Université TEK-UP — Cours Séries Temporelles pour Data Science & IA

Mini-projet 2026 — Option A (données financières)

# Blé, maïs, soja

## Volatilité et interdépendances des prix mondiaux des céréales face aux chocs climatiques et géopolitiques

*Application à la sécurité alimentaire de la Tunisie*

Réalisé par : Alaa Eddine Zamouri, Molka Cherni et Sarra Touati

Enseignant : M. Ahmed Dhouibi

Année universitaire 2025-2026

# Résumé

Ce projet modélise les prix journaliers des contrats à terme sur le blé, le maïs et le soja cotés à Chicago (Yahoo Finance, janvier 2010 à décembre 2025, 4 022 séances) afin de répondre à une question concrète pour un pays importateur comme la Tunisie : le prix du blé est-il prévisible à 10 séances, et sa volatilité l'est-elle ?

Les tests ADF et KPSS montrent que les log-prix sont non stationnaires et les rendements stationnaires. Aucun modèle ne prévoit significativement mieux le **prix** que la marche aléatoire au seuil de 5 % (test de Diebold-Mariano) : ARIMA, VAR, LSTM, GRU et hybride ARIMA + LSTM ont tous une erreur moyenne d'environ 2,5 %. Le meilleur modèle est le **VECM**, qui exploite deux relations de cointégration entre les trois céréales : il réduit le RMSE de 3,1 % (18,7 contre 19,3 cents par boisseau), avec une p-value de Diebold-Mariano de 0,058 à un jour, à la limite de la significativité.

La **volatilité**, elle, est fortement prévisible. Un GARCH(1,1) à loi de Student décrit une persistance de 0,986 (demi-vie d'un choc : 51 séances) et ses pics coïncident avec l'embargo russe d'août 2010, l'invasion de l'Ukraine (volatilité annualisée de 107 % le 11 mars 2022) et la fin de l'accord céréalier de la mer Noire (juillet 2023). Enfin, une corrélation dynamique DCC-GARCH montre que le lien blé–maïs varie entre 0,37 et 0,74 selon les périodes.

**Mots-clés :** séries temporelles, ARIMA, VAR, cointégration, VECM, GARCH, DCC-GARCH, LSTM, GRU, Diebold-Mariano, céréales, sécurité alimentaire.

# Table des matières

**Résumé** ..... **2**

**1. Introduction** ..... **4**
1.1 Contexte et justification du choix ..... 4
1.2 Problématique et questions de recherche ..... 4
1.3 Démarche et organisation du code ..... 4

**2. Données et préparation** ..... **4**
2.1 Source et variables ..... 4
2.2 Transformations du fournisseur et nettoyage ..... 5
2.3 Indicateurs dérivés et découpage ..... 5

**3. Analyse exploratoire** ..... **7**
3.1 Visualisation ..... 7
3.2 ACF et PACF ..... 8
3.3 Décomposition STL ..... 9
3.4 Valeurs aberrantes ..... 10
3.5 Stationnarité : ADF et KPSS ..... 10

**4. Modélisation classique** ..... **11**
4.1 ARIMA ..... 11
4.2 VAR et causalité de Granger ..... 12
4.3 Cointégration et VECM ..... 12
4.4 Volatilité conditionnelle : ARCH et GARCH ..... 13

**5. Modélisation avancée** ..... **14**
5.1 LSTM et GRU ..... 15
5.2 Approche hybride ARIMA + LSTM ..... 16
5.3 Modèle additionnel : Prophet ..... 16

**6. Prévision et évaluation** ..... **16**
6.1 Protocole ..... 16
6.2 Tableau récapitulatif ..... 16
6.3 Comparaison par critères d'information (AIC, AICc, BIC) ..... 17
6.4 Test de Diebold-Mariano ..... 19
6.5 Prévision de la volatilité ..... 19
6.6 Choix du meilleur modèle ..... 19

**7. Extensions** ..... **20**
7.1 Corrélation dynamique blé–maïs (DCC-GARCH) ..... 20
7.2 Variables exogènes : dollar et pétrole ..... 20
7.3 Le site web AgriScope ..... 20

**8. IA générative et LLM** ..... **21**
8.1 Notre interprétation (rédigée avant le LLM) ..... 21
8.2 Génération d'hypothèses de modèles (prompt 1) ..... 21
8.3 Vulgarisation pour un décideur (prompt 2) ..... 22
8.4 Recommandations simulées et leurs limites (prompt 3) ..... 22
8.5 Comparaison humain / LLM et erreurs détectées ..... 23

**9. Discussion critