



BUT  
Science des  
données



# Rapport SAE 3.03 : Analyse de l'Orbite de l'Astéroïde 1221 Amor

Jonathan WILD

13/11/2025

## TABLE DES MATIÈRES

INTRODUCTION.....	2
PRÉSENTATION DES DONNÉES .....	2
DÉCOMPOSITION D'UNE SÉRIE TEMPORELLE.....	4
COMPARAISON DE MODÈLES .....	10
PREDICTION.....	13
CONCLUSION GENERALE.....	15

## INTRODUCTION

Dans le cadre de cette SAE, j'ai choisi d'étudier l'évolution de la distance entre la Terre et l'astéroïde (1221) Amor. Passionné par l'espace, j'ai voulu comprendre comment cet astéroïde se déplace au fil du. Bien que son orbite soit considérée comme instable — au sens où à chaque cycle l'astéroïde s'écarte un peu plus de la Terre — Amor suit néanmoins un schéma orbital répétitif. Sa distance par rapport à la Terre varie donc d'un cycle à l'autre, mais selon une structure régulière qui permet d'identifier des motifs récurrents.

Cette particularité rend l'astéroïde particulièrement intéressant pour une analyse statistique. Les fluctuations de distance qu'il présente peuvent en effet être décrites par un modèle additif, combinant une tendance globale, une composante périodique liée à son orbite, et un reste correspondant aux variations plus irrégulières.

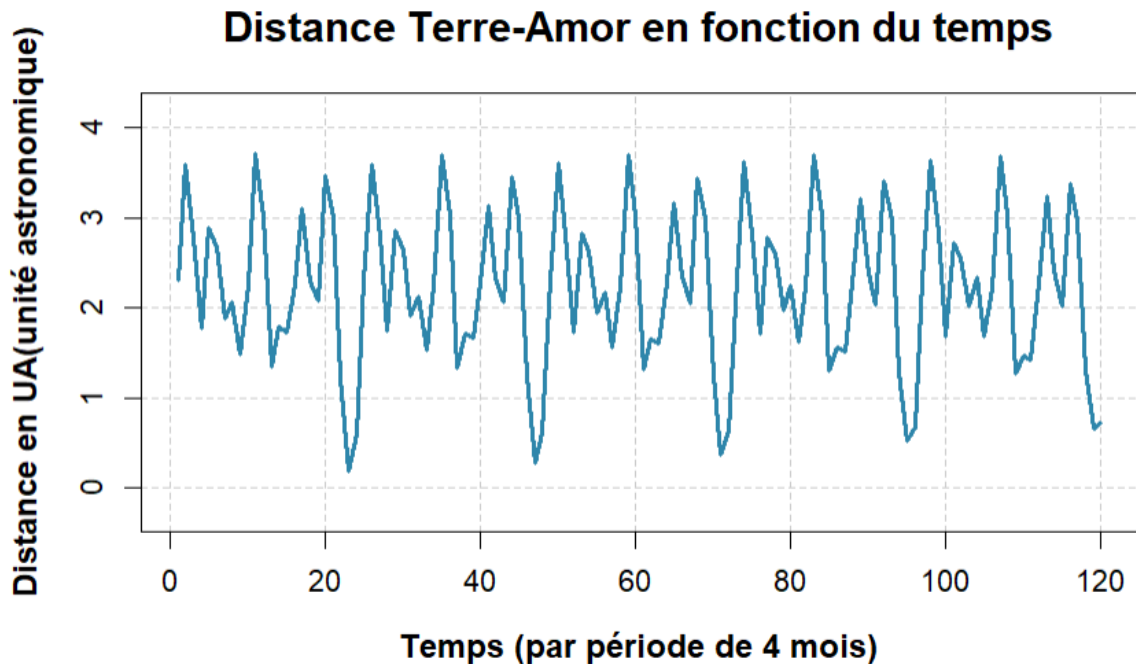
Problématique : Dans quelle mesure la décomposition statistique en modèle additif permet-elle de prédire l'évolution complexe de la distance entre la Terre et l'astéroïde Amor ?

## PRÉSENTATION DES DONNÉES

Les données exploitées dans ce projet correspondent à l'évolution de la distance entre la Terre et l'astéroïde Amor (1221 Amor) au cours du temps. J'ai récupéré ces données sur le site de la NASA. Le fichier de données contient donc les données suivantes :

- La distance entre la Terre et l'astéroïde Amor : L'astéroïde ayant une orbite quasiment stable, ce dernier viens s'éloigné un petit peu de la terre à chaque cycle ce qui nous permet d'y appliquer un modèle additif. Elle est exprimée en UA (Unité Astronomique)  $1 \text{ UA} \approx 149\,597\,870 \text{ km}$ .
- Le temps, en effet les données sont exprimées en fonctions du temps (j'ai récupéré une donnée tous les 4 mois afin d'avoir un cycle sur mes données de 24 données et non plus 96 car c'était beaucoup trop élevé). En effet l'astéroïde ayant une période de révolution de 8 ans cela représente  $24 \times 4$  mois, cette périodicité sur 24 données nous permettras donc d'appliqué les tests même celui de Holt-Winters, ce qui n'aurait pas était le cas avec une période de 96.

