



*12-11-2021*

---

## Rapport d'avancement

Tentative d'apprentissage des corrélations entre phases de la lune et du soleil et coefficients des marées

---

*Auteurs :*

Otmane EL ALOI, Abdelhafid SAOUD, Abdelhadi ZIANE

*Encadrant :*

Mr. Olivier ROUX

Option :  
Informatique

## Table des matières

|     |  |   |
|-----|--|---|
| 1   | Recherche exhaustive des paramètres du modèle des forêts aléatoires . .  | 2 |
| 2   | Recherche du meilleur nombre d'échantillon des anciennes valeurs à utiliser pour faire la prédiction . . . . . | 3 |
| 3   | Premier pas de conception de l'interface Web . . . . .   | 4 |
| 4   | Planning . . . . .   | 6 |
| 4.1 | Planning de la semaine en cours . . . . .  | 6 |
| 4.2 | Planning de la semaine prochaine . . . . .   | 7 |

## Avancement

Au cours de cette semaine, nous avons continué l'entraînement de nos modèles.

Nous décrivons dans les sections qui suivent les résultats des expériences effectuées jusqu'au présent.

## 1 Recherche exhaustive des paramètres du modèle des forêts aléatoires

Le but de cette expérience est d'étudier l'effet du changement des paramètres sur la performance du modèle. Pour comparer les différentes variantes du modèle de forêts aléatoires. Nous utilisons la racine de l'erreur quadratique moyenne pour chaque  $j$  ème valeur prédite :

$$RMSE_j(Y_{pred}, Y_{reel})$$

Le schéma ci-dessous explique comment on calcule les erreurs de prédiction :

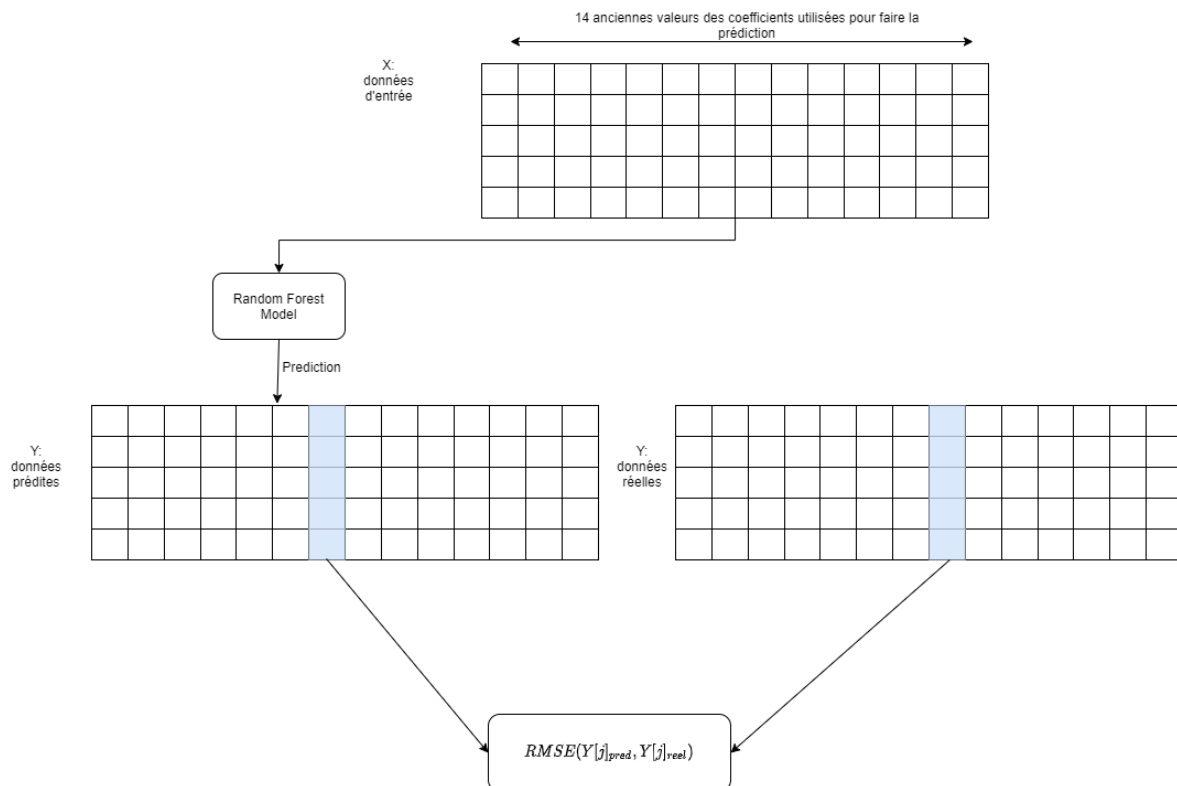


FIGURE 1 – Illustration du calcul de l'erreur pour le  $j$  ème jour de chaque semaine de test.

