

# Simulation RF TS

## 1 Simulation pure

Plusieurs modèles:

1. Modèle 1 :

$$X_t = \phi X_{t-1} + a_{t \bmod 12} \epsilon_t \quad (1)$$

où  $\epsilon_t \sim \mathcal{N}(0, 1)$ ,  $a_t = \{1, 1, 1, 2, 3, 1, 1, 1, 1, 2, 4, 6\}$  et  $\phi = \{0.3, 0.5, 0.7, 0.9, 0.95\}$ . Exemple en fig. 1.

2. Modèle 2 :

$$X_t = U_t(t \bmod d + 1) + \sin\left(\frac{2\pi t}{d}\right) \quad (2)$$

où  $U_t \sim \mathcal{U}(0, 1)$  et  $d = \{10, 50, 100\}$ . Exemple en fig. 2.

Training set composé de  $T_{train} = 1000$  observations, test set composé de 250 observations.

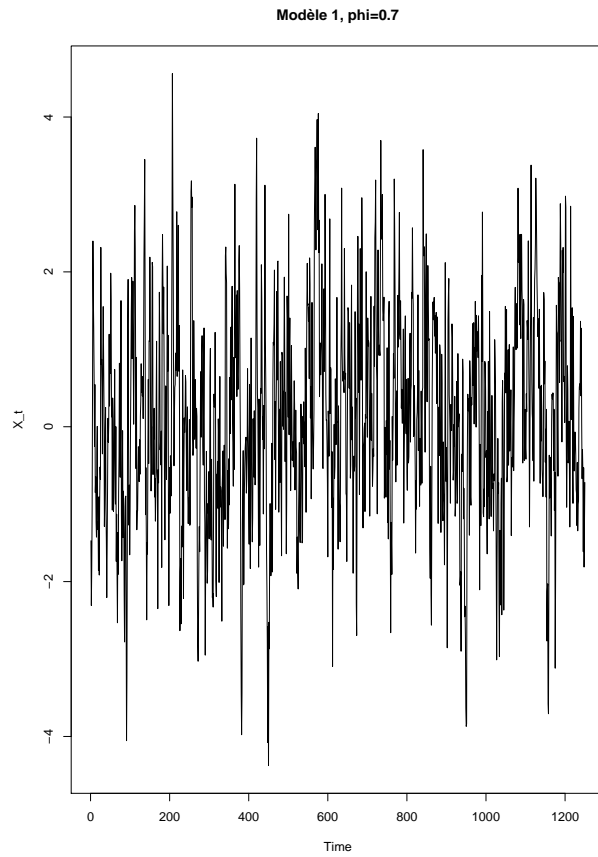


Figure 1: Exemple modèle 1 pour  $\phi = 0.7$ .

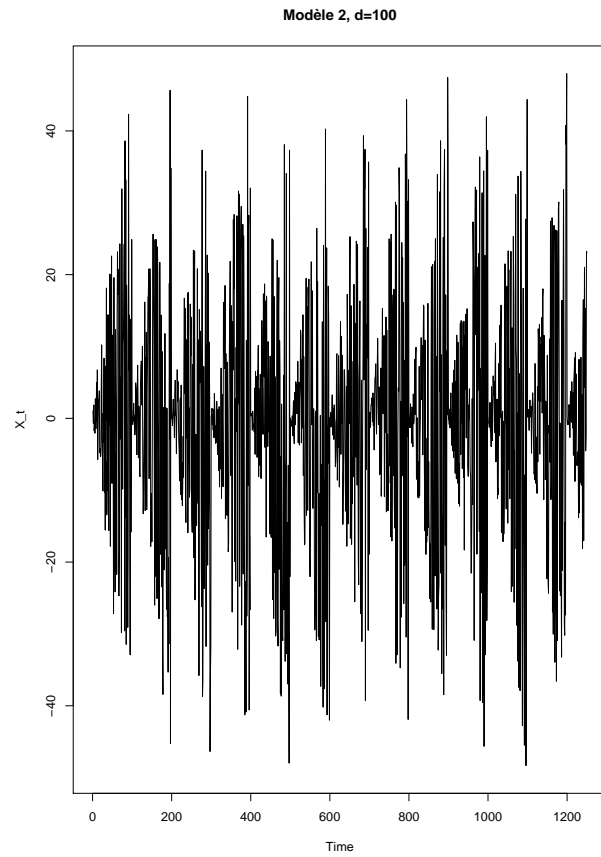


Figure 2: Exemple modèle 2 pour  $d = 100$ .

## 1.1 Graphiques modèle 1

Evolution de l'erreur sur le test set selon la taille du blocs (sur une grille fixée à l'avance). Répété 20 fois. Voir figs. 3 to 7.

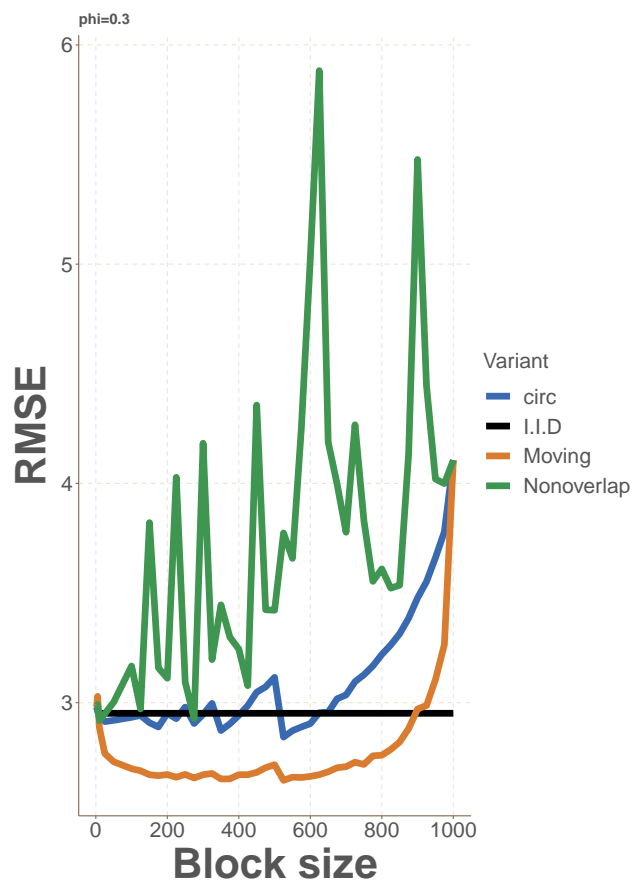


Figure 3: Modèle pour  $\phi = 0.3$ .