On peut donc à partir de ces données faire un graphique représentant la distance entre l'astéroïde et la Terre au cours du temps ce qui nous donne le graphique suivant :



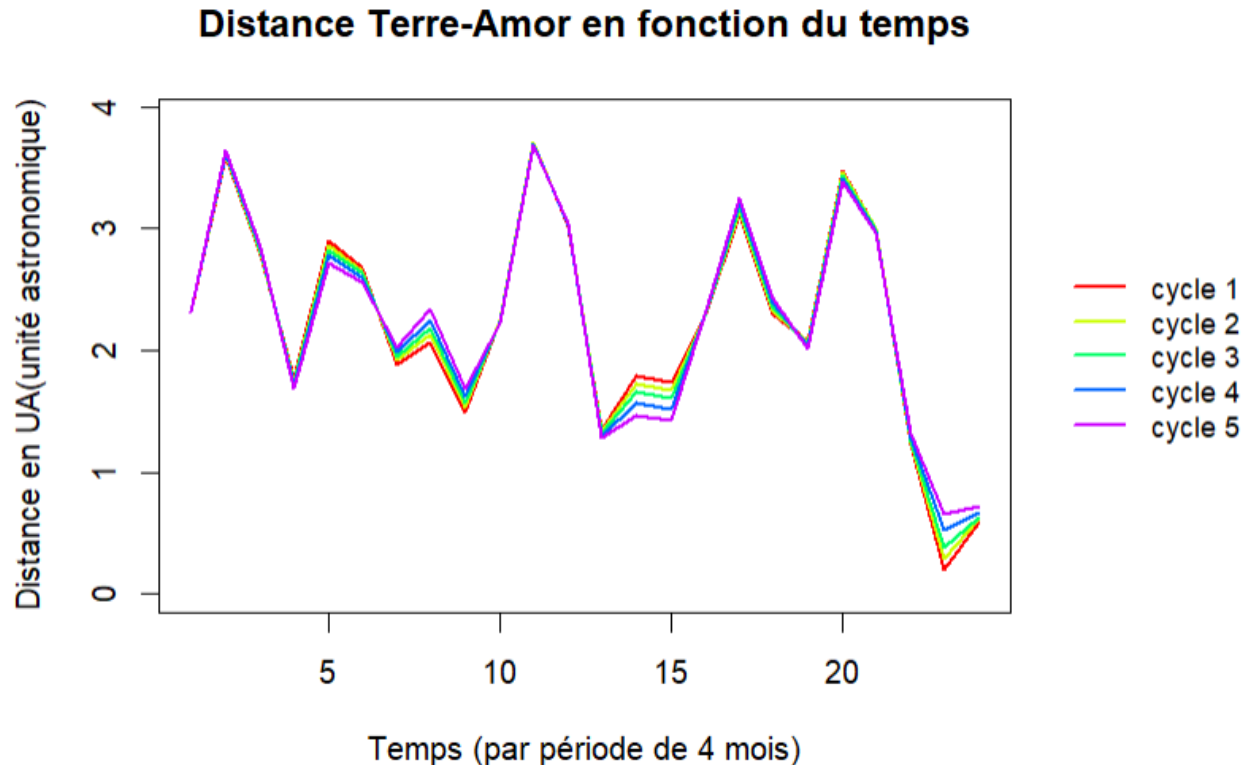
Grâce à ce graphique, on peut dire que ce modèle est additif car l'amplitude des variations est constante :

- Les pics montent régulièrement vers 3.5 - 3.8 UA.
- Les creux descendent régulièrement vers 0.5 - 1.5 UA.

L'écart global (l'enveloppe du graphique) reste sensiblement le même. Le graphique ressemble à une "bande" horizontale plutôt qu'à un entonnoir ou un cône. Il est aussi possible de confirmer cela de manière mathématique car les distances entre corps célestes suivent des cycles périodiques basés sur des orbites elliptiques. Mathématiquement, ces mouvements se décrivent souvent par des sommes de fonctions périodiques (sinus/cosinus), ce qui est par nature une structure additive.

## DÉCOMPOSITION D'UNE SÉRIE TEMPORELLE

Afin d'analyser les données, je vais donc décomposer les données en commençant par analysé les cycles de la distance entre la terre et l'astéroïde.



Ce graphique illustre la saisonnalité de nos données en superposant les 5 derniers cycles observés. L'analyse visuelle révèle que l'orbite de l'astéroïde conserve une forme globale similaire, mais présente un léger décalage à chaque itération.

En effet, les courbes ne se superposent pas parfaitement : on observe une variation au niveau des extremums d'un cycle à l'autre. Cette fluctuation, semble indiquer une légère tendance, ce qui pourrait compliquer la modélisation du fait que les écarts entre les cycles ne sont pas constants au cours des cycles.

L'œil humain a du mal à voir si la distance globale augmente ou diminue juste en regardant les pics bleus. C'est pour cela que le graphique suivant va analyser via la moyenne mobile l'évolution de notre jeu de données.