Dans cette expérience nous avons gardé le nombre des anciennes valeurs fixe : 14 valeurs des coefficients est donnée en entrées au modèle des forêts aléatoire pour prédire

FIGURE 3 – Caption

14 futures valeurs. Nous avons effectués une recherche exhaustive effectuée sur quelques paramètres :

```
param_grid = {"n_estimators": [100, 200, 300, 400, 500],  
              "max_depth": [2, 5, 10, 15],  
              "min_samples_split": [2, 5, 10, 15],  
              "min_samples_leaf": [1, 5, 10, 15]}
```

FIGURE 2 – Exemple de paramètres essayés

On constate que les performance du modèle n'ont pas améliorées.

## 2 Recherche du meilleur nombre d'échantillon des anciennes valeurs à utiliser pour faire la prédiction

Le but de cette expérience est de trouver la meilleur valeur pour le paramètre *past\_value* qui indique le nombre d'ancienne valeurs des coefficients à utiliser pour prédire les coefficients de la semaine suivante.

Pour aujourd'hui nous avons essayé un échantillon de valeur qui s'étend de 14 anciennes valeurs à 120 anciennes valeurs (i.e deux mois d'anciennes valeurs des coefficients).

Nous avons constaté une augmentation remarquable des performance lorsque le nombre d'anciennes valeurs prise en compte augmente, la figure ci-dessous présente l'erreur de 3 modèles sur l'année de test 2021 :

- L'erreur du modèle en vert : 14 anciennes valeurs
- L'erreur du modèle en orange : 60 anciennes valeurs (1 mois)
- L'erreur du modèle en bleu : 120 anciennes valeurs (2 mois)

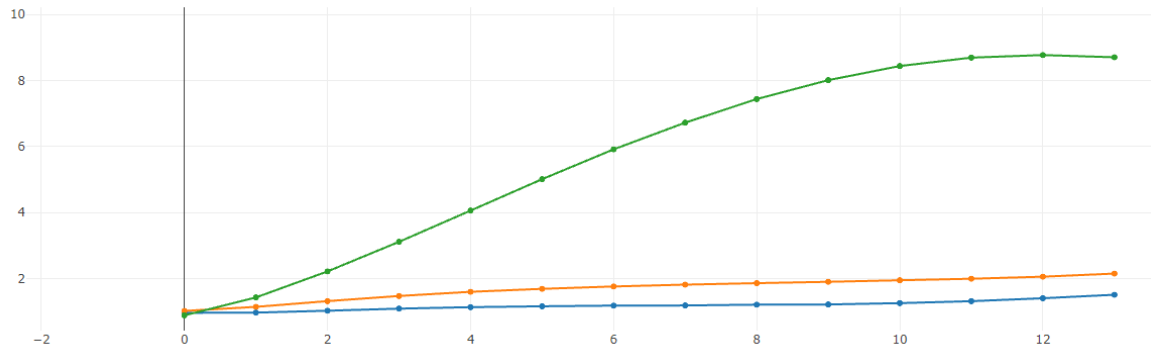


FIGURE 4 – Racine de l'erreur moyenne quadratique pour 14 valeurs prédite

### 3 Premier pas de conception de l'interface Web

Pour cette partie, nous avons commencé à mettre les premières briques de notre interface Web. Cela en commençant par la première page Web qui permettra à l'utilisateur d'interagir avec le système de prédiction. Cette page regroupe les différentes fonctionnalités principales d'un site Web à savoir l'objectif de l'interface Web, méthode de fonctionnement et contact. En plus, elle comporte les fonctionnalités permettant de lancer et visualiser la prédiction.