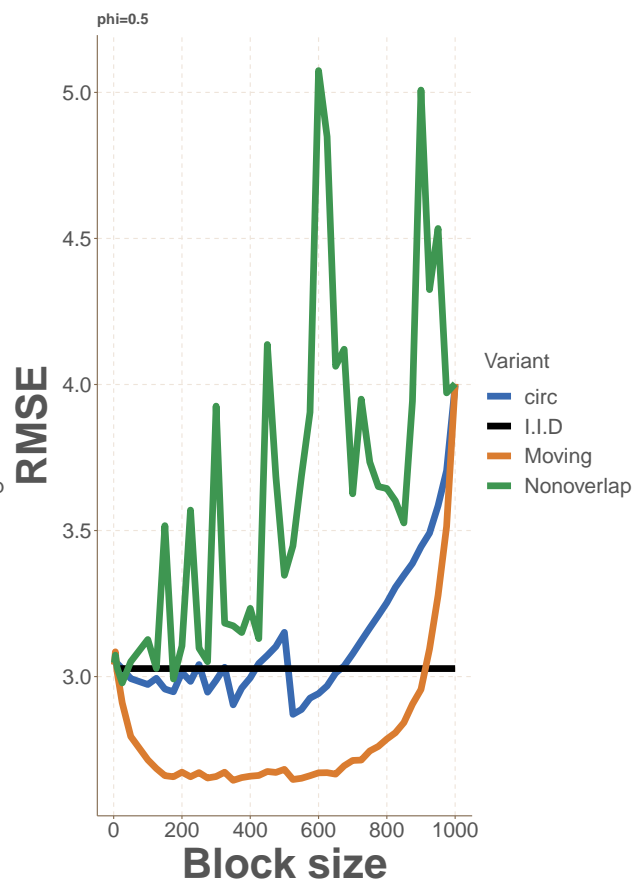


Figure 4: Modèle pour  $\phi = 0.5$ .

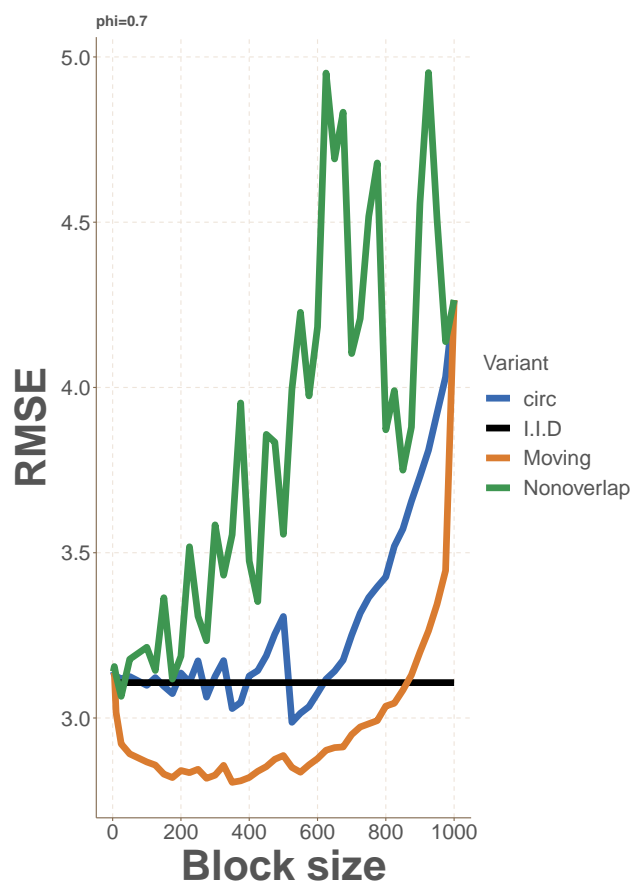


Figure 5: Modèle pour  $\phi = 0.7$ .

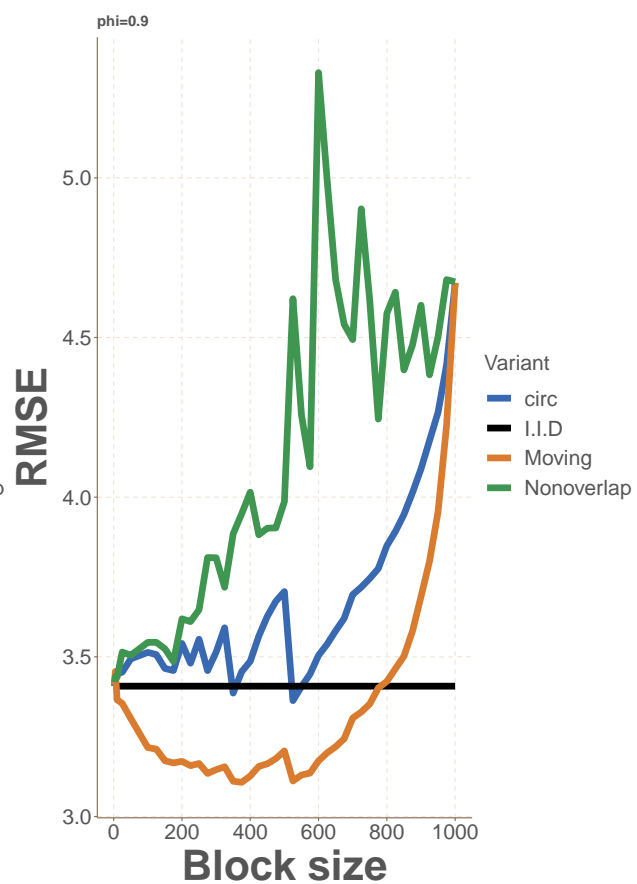


Figure 6: Modèle pour  $\phi = 0.9$ .

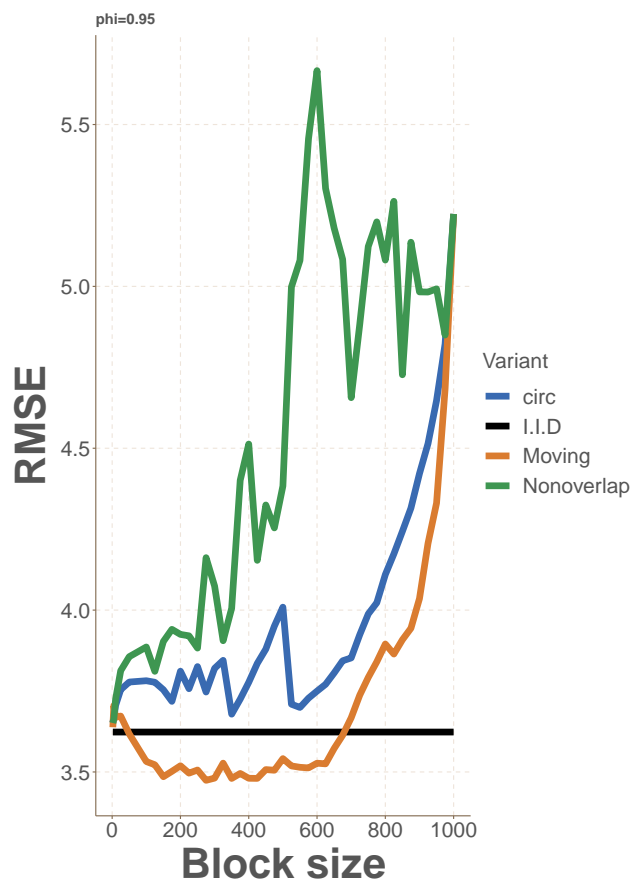


Figure 7: Modèle pour  $\phi = 0.95$ .

## 1.2 Graphiques Modèle 2

Evolution de l'erreur sur le test set selon la taille du blocs (sur une grille fixée à l'avance). Répété 20 fois. Voir figs. 8 to 10.