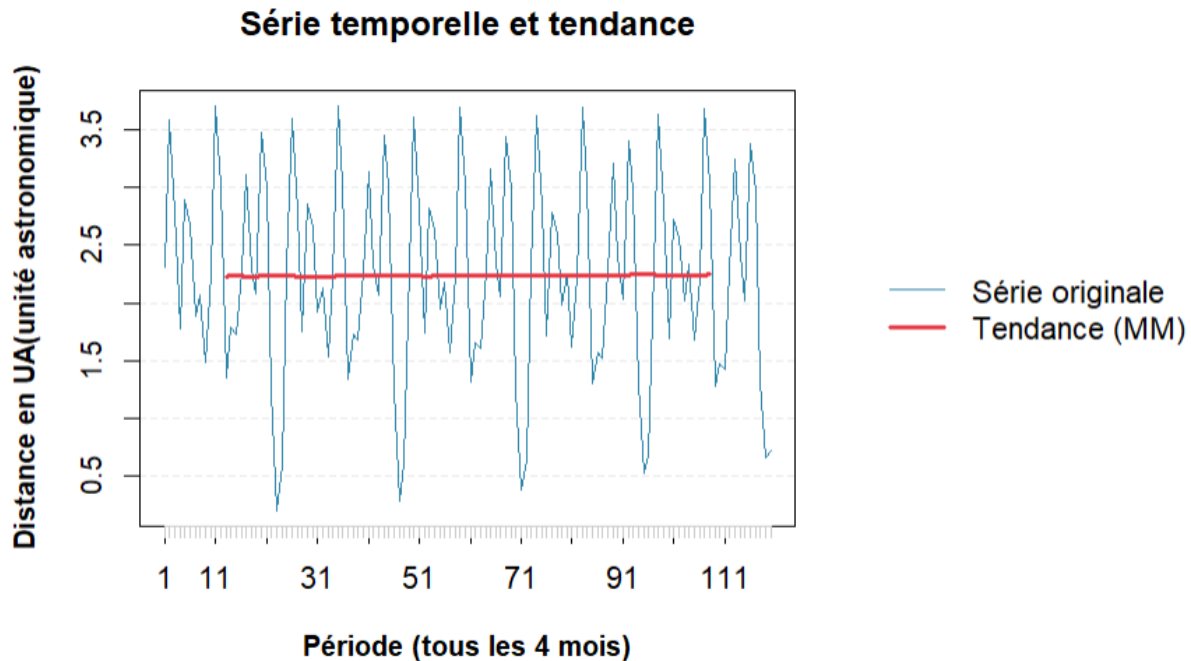
## 1. Tableau des Coefficients Saisonniers

### Décomposition Saisonnière de l'Astéroïde (1221) Amor

Phase du Cycle (t)	Coeff. Saisonnier (UA)	Interprétation Physique
1	+0.05	Position neutre
2	+0.55	Début d'éloignement
3	+1.35	Pic local (Éloignement)
4	+0.60	Rapprochement relatif
5	-0.50	Passage sous la moyenne
6	+0.55	Remontée
7	+0.40	Éloignement modéré
8	-0.25	Léger rapprochement
9	-0.05	Point neutre
10	-0.65	Creux modéré
11	+0.00	Passage à zéro
12	+1.45	PIC MAJEUR (Aphélie)
13	+0.80	Chute rapide
14	-0.90	Creux intermédiaire
15	-0.60	Stabilisation basse
16	+0.05	Retour à l'équilibre
17	+0.90	Pic secondaire
18	+0.10	Léger repli
19	-0.20	Léger creux
20	+1.20	Pic tertiaire
21	+0.75	Rapprochement
22	-1.00	Chute brutale
23	-1.90	CREUX ABSOLU (Périhélie)
24	-1.60	Remontée finale

On constate une asymétrie : les creux (rapprochements) sont plus profonds et brutaux (période 23) que les pics (éloignements) qui sont plus étalés, reflétant la nature elliptique de l'orbite.

## 2. Série Corrigée des Variations Saisonnnières (CVS)

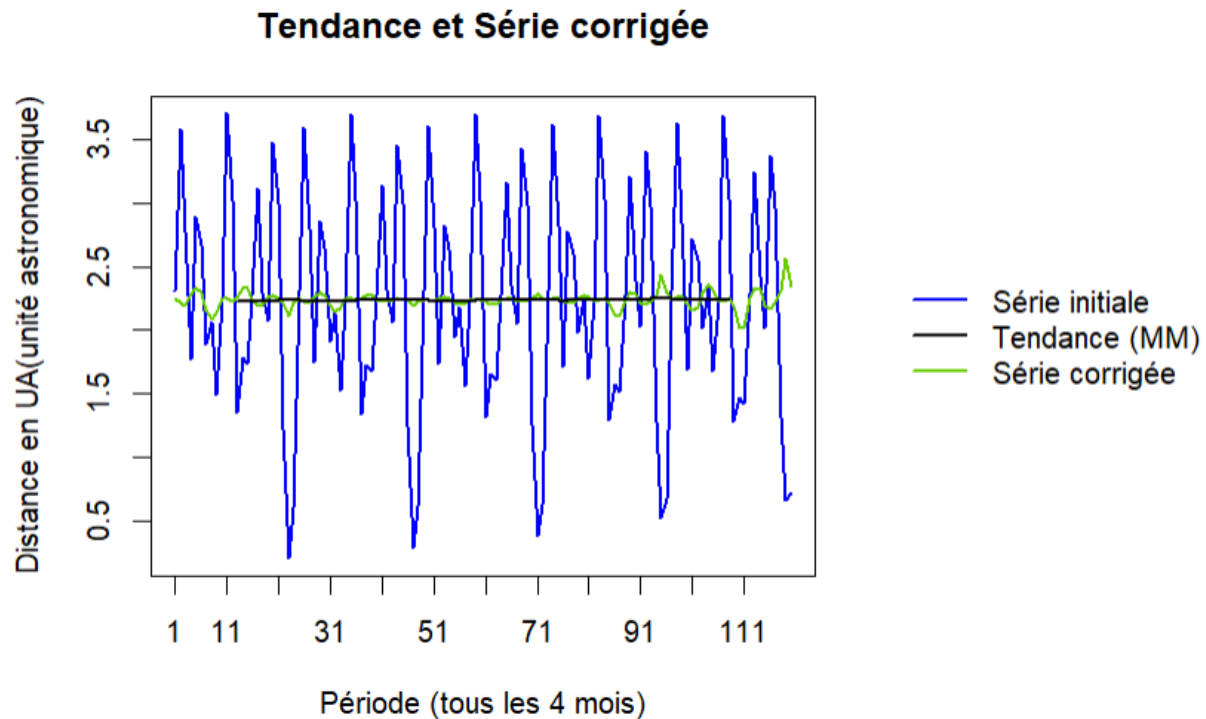


Le graphique ci-après superpose l'évolution brute de la distance Terre-Amor (courbe bleue) et sa tendance de fond (courbe rouge). L'axe des ordonnées représente la distance en Unités Astronomiques (UA), tandis que l'axe des abscisses correspond à la chronologie des observations, échantillonnées par périodes de 4 mois.

