUR LE PREMIER ET LE TROISIÈME QUARTILE (LIMITES DE LA BOÎTE). LES "MOUSTACHES" S'ÉTENDENT JUSQU'AUX VALEURS SITUÉES DANS UNE PLAGE RAISONNABLE AUTOUR DE LA BOÎTE, ET LES POINTS SITUÉS AU-DELÀ SONT CONSIDÉRÉS COMME DES VALEURS ATYPIQUES OU OUTLIERS.

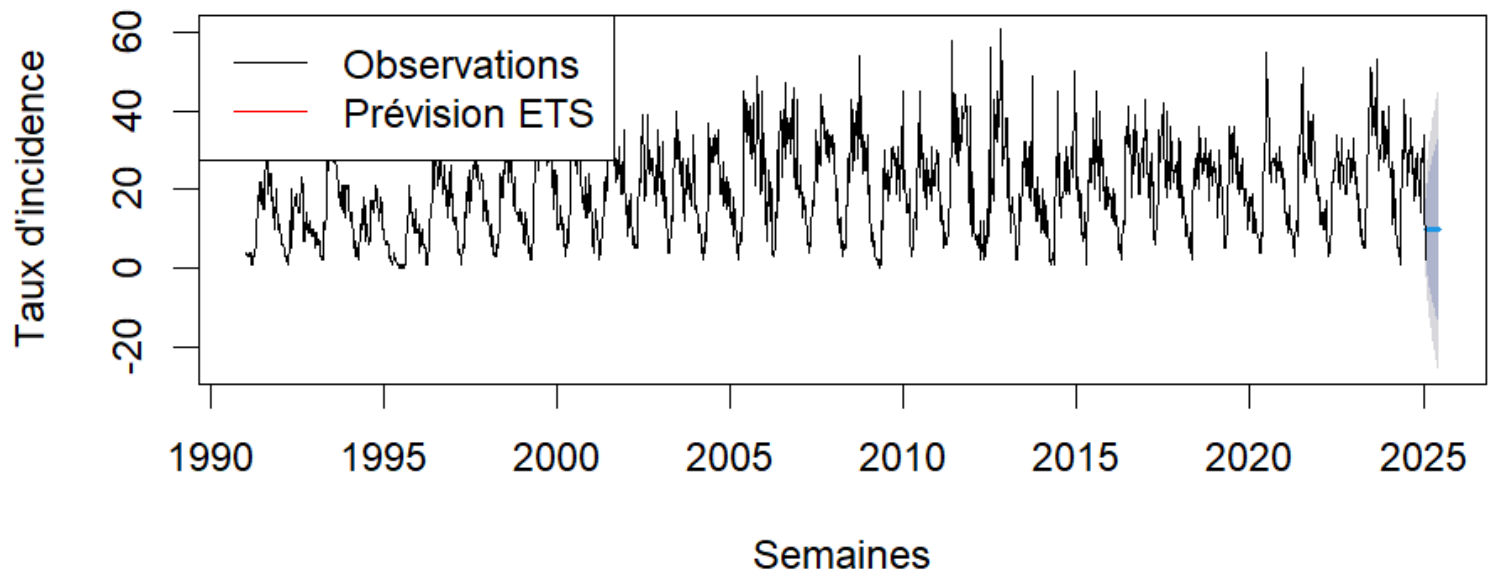
LES VALEURS ATYPIQUES, REPRÉSENTÉES PAR DES CERCLES AU-DESSUS ET EN DESSOUS DES MOUSTACHES, INDICENT DES RÉSIDUS QUI S'ÉCARTENT FORTEMENT DES PRÉVISIONS DU MODÈLE. CES POINTS POURRAIENT CORRESPONDRE À DES ANOMALIES OU À DES VARIATIONS NON EXPLIQUÉES PAR LE MODÈLE. CE GRAPHIQUE EST UN OUTIL UTILE POUR VÉRIFIER SI LES HYPOTHÈSES DE DISTRIBUTION NORMALE ET DE VARIANCE CONSTANTE DES RÉSIDUS SONT RESPECTÉES, DEUX CONDITIONS IMPORTANTES POUR VALIDER LE MODÈLE STATISTIQUE.