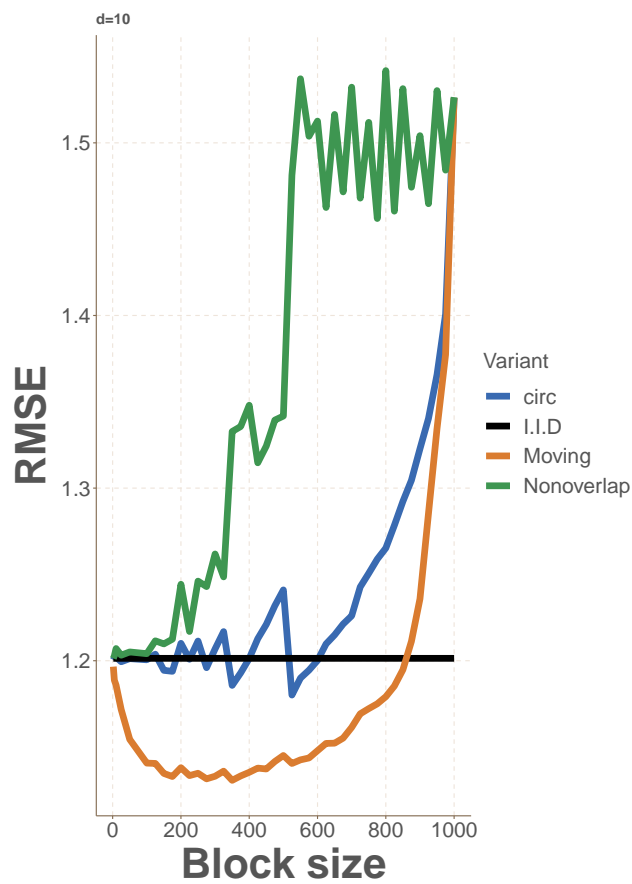


Figure 8: Modèle 2 pour  $d = 10$ .

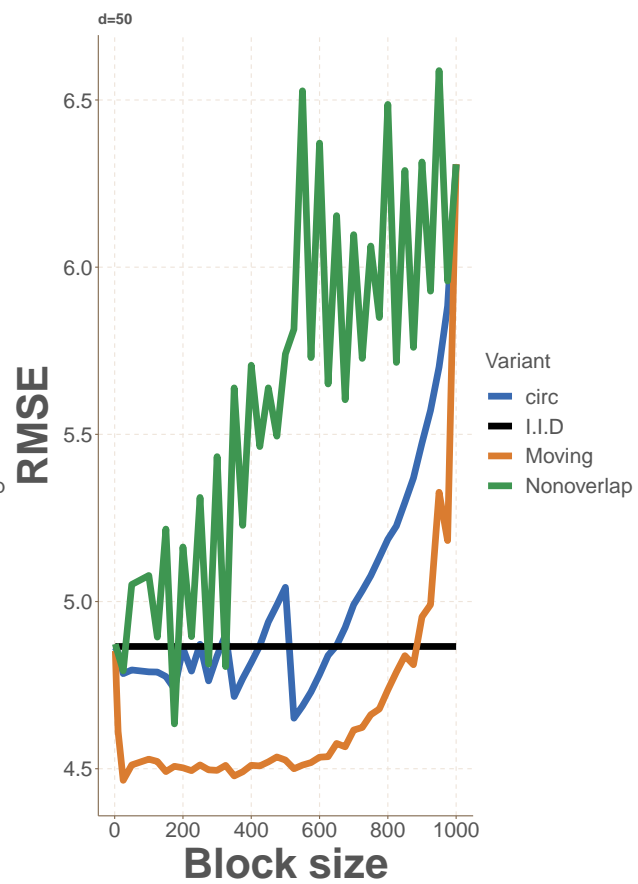


Figure 9: Modèle 2 pour  $d = 50$ .

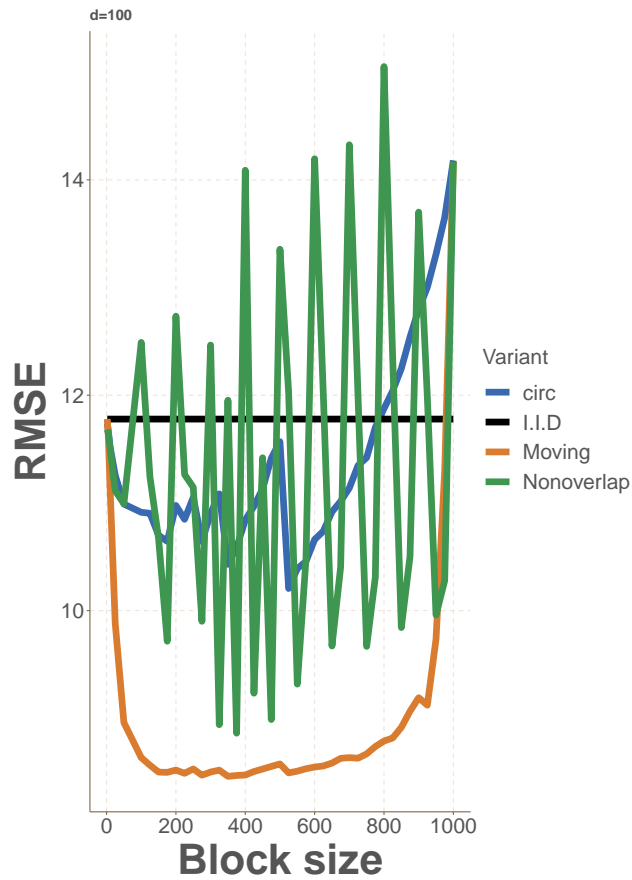


Figure 10: Modèle 2 pour  $d = 100$ .

## 2 Tentative de choix de blocs par pacf

Expérience :

1. Première étape: (Répéter M fois)
  - (a) Simuler modèle (eq. (1) ou eq. (2))
  - (b) Fit auto.arima
  - (c) Regarder le pacf des résidus
  - (d) Garder les lags où la corrélation partielle  $\geq \sqrt{\frac{2}{T_{train}}} qnorm(\frac{\alpha}{2})$  (ici  $\alpha = 0.05$ )
2. Deuxième étape: (Répéter M fois)
  - (a) Simuler à partir du même modèle.
  - (b) Utiliser les lags obtenus de la première étape comme input pour le block size dans la forêt.

3. Admirer.

Ici répétition = 20.

## 2.1 Modèle 1

Evolution de l'erreur sur le test set selon la taille du blocs choisi au préalable sur répétition du modèle. Voir figs. 11 to 15.

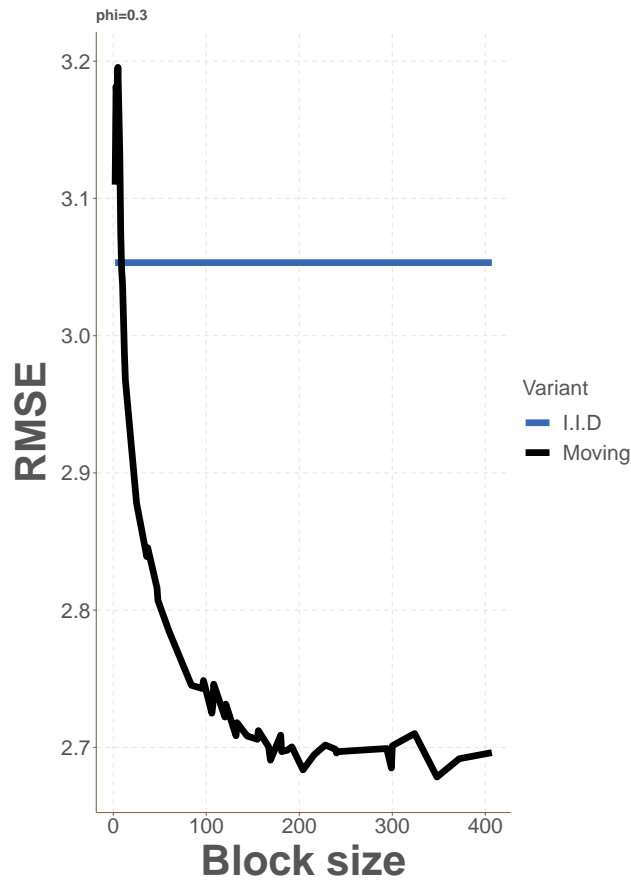


Figure 11: PACF pour modèle 1 pour  $\phi = 0.3$ .