La courbe bleue (« Série originale ») met en évidence la forte périodicité du phénomène : on observe des oscillations de grande amplitude, alternant entre des rapprochements significatifs (les creux, descendant sous 0,5 UA) et des éloignements maximaux (les pics, atteignant environ 3,5 UA). Ces fluctuations correspondent aux cycles synodiques de l'astéroïde.

La courbe rouge représente la tendance, calculée ici par Moyenne Mobile (MM). Contrairement aux variations cycliques rapides, cette ligne apparaît relativement horizontale et stable autour de 2,2 - 2,3 UA sur la période observée. Cela indique que, malgré les variations extrêmes de distance dues à l'orbite, la distance moyenne entre la Terre et l'astéroïde reste quasi-stationnaire sur cette fenêtre temporelle, la moyenne mobile filtrant efficacement la composante saisonnière.

Notre Moyenne mobile semble très plate c'est pour cela que l'on fait la série corrigée qui est sensé être plus proche des données réelles. (Graphique ci-dessous).

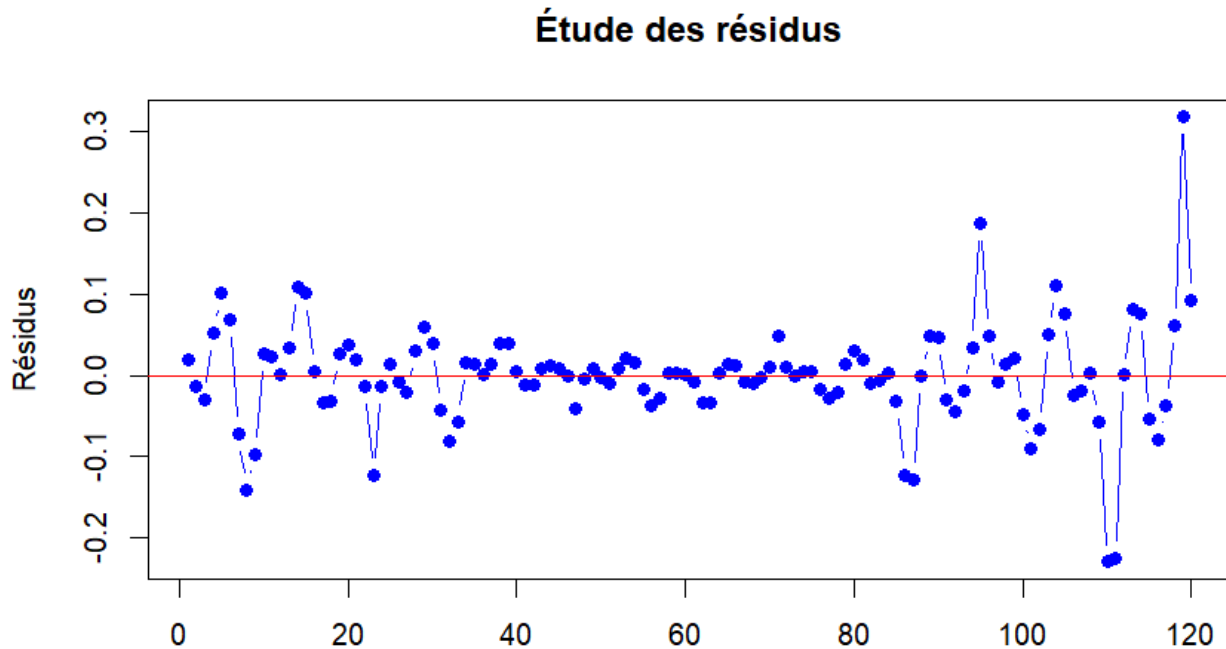


L'examen de la courbe verte (**Série corrigée**) par rapport à la courbe bleue (**Série initiale**) permet d'apprécier l'efficacité de la désaisonnalisation. Alors que la série brute présente des oscillations violentes comprises entre 0,5 et 3,8 UA, la série CVS apparaît considérablement "aplatie", oscillant beaucoup plus faiblement autour de la tendance (ligne noire).

Cependant, la correction n'est pas parfaite. On observe sur la courbe verte une agitation résiduelle (des pics et des creux persistants), particulièrement visible sur la fin de la série temporelle. Cela confirme l'observation faite lors de l'étude des résidus : la saisonnalité de l'astéroïde Amor n'est pas strictement constante sur la durée<sup>5</sup>. Bien que la série CVS permette de mieux visualiser la trajectoire moyenne en éliminant les variations orbitales majeures, elle laisse apparaître les irrégularités (instabilités) propres à ce corps céleste, prouvant que le mouvement de l'astéroïde ne se résume pas à une simple répétition mécanique parfait.