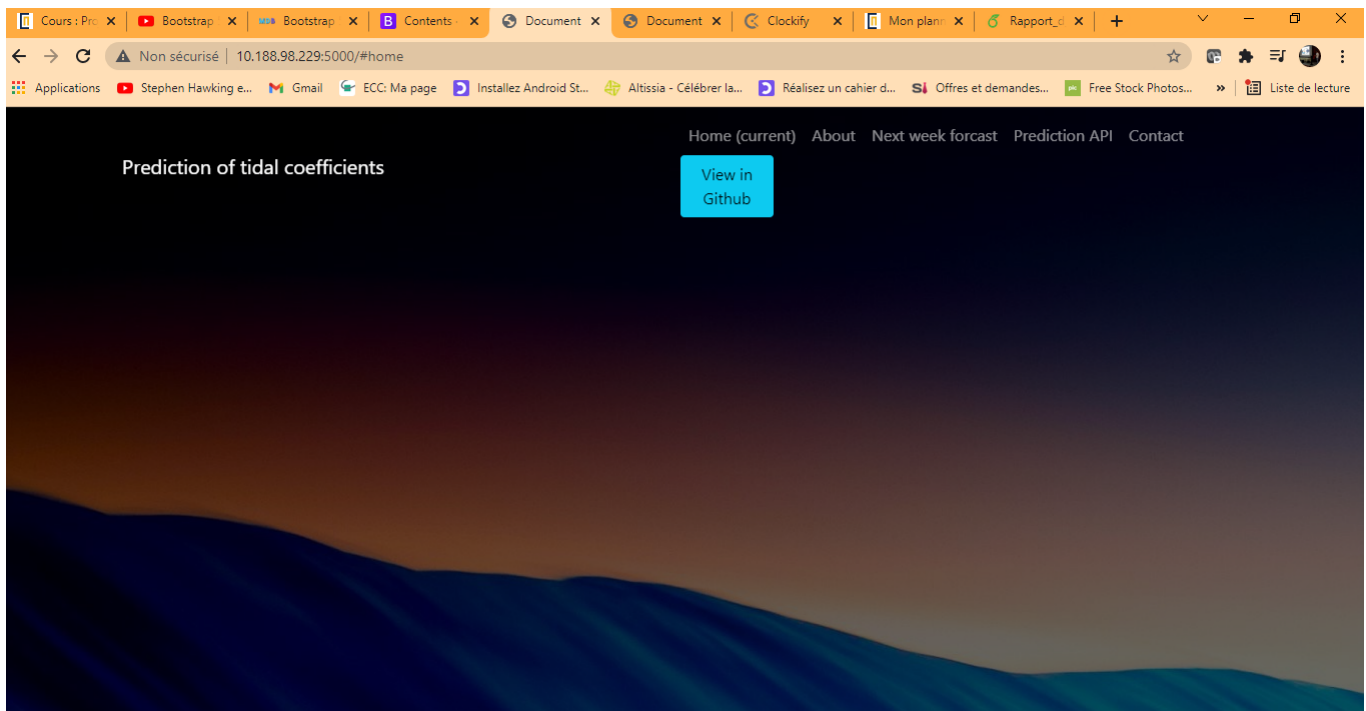


FIGURE 5 – Planning de la semaine en cours

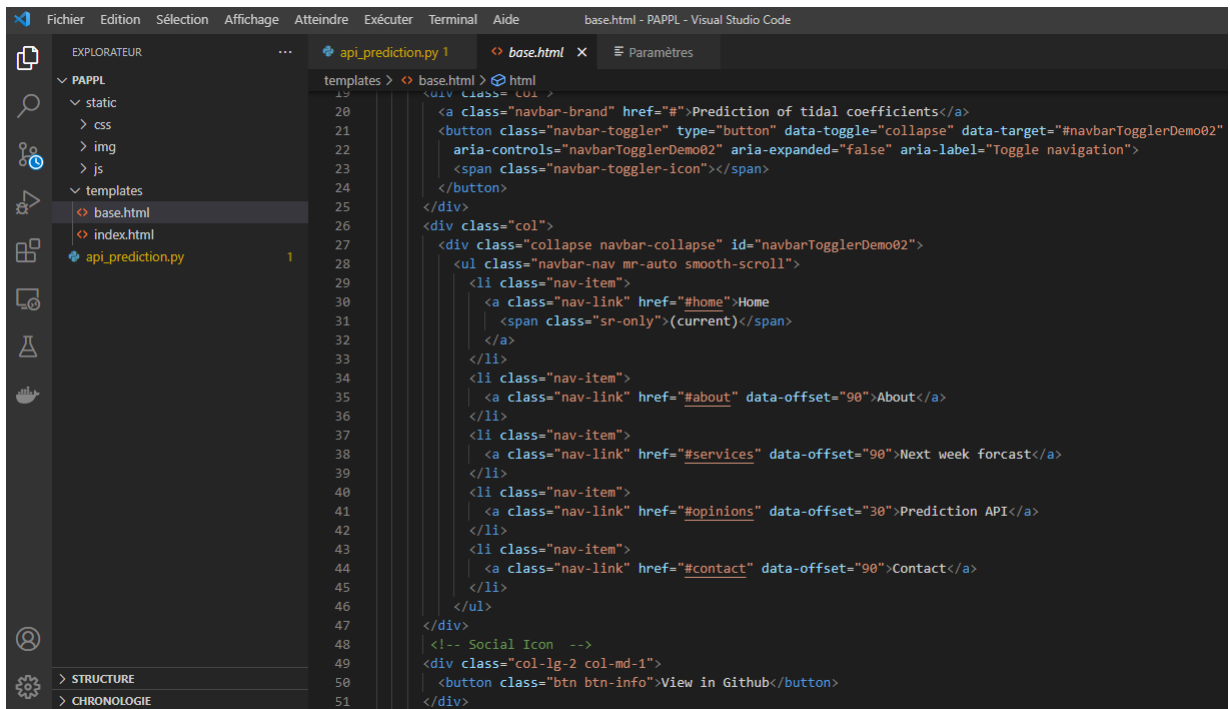


FIGURE 6 – Planning de la semaine en cours

## 4 Planning

### 4.1 Planning de la semaine en cours



FIGURE 7 – Planning de la semaine en cours

## 4.2 Planning de la semaine prochaine



FIGURE 8 – Planning de la semaine prochaine