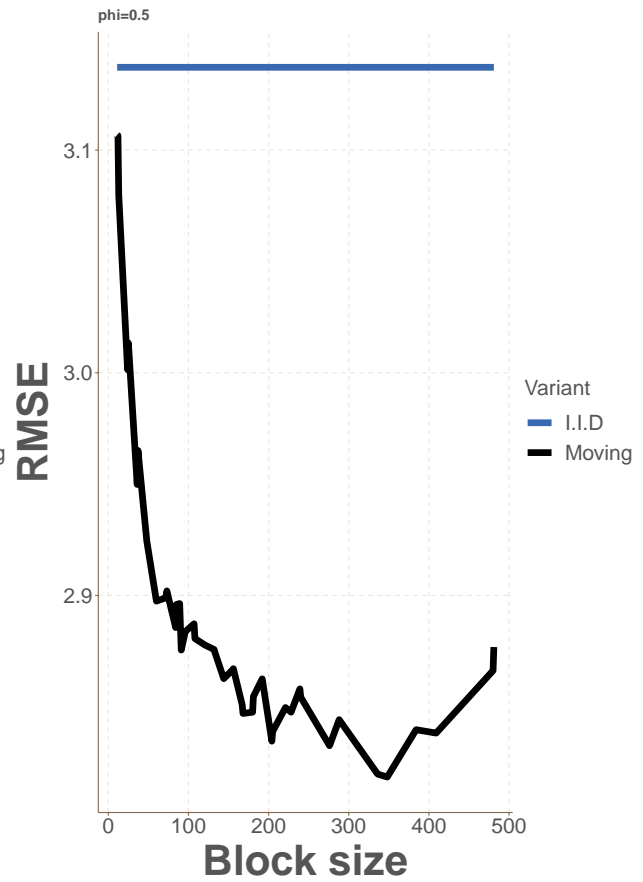


Figure 12: PACF pour modèle 1 pour  $\phi = 0.5$ .



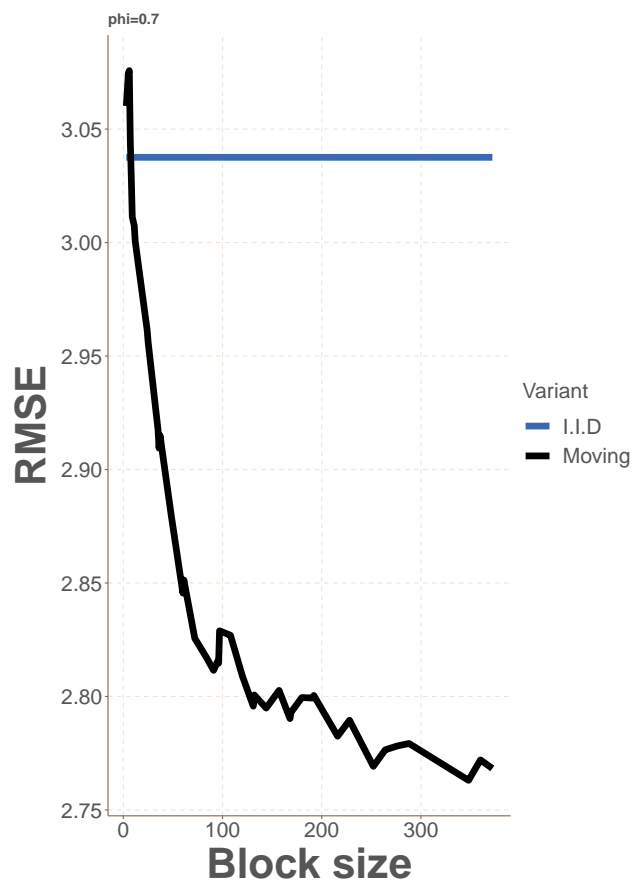


Figure 13: PACF pour modèle 1 pour  $\phi = 0.7$ .

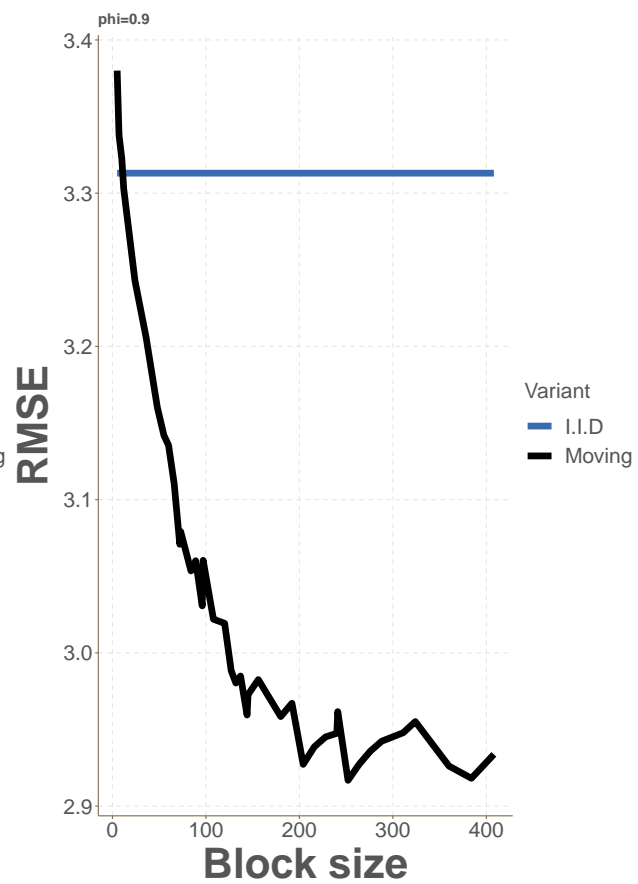


Figure 14: PACF pour modèle 1 pour  $\phi = 0.9$ .

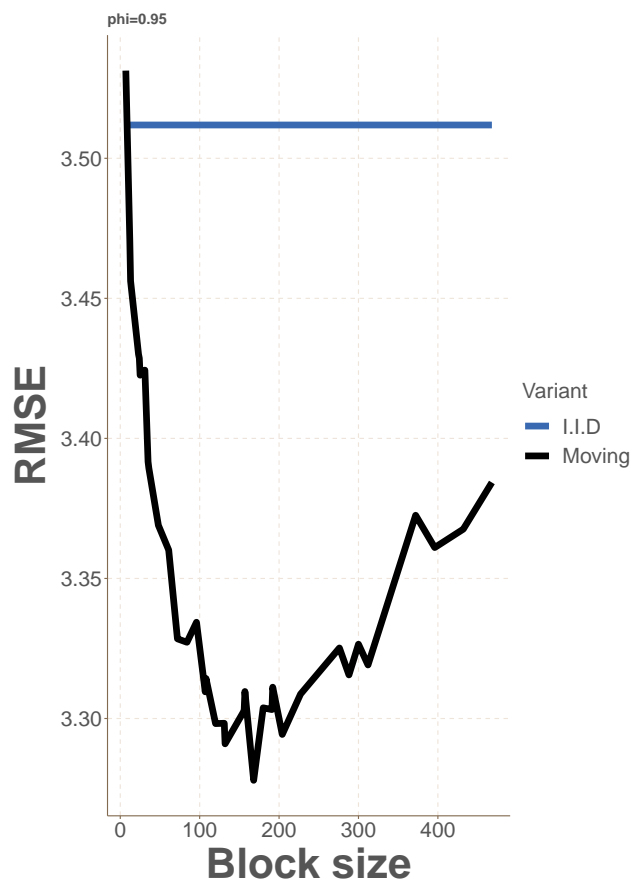


Figure 15: PACF pour modèle 1 pour  $\phi = 0.95$ .