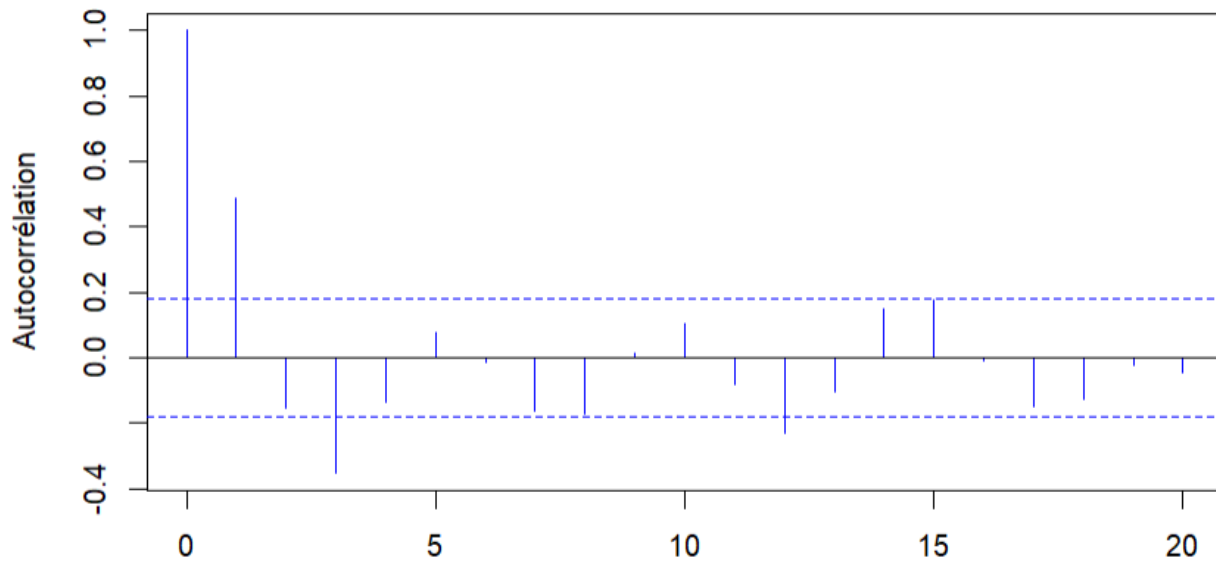


### 3. Étude des résidus



L'étude graphique des résidus permet de qualifier la pertinence du modèle additif utilisé : si la distribution des points s'organise globalement autour de l'axe zéro, attestant de l'absence de biais systématique dans l'estimation des distances, elle révèle néanmoins une limite structurelle importante caractérisée par une forte hétéroscédasticité. En effet, alors que le modèle présente une excellente précision sur la plage médiane (périodes 30 à 80) avec des écarts résiduels minimes, on observe une dégradation nette de la performance sur le dernier tiers de la série temporelle, où l'amplitude des erreurs augmente considérablement. Cette "explosion" de la variance en fin de période indique que la supposition d'une saisonnalité strictement constante ne suffit plus à décrire parfaitement la réalité physique sur le long terme, trahissant ainsi les instabilités orbitales propres à l'astéroïde Amor qui finissent par générer un décalage croissant entre le schéma théorique répétitif et la trajectoire réelle observée.

### Corrélogramme des résidus relatifs



L'analyse de ce graphique permet de vérifier si les erreurs commises par le modèle sont dues au pur hasard, ce qui serait le signe d'un résultat idéal. Cependant, le fait que certaines barres bleues dépassent les limites en pointillés indique que ce n'est pas le cas ici : les erreurs ne sont pas isolées mais liées entre elles, c'est-à-dire que si le modèle se trompe à un instant donné, il a tendance à reproduire une erreur similaire juste après. Cette "mémoire" dans les erreurs prouve qu'il reste encore une petite partie du comportement de l'astéroïde que la formule mathématique actuelle n'a pas réussi à deviner, laissant échapper une information cachée dans les mouvements de l'objet céleste.

On peut donc en conclure que notre modèle est mauvais, pour s'assurer de cela et essayer de prédire aux mieux nos futures données, il va donc falloir essayer d'autres modèles afin d'en trouver un meilleur. C'est ce que nous allons faire dans la partie suivante.

## COMPARAISON DE MODÈLES

### 1. Le Lissage Exponentiel Simple (Modèle ANN)

C'est le modèle le plus basique.

Le principe : Il calcule une moyenne pondérée des données passées, en donnant plus d'importance aux observations récentes qu'aux anciennes. Il considère que la série est **stationnaire** (elle ne va ni vers le haut, ni vers le bas sur le long terme).

Ce qu'il détecte : Il ne détecte ni tendance (montée/baisse globale), ni saisonnalité (cycles). Il voit juste un niveau moyen qui fluctue un peu.

Sa prévision : Une ligne plate horizontale.

Pour l'astéroïde Amor : C'est un mauvais modèle car il ignore que l'astéroïde tourne (cycles) et s'éloigne (tendance). Il servira juste de point de comparaison (le "niveau zéro").

### 2. Le Lissage de Holt (Modèle AAN)

Le principe : Il reprend le lissage simple mais ajoute une seconde équation pour capter la tendance (la pente). Il est capable de comprendre si les données augmentent ou diminuent structurellement.

Ce qu'il "voit" : Il détecte le niveau moyen et la direction globale (l'éloignement séculaire de l'astéroïde), mais il ignore totalement les oscillations répétitives.

Sa prévision : Une ligne droite oblique (qui continue de monter ou descendre à l'infini).

Pour l'astéroïde Amor : Il est mieux que le simple car il comprend que l'astéroïde s'éloigne, mais il est incomplet car il "coupe tout droit" à travers les cycles orbitaux sans les reproduire.