Prévisions avec Modèle ARMA



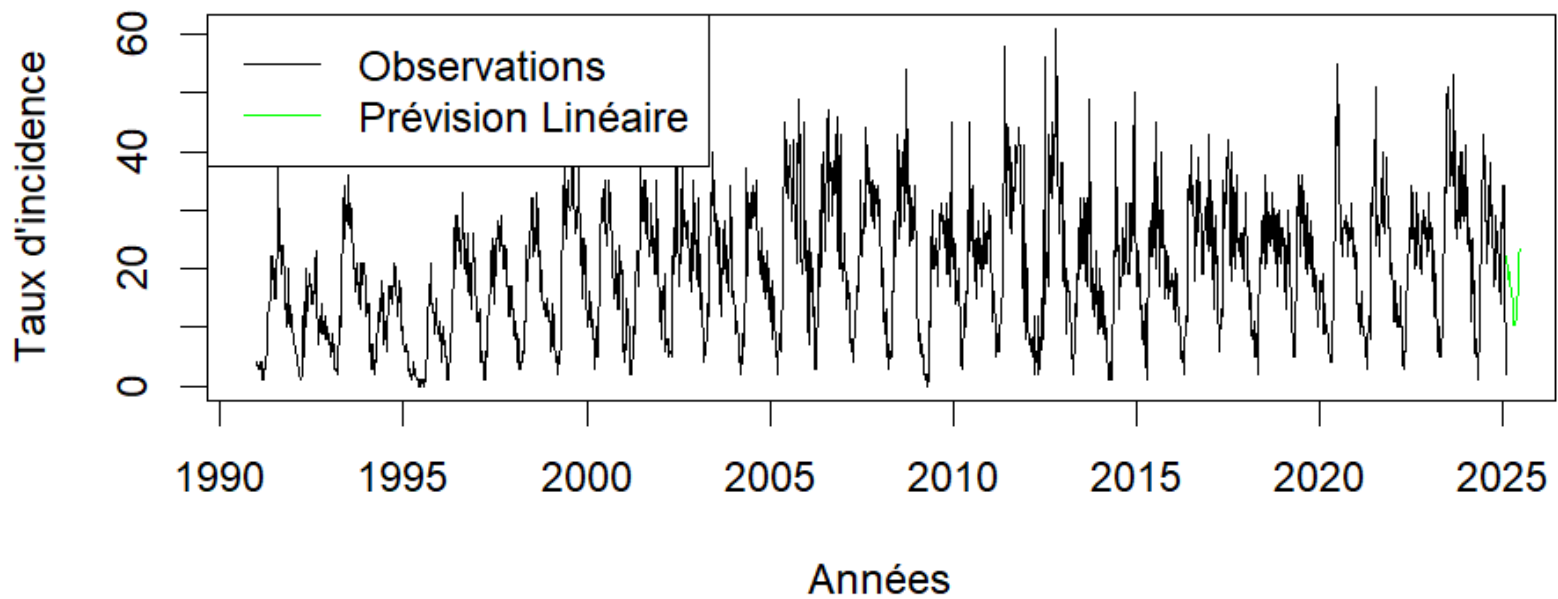
LE GRAPHIQUE DES PRÉVISIONS ARMA MONTRE LES TAUX D'INCIDENCE OBSERVÉS (LIGNE NOIRE) ET PRÉVUS (LIGNE BLEUE) DES SYNDROMES GRIPPAUX SUR UNE PÉRIODE ALLANT DE 1990 À 2025. LA COURBE DES OBSERVATIONS MET EN ÉVIDENCE DES VARIATIONS SAISONNIÈRES RÉGULIÈRES AVEC DES PICS ANNUELS, TYPQUES DES ÉPIDÉMIES DE GRIPPE EN HIVER. LA COURBE DES PRÉVISIONS ARMA SUIT DE PRÈS CETTE TENDANCE, INDIQUANT QUE LE MODÈLE CAPTE BIEN LA SAISONNALITÉ ET LES VARIATIONS DES DONNÉES HISTORIQUES. LES PRÉVISIONS JUSQU'EN 2025 SUGGÈRENT UNE CONTINUITÉ DES SCHÉMAS SAISONNIERS OBSERVÉS, BIEN QUE DES ÉCARTS PUISSENT APPARAÎTRE EN RAISON DE FACTEURS IMPRÉVUS OU DE LIMITATIONS DU MODÈLE. L'ÉVALUATION DE LA PRÉCISION DES PRÉVISIONS PAR RAPPORT AUX DONNÉES RÉCENTES EST ESSENTIELLE POUR VALIDER LA FIABILITÉ DU MODÈLE ARMA.