## 2.2 Modèle 2

Evolution de l'erreur sur le test set selon la taille du blocs choisi au préalable sur répétition du modèle. Voir figs. 16 to 18.

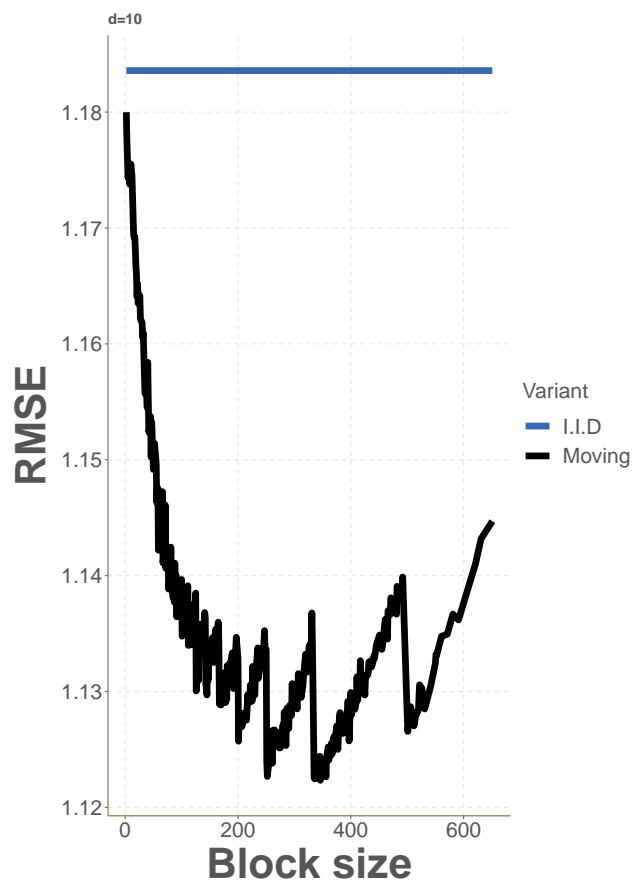


Figure 16: PACF Modèle 2 pour  $d = 10$ .

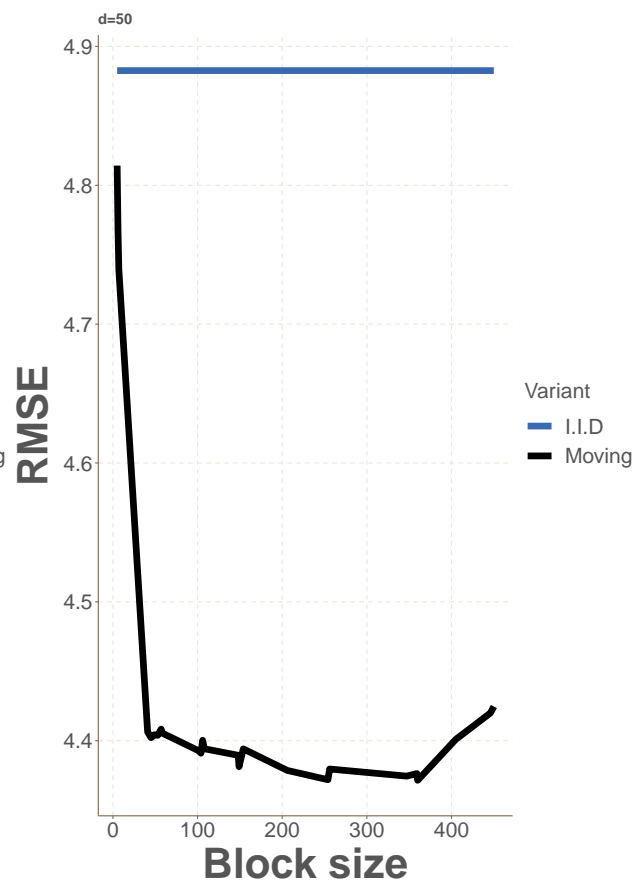


Figure 17: PACF Modèle 2 pour  $d = 50$ .

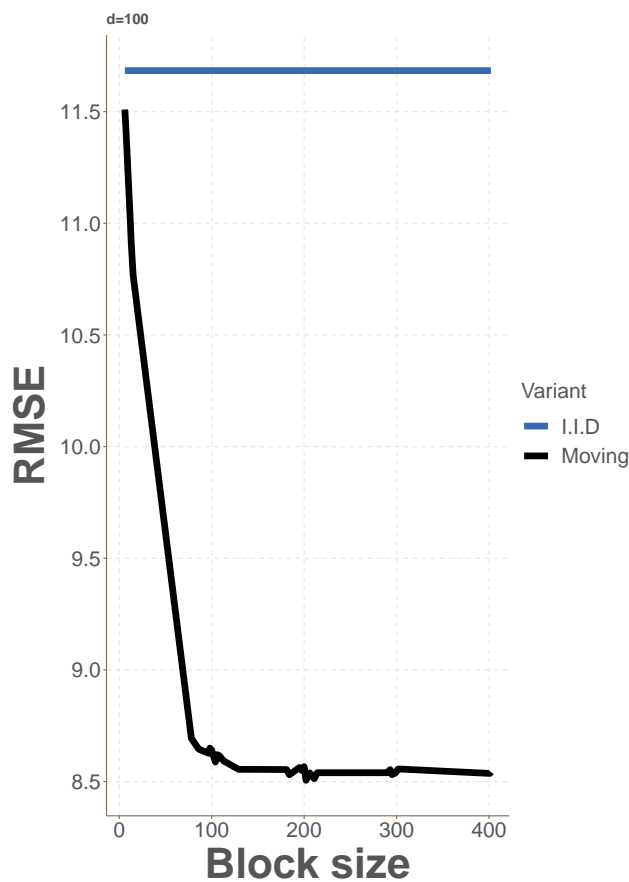


Figure 18: PACF Modèle 2 pour  $d = 100$ .

### 3 Tentative choix de blocs continues par pacf

Expérience (Répéter X fois) :

1. Simuler modèle à partir de eq. (1) ou eq. (2).
2. Fit auto.arima
3. Regarder le pacf des résidus
4. Garder les lags où la corrélation partielle  $\geq \sqrt{\frac{2}{T_{train}}} qnorm(\frac{\alpha}{2})$  (ici  $\alpha = 0.05$ )
5. Utiliser les lags de l'étape 4 comme input pour le block size dans la forêt.