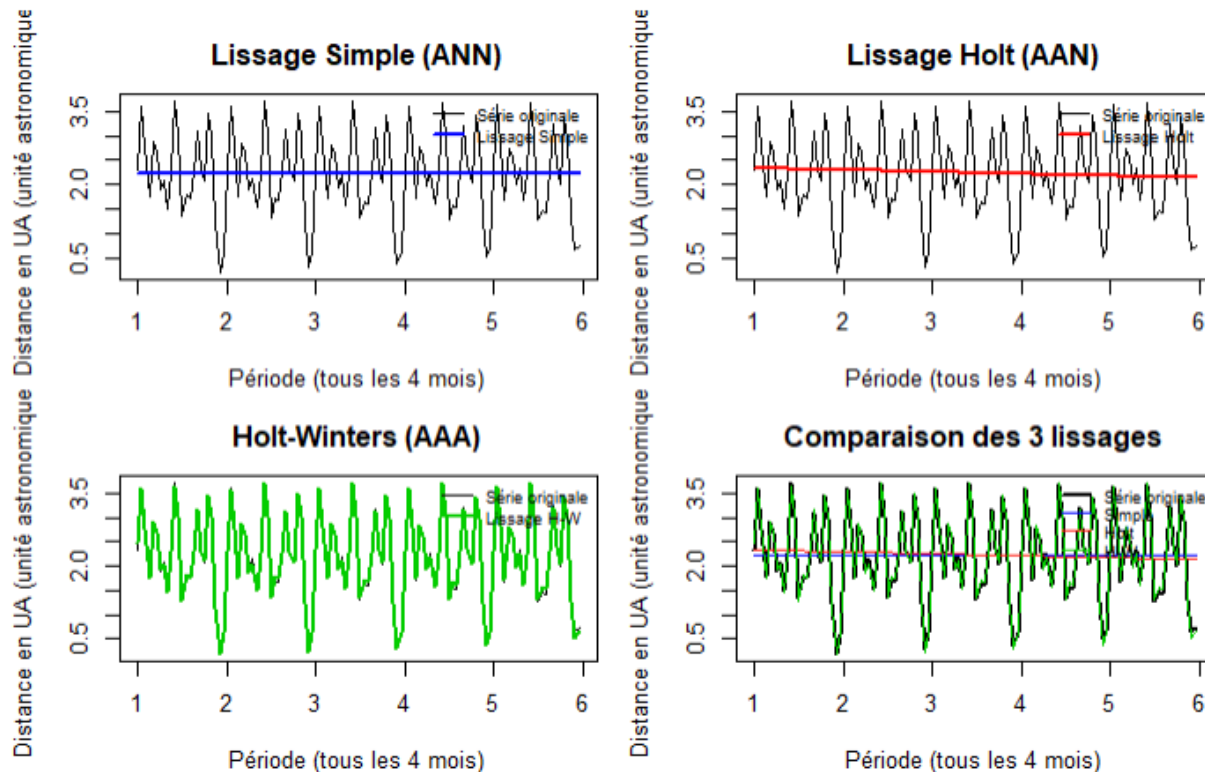
### 3. Le Lissage de Holt-Winters (Modèle AAA)

Le principe : C'est le modèle le plus abouti pour ce type de données. Il ajoute une troisième équation pour capter la saisonnalité (le motif qui se répète).

Ce qu'il "voit" : Il décompose le signal en trois parties : le Niveau (base), la Tendance (évolution long terme) et la Saisonnalité (le cycle orbital de 8 ans/24 périodes).

Sa prévision : Une courbe qui reproduit les vagues (pics et creux) tout en suivant la pente générale.

Pour l'astéroïde Amor : C'est théoriquement le meilleur modèle. Il colle à la réalité physique : il prend en compte l'éloignement progressif (Tendance) tout en respectant la mécanique céleste des retours périodiques (Saisonnalité).



Le graphique présente les résultats de l'ajustement de trois modèles de lissage exponentiel.

**1. Lissage Simple (ANN - courbe bleue)** Le graphique en haut à gauche montre le résultat du lissage exponentiel simple. La courbe bleue forme une **ligne horizontale stricte** (constante).

Interprétation : Ce modèle suppose que la série est stationnaire (sans tendance ni cycle) et que les variations observées ne sont que du "bruit". Il se contente de calculer une moyenne locale.

Verdict : Il est inadapté pour décrire une orbite, car il échoue totalement à capturer les variations cycliques de la distance.

**2. Lissage de Holt (AAN - courbe rouge)** Le graphique en haut à droite présente le modèle de Holt (avec tendance). La courbe rouge est une ligne droite, très légèrement inclinée.

Interprétation : Ce modèle détecte la tendance de fond (l'éloignement ou le rapprochement séculaire moyen), mais il reste "aveugle" aux oscillations. Il traverse la série temporelle en son centre sans suivre les pics et les creux.

Verdict : Bien qu'il capte la direction générale, il est insuffisant car il ignore la composante principale du phénomène : la périodicité.

**3. Lissage de Holt-Winters (AAA - courbe verte)** Le graphique en bas à gauche montre le modèle de Holt-Winters (avec tendance et saisonnalité). La courbe verte se superpose presque parfaitement à la courbe noire (données réelles).

Interprétation : Grâce à sa composante saisonnière (le troisième paramètre du modèle), cet algorithme a "appris" le motif répétitif de l'orbite. Il parvient à reproduire fidèlement l'amplitude des rapprochements et des éloignements à chaque période.

Verdict : C'est le seul modèle capable de décrire la réalité physique du mouvement de l'astéroïde.