Prévisions avec Modèle ETS



LE GRAPHIQUE DES PRÉVISIONS ETS MONTRE LES TAUX D'INCIDENCE OBSERVÉS (LIGNE NOIRE) ET PRÉVUS (LIGNE ROUGE) DES SYNDROMES GRIPPAUX POUR LA PÉRIODE ALLANT DE 1990 À 2025. LA COURBE DES OBSERVATIONS ILLUSTRÉ DES VARIATIONS SAISONNIÈRES RÉGULIÈRES AVEC DES PICS ANNUELS PENDANT LES PÉRIODES HIVERNALES. LA COURBE DES PRÉVISIONS ETS SUIT DE PRÈS CETTE TENDANCE, SUGGÉRANT QUE LE MODÈLE ETS CAPTURE BIEN LA SAISONNALITÉ ET LES VARIATIONS DES DONNÉES HISTORIQUES. LES PRÉVISIONS JUSQU'EN 2025 MONTRENT UNE CONTINUITÉ DES SCHÉMAS SAISONNIERS OBSERVÉS, BIEN QUE DES ÉCARTS PUISSENT APPARAÎTRE EN RAISON DE FACTEURS IMPRÉVUS OU DES LIMITATIONS DU MODÈLE. L'ÉVALUATION DE LA PRÉCISION DES PRÉVISIONS EN COMPARANT AVEC LES DONNÉES RÉCENTES EST ESSENTIELLE POUR VALIDER LA FIABILITÉ DU MODÈLE ETS.

Prévisions avec Modèle Linéaire



LE GRAPHIQUE DES PRÉVISIONS LINÉAIRES MONTRE LES TAUX D'INCIDENCE OBSERVÉS (LIGNE NOIRE) ET PRÉVUS (LIGNE VERTE) DES SYNDROMES GRIPPAUX ENTRE 1990 ET 2025. LA COURBE DES OBSERVATIONS ILLUSTRE LES FLUCTUATIONS SAISONNIÈRES TYPIQUES AVEC DES PICS ANNUELS, EN PARTICULIER DURANT LES PÉRIODES HIVERNALES. LA COURBE DE PRÉVISION LINÉAIRE SUIT UNE TENDANCE QUI SEMBLE CAPTURER LES VARIATIONS GÉNÉRALES DES DONNÉES HISTORIQUES, MAIS POURRAIT NE PAS REFLÉTER PARFAITEMENT LA SAISONNALITÉ DÉTAILLÉE DES INCIDENCES. LES PRÉVISIONS JUSQU'EN 2025 MONTRENT UNE TENDANCE À LA BAISSSE DES TAUX D'INCIDENCE, BIEN QUE DES ÉCARTS SOIENT POSSIBLES EN RAISON DE FACTEURS IMPRÉVUS OU DES LIMITATIONS INHÉRENTES AU MODÈLE LINÉAIRE. UNE ÉVALUATION DES PRÉVISIONS COMPARÉES AUX DONNÉES RÉCENTES EST NÉCESSAIRE POUR VALIDER LA FIABILITÉ DE CE MODÈLE.

```
> cat("EQM Modèle ARMA :", round(EQM_arma, 2), "\n")
EQM Modèle ARMA : 43.4
> cat("EQM Modèle ETS :", round(EQM_ets, 2), "\n")
EQM Modèle ETS : 103.39
> cat("EQM Modèle Linéaire :", round(EQM_l, 2), "\n")
EQM Modèle Linéaire : 105.97
```

L'IMAGE PRÉSENTE UNE ÉVALUATION DE LA PRÉCISION DE TROIS MODÈLES STATISTIQUES DIFFÉRENTS : ARMA, ETS ET LINÉAIRE, EN UTILISANT L'EQM (ERREUR QUADRATIQUE MOYENNE). LES RÉSULTATS MONTRENT QUE LE MODÈLE ARMA A L'EQM LA PLUS BASSE À 43.4, INDIQUANT QU'IL EST LE PLUS PRÉCIS DES TROIS POUR PRÉDIRE LES TAUX D'INCIDENCE DES SYNDROMES GRIPPAUX. LE MODÈLE ETS, AVEC UNE EQM DE 103.39, EST MOINS PRÉCIS, SUIVI PAR LE MODÈLE LINÉAIRE, QUI A L'EQM LA PLUS ÉLEVÉE À 105.97. CES RÉSULTATS SUGGÈRENT QUE LE MODÈLE ARMA CAPTURE MIEUX LES TENDANCES ET LA SAISONNALITÉ DES DONNÉES HISTORIQUES COMPARÉ AUX MODÈLES ETS ET LINÉAIRE. AINSI, LE MODÈLE ARMA SERAIT LE PLUS RECOMMANDÉ POUR DES PRÉVISIONS FUTURES SUR L'INCIDENCE DES SYNDROMES GRIPPAUX.

Analyse en anglais :

MODELING AND FORECASTING THE INCIDENCE RATE OF CHICKENPOX IN FRANCE

THIS REPORT FOCUSES ON ANALYZING AND PREDICTING WEEKLY INCIDENCE RATES OF CHICKENPOX IN FRANCE USING TIME SERIES DATA COLLECTED SINCE 1991. THE MAIN GOAL IS TO IDENTIFY PATTERNS, SEASONAL TRENDS, AND FORECAST FUTURE RATES. THE PROJECT ALSO EVALUATES DIFFERENT FORECASTING MODELS TO DETERMINE THE MOST EFFECTIVE ONE.