Graphique dans section 3.1 et 3.2, répétition = 200. Variance énorme car pour la majorité des paramètres il n'y a que peu d'observations (entre 1 et 20 en moyenne)

### 3.1 Modèle 1

Evolution de l'erreur sur le test set selon la taille du blocs choisi au préalable sur répétition du modèle. Voir figs. 19 to 23.

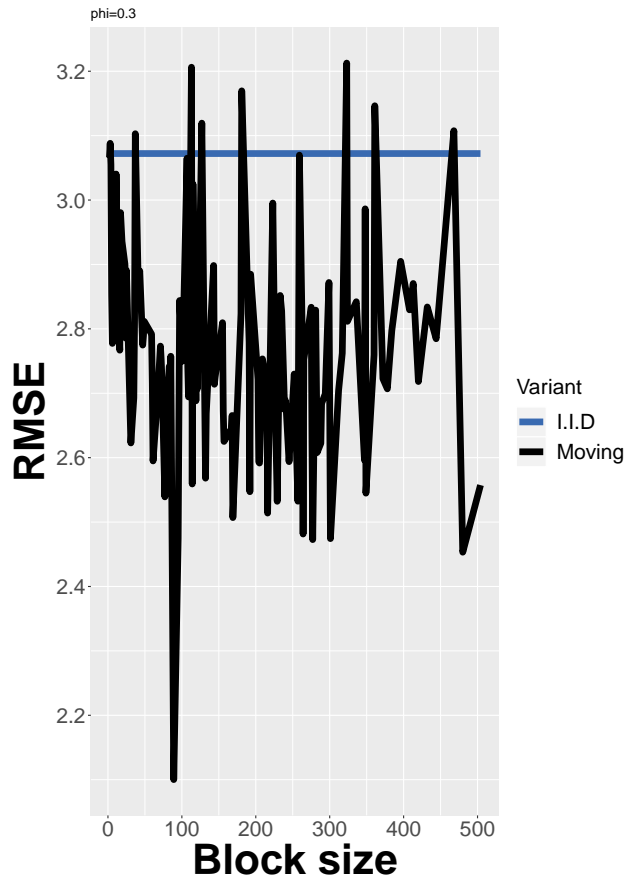


Figure 19: PACF pour modèle 1 pour  $\phi = 0.3$ .

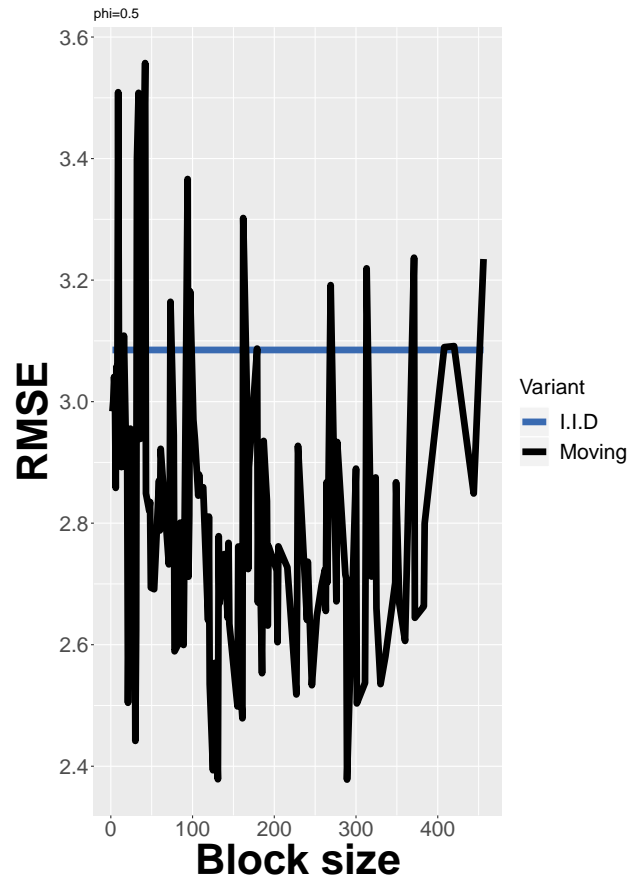


Figure 20: PACF pour modèle 1 pour  $\phi = 0.5$ .

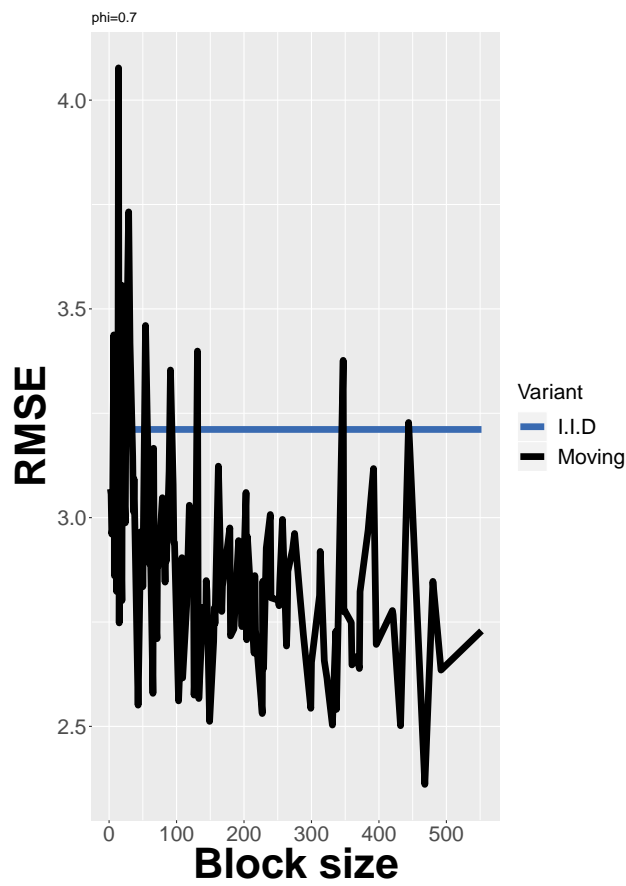


Figure 21: PACF pour modèle 1 pour  $\phi = 0.7$ .

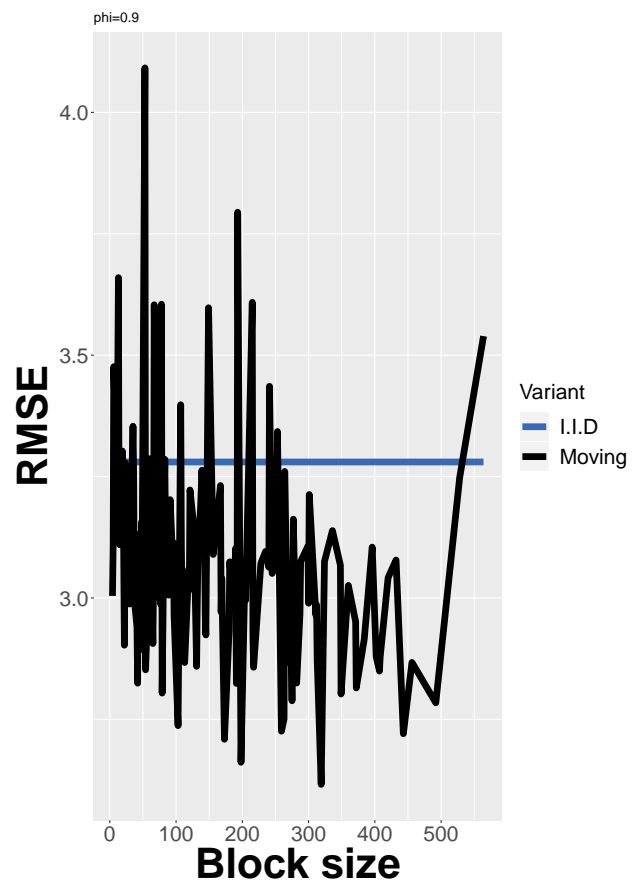


Figure 22: PACF pour modèle 1 pour  $\phi = 0.9$ .