**4. Synthèse :** Le dernier graphique (bas droite) superpose les trois approches et confirme visuellement la supériorité du modèle Holt-Winters. Alors que les modèles Simple (bleu) et Holt (rouge) restent passifs face aux variations, le modèle Holt-Winters (vert) colle à la dynamique de la série. On peut vérifier cela en analysant les données fournies par le modèle :

Performances des modèles de lissage				
	AIC	MSE	RMSE	MAPE
Simple	544.094	0.738	0.859	58.117
Holt	547.727	0.736	0.858	58.237
HoltWinters	-104.209	0.002	0.046	2.531

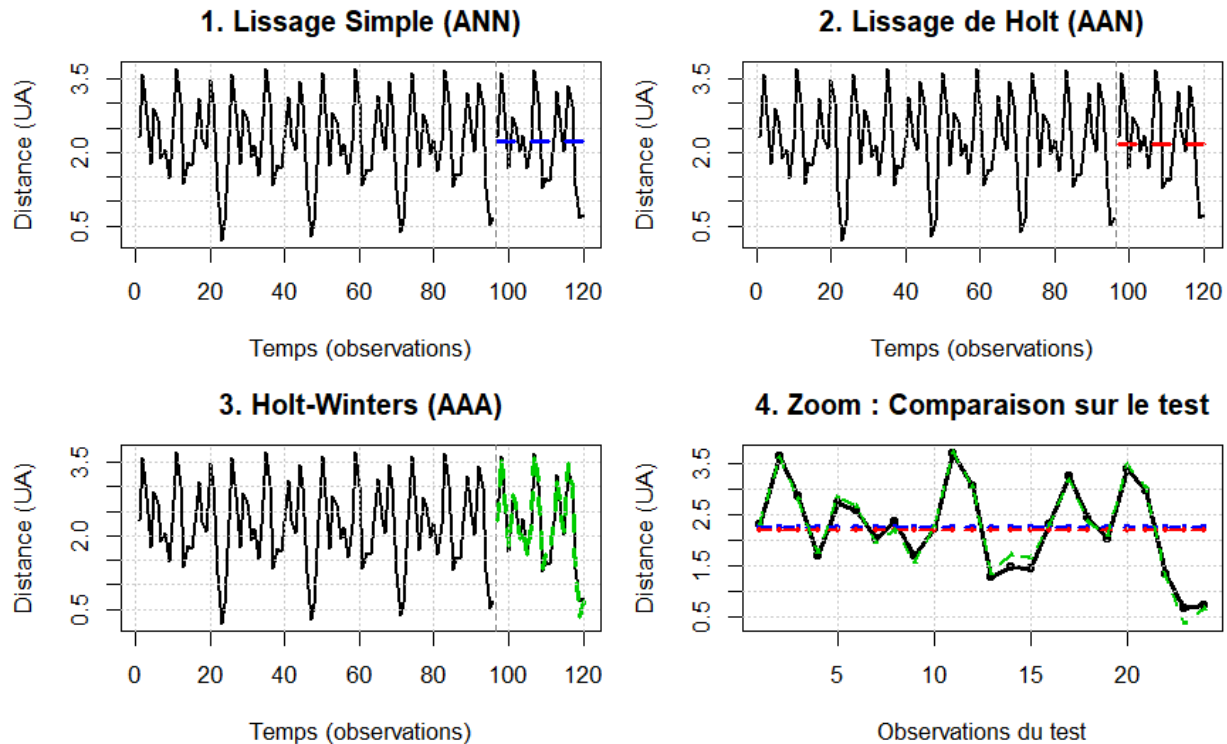
Ce tableau nous confirme bien que le modèle Holt-Winters est le plus performant, car il minimise drastiquement l'ensemble des critères d'erreur par rapport aux modèles Simple et Holt :

- **Le critère AIC :** Il est le seul à être négatif (**-104.209**), ce qui est infiniment meilleur que les scores élevés (> 540) des deux autres modèles. En statistiques, plus l'AIC est faible, meilleur est le modèle.
- **La précision (MSE/RMSE) :** L'erreur quadratique moyenne (MSE) est quasi-nulle (**0.002**), contre environ 0.74 pour les autres. Cela signifie que la distance entre la courbe de prévision et la courbe réelle est infime.
- **L'erreur relative (MAPE) :** C'est l'indicateur le plus parlant. Le modèle Holt-Winters ne commet qu'une erreur moyenne de **2.53%**, alors que les modèles Simple et Holt se trompent en moyenne de **58%**.

**Conclusion :** L'analyse graphique valide le choix du modèle **Holt-Winters (AAA)**. La distance Terre-Amor étant un phénomène intrinsèquement saisonnier (dû à la mécanique céleste), seuls les modèles intégrant cette saisonnalité peuvent fournir un ajustement précis et des prévisions fiables.