THE DATA COMES FROM THE "RÉSEAU SENTINELLES" AND INCLUDES WEEKLY CASE COUNTS AND INCIDENCE RATES PER 100,000 INHABITANTS. KEY STEPS IN THE ANALYSIS INCLUDE:

DATA PREPARATION AND EXPLORATION: CLEANING AND ORGANIZING THE DATASET TO IDENTIFY KEY TRENDS.

TIME SERIES DECOMPOSITION: BREAKING DOWN THE DATA INTO TREND, SEASONAL, AND RANDOM COMPONENTS.

FORECASTING MODELS: USING ARMA, ETS, AND LINEAR REGRESSION MODELS TO PREDICT FUTURE INCIDENCE RATES.

MODEL EVALUATION: COMPARING MODEL PERFORMANCES USING MEAN SQUARED ERROR (MSE).

KEY FINDINGS INCLUDE STRONG SEASONAL PATTERNS WITH REGULAR ANNUAL PEAKS AND A NOTICEABLE DECLINE IN INCIDENCE SINCE 2000, LIKELY DUE TO VACCINATION AND OTHER HEALTH MEASURES. AMONG THE MODELS, ARMA PROVIDED THE MOST ACCURATE PREDICTIONS.

THIS STUDY HIGHLIGHTS THE IMPORTANCE OF TIME SERIES ANALYSIS IN UNDERSTANDING AND MANAGING EPIDEMIOLOGICAL TRENDS.

RÉPARTITION DES TÂCHES

DANS LE CADRE DE CE PROJET, NOUS AVONS RÉPARTI LES TÂCHES DE MANIÈRE ÉQUILIBRÉE TOUT EN COLLABORANT ÉTROITEMENT SUR PLUSIEURS ÉTAPES CLÉS. CETTE ORGANISATION A PERMIS À CHACUNE D'ENTRE NOUS DE SE CONCENTRER SUR DES ASPECTS SPÉCIFIQUES, TOUT EN TRAVAILLANT ENSEMBLE POUR GARANTIR UNE COHÉRENCE ET UNE MAÎTRISE COMPLÈTE DES ANALYSES EFFECTUÉES.

ASSIA AOUIMEUR :

A PRÉPARÉ ET EXPLORÉ LES DONNÉES, EN S'OCCUPANT DU NETTOYAGE, DE L'ORGANISATION DES INFORMATIONS ET DE LA CRÉATION DES BASES NÉCESSAIRES POUR L'ANALYSE.

A RÉALISÉ L'ANALYSE DE LA TENDANCE EN APPLIQUANT LES MOYENNES MOBILES ET EN INTERPRÉTANT LES RÉSULTATS ASSOCIÉS.

A CONTRIBUÉ À LA CRÉATION DES PRÉVISIONS AVEC LE MODÈLE ARMA, EN AJUSTANT LES PARAMÈTRES ET EN ÉVALUANT SES PERFORMANCES.

A PARTICIPÉ À LA RÉDACTION DES SECTIONS EXPLIQUANT LA DÉCOMPOSITION DE LA SÉRIE TEMPORELLE, EN SE CONCENTRANT SUR L'ANALYSE DE LA SAISONNALITÉ.

INÈS DRIS :

A CALCULÉ LES COEFFICIENTS SAISONNIERS ET INTERPRÉTÉ LEURS VARIATIONS, TOUT EN TRAÇANT ET COMMENTANT LES GRAPHIQUES ASSOCIÉS.

A PRIS EN CHARGE L'ANALYSE DE LA SÉRIE DÉSAISONNALISÉE, EN IDENTIFIANT LES TENDANCES GÉNÉRALES ET LES ANOMALIES.

A RÉALISÉ LES PRÉVISIONS EN UTILISANT LES MODÈLES ETS ET RÉGRESSION LINÉAIRE, ET COMPARÉ LEURS PERFORMANCES.

A RÉDIGÉ LA SECTION SUR L'ÉVALUATION DES MODÈLES AVEC L'EQM, EN EXPLIQUANT POURQUOI LE MODÈLE ARMA ÉTAIT LE PLUS PERFORMANT.

COLLABORATION COMMUNE :

NOUS AVONS RÉDIGÉ ENSEMBLE L'INTRODUCTION, LA CONCLUSION, ET LE RÉSUMÉ EN ANGLAIS POUR ASSURER UNE COHÉRENCE GLOBALE DU RAPPORT.

LA VALIDATION FINALE DES GRAPHIQUES ET DES ANALYSES A ÉTÉ EFFECTUÉE EN ÉQUIPE POUR GARANTIR L'EXACTITUDE ET LA QUALITÉ DU CONTENU.