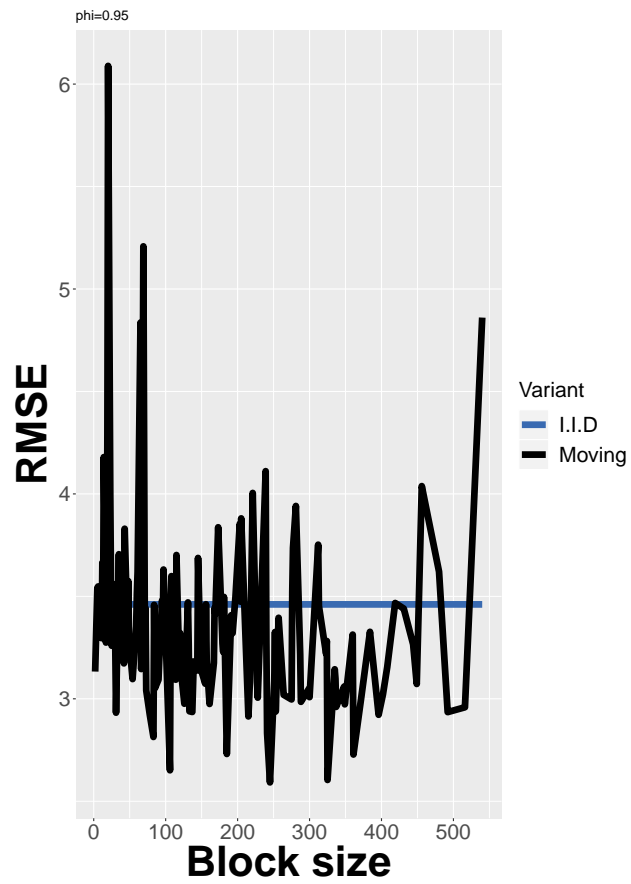


Figure 23: PACF pour modèle 1 pour  $\phi = 0.95$ .

### 3.2 Modèle 2

Evolution de l'erreur sur le test set selon la taille du blocs choisi au préalable sur répétition du modèle. Voir figs. 24 to 26.

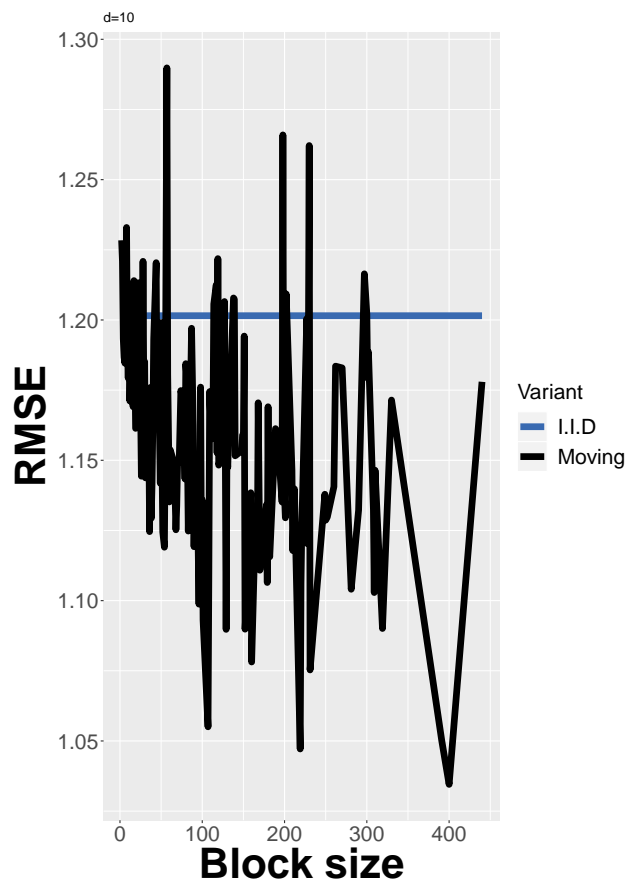


Figure 24: PACF Modèle 2 pour  $d = 10$ .

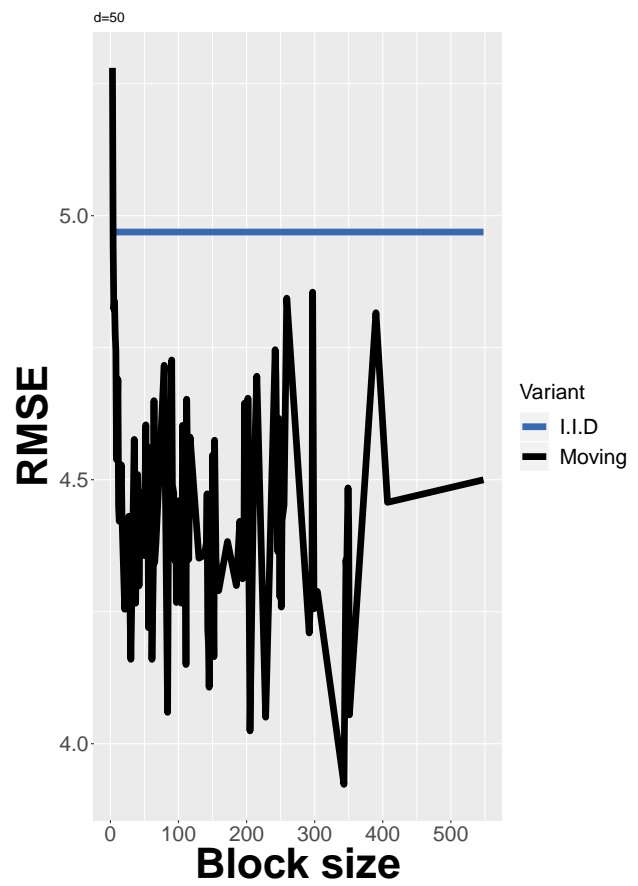


Figure 25: PACF Modèle 2 pour  $d = 50$ .