## PREDICTION

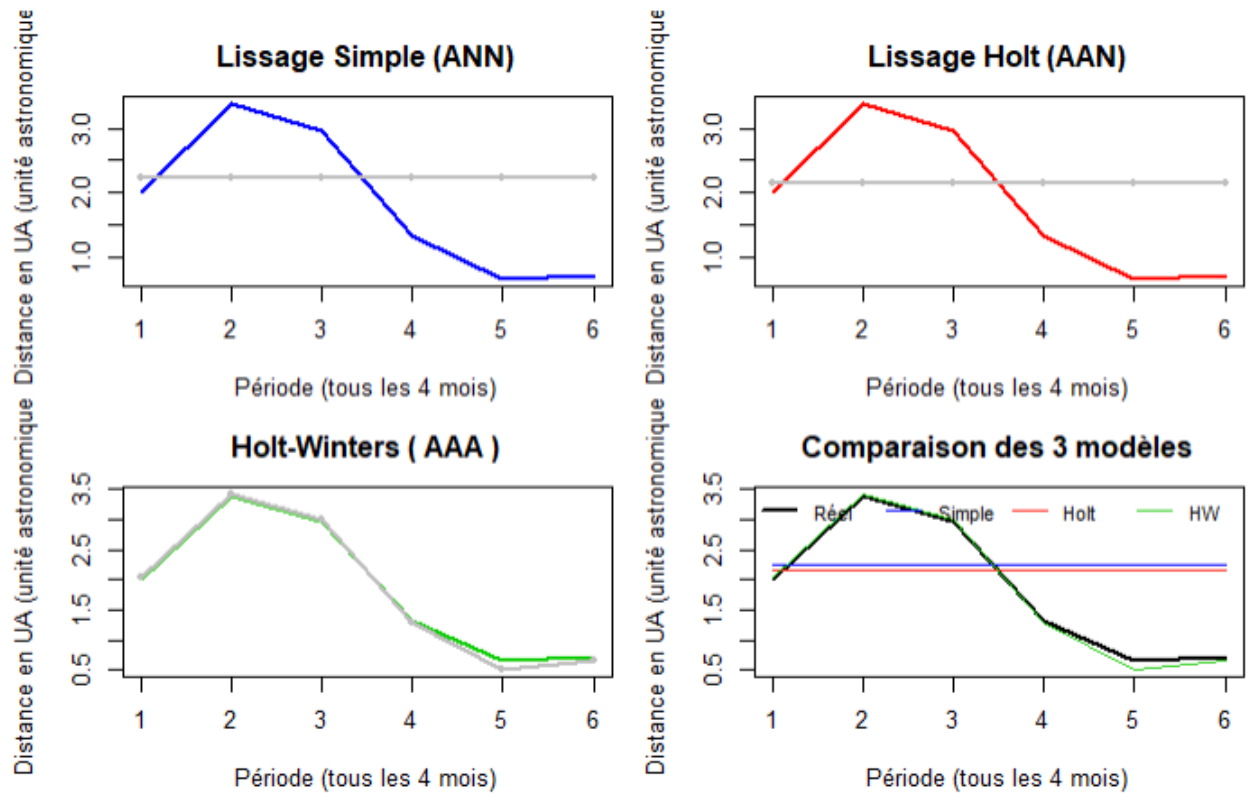
Au début de ce rapport, nous avons mis de cotés le dernier cycle de nos données afin de pouvoir comparer nos prédictions, c'est ce que nous allons maintenant faire.



Conformément à notre méthodologie, nous avons isolé le dernier cycle complet de données (le jeu de donnée, correspondant aux observations 100 à 120) pour vérifier la capacité prédictive de nos modèles face à la réalité.

On constate l'échec des modèles linéaires (Graphiques 1 & 2) Les graphiques du Lissage Simple (ANN) et du Lissage de Holt (AAN) confirment leur inadaptation pour ce type de données. Leurs prévisions (traits bleu et rouge) sont des lignes droites qui traversent le cycle sans en suivre les variations.

Cependant on constate aussi le succès du modèle de Holt-Winters (AAA) qui démontre une excellente performance. La courbe de prévision (en vert) ne se contente pas de suivre une moyenne, elle reproduit la dynamique oscillatoire. Le modèle a compris que le motif observé dans le passé allait se reproduire dans le futur. C'est encore plus flagrant lorsque l'on se concentre uniquement sur la prédiction :



Conclusion : Ce test visuel est sans appel : seul le modèle Holt-Winters est valide pour effectuer des prévisions astronomiques sur l'astéroïde Amor. Il capture la saisonnalité essentielle du phénomène, là où les autres modèles échouent par excès de simplification.

## CONCLUSION GENERALE

Ce projet avait pour objectif d'analyser et de modéliser l'évolution de la distance entre la Terre et l'astéroïde (1221) Amor, un corps céleste présentant une orbite instable mais suivant un schéma répétitif. L'exploitation des données de la NASA, échantillonnées sur une période de révolution de 8 ans, a permis de tester la pertinence d'une décomposition statistique en modèle additif.

L'analyse préliminaire de la série temporelle a révélé une structure saisonnière marquée, caractérisée par des oscillations régulières entre 0,5 et 3,8 UA. Bien que la désaisonnalisation ait permis d'isoler une tendance moyenne quasi-stationnaire autour de 2,3 UA, l'étude des résidus a mis en lumière une limite structurelle : une forte hétéroscédasticité en fin de période, trahissant les instabilités orbitales propres à cet astéroïde qui ne peuvent être parfaitement décrites par une saisonnalité strictement constante

La comparaison des modèles de prévision a cependant livré un verdict sans appel :

- **L'échec des modèles naïfs** : Le Lissage Simple (ANN) et le Lissage de Holt (AAN) se sont révélés inadaptés, produisant des prévisions linéaires incapables de capturer la dynamique cyclique de l'orbite.
- **La supériorité de Holt-Winters (AAA)** : Ce modèle s'est imposé comme le seul capable de reproduire fidèlement la réalité physique et les variations d'amplitude des cycles.

Du fait de du caractère réguliers de nos données on peut donc considérer que si ce modèle est efficace pour des données actuelles il le sera pour des données futures car elles varient très peu, cela peut donc nous permettre de modéliser plusieurs cycles en avance afin de connaître la distance de l'astéroïde :