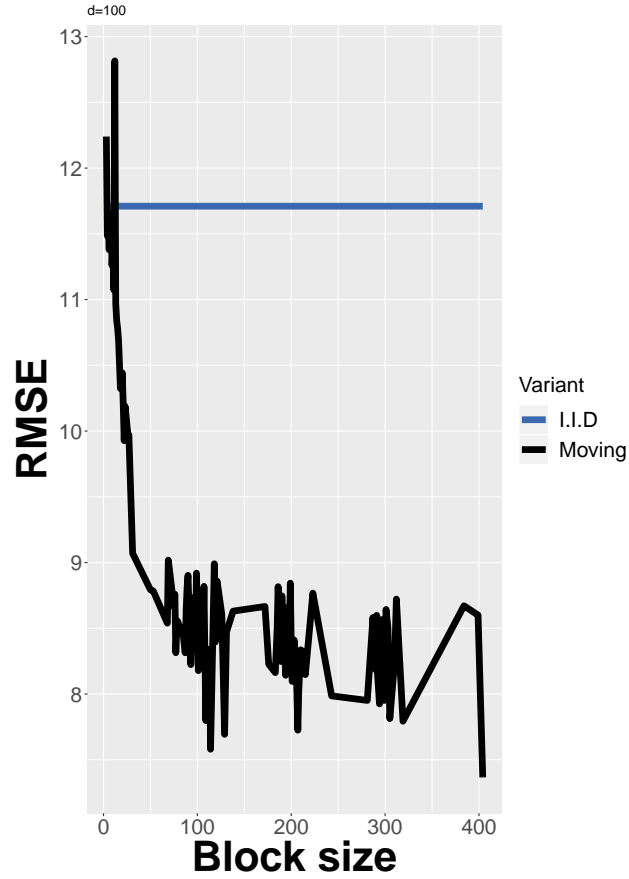


Figure 26: PACF Modèle 2 pour  $d = 100$ .

## 4 Expérience données FR

Simulation à partir du gam :

$$\begin{aligned}
 \text{Load} = & TdchTempLisse975 + TypeJour1 : rupture2 + td + s(posan, k = 20, bs = 'cc', by = TypeJour2) \\
 & + s(EcartTlag14, k = 12) + s(nebulosite, k = 6, bs = 'cr', by = FactorChauff_{975}) \\
 & + s(temperature) + s(TempLisse_{998}, k = 5) + s(TempLisse97, k = 5) \\
 & + s(TempMaxLisse79, k = 6, bs = 'cr')
 \end{aligned}$$

Training set avec les années 2013- $X$  et test set pour  $X$ -2018 avec  $X = \{2016, 2017, 2018\}$ .  
 Voir figs. 27 to 29

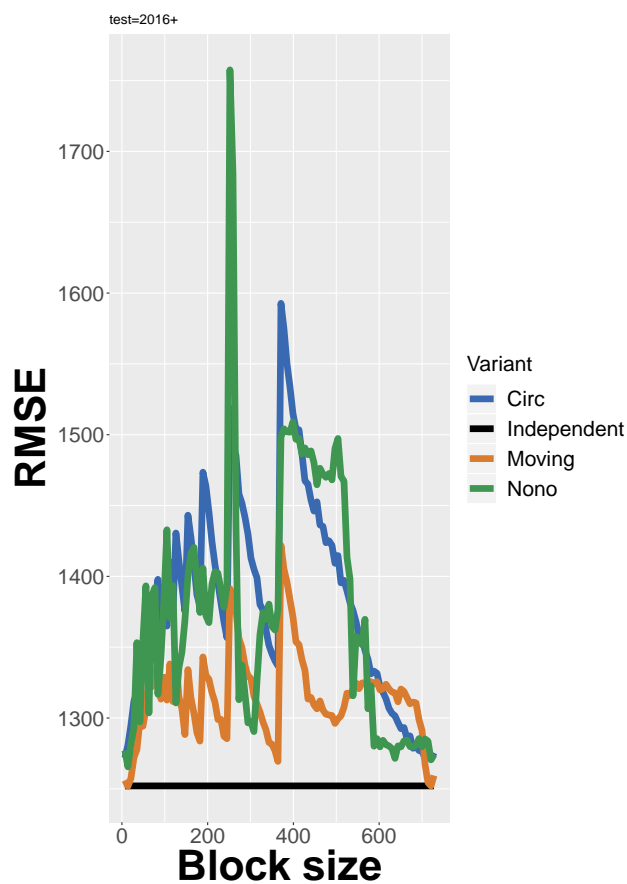


Figure 27: Erreur sur test 2016 et plus

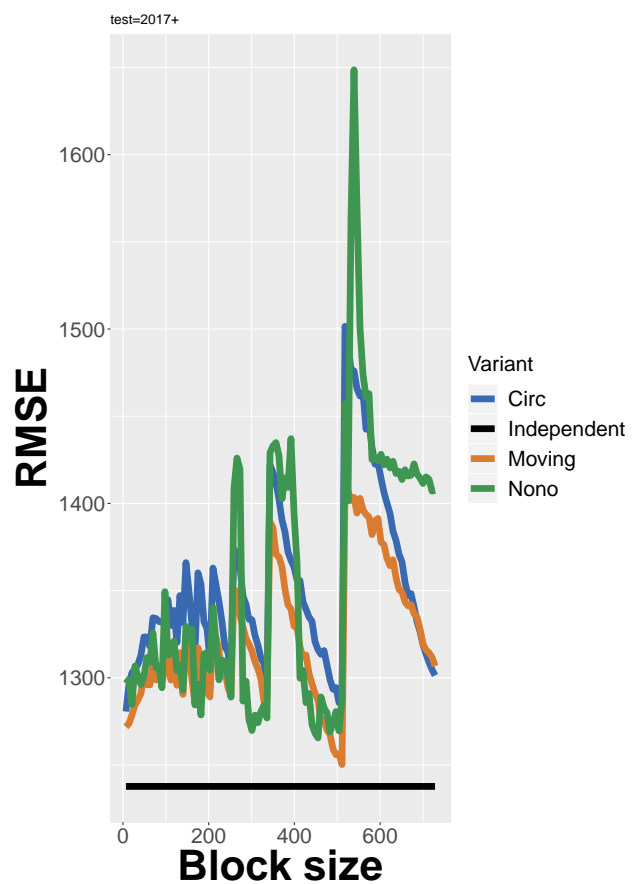


Figure 28: Erreur sur test 2017 et plus

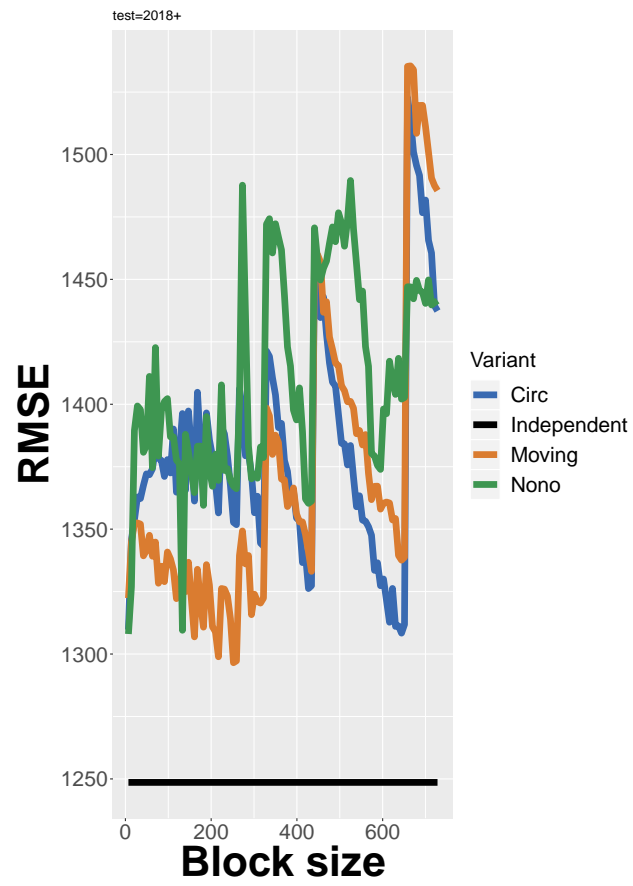


Figure 29: Erreur sur test 2018 et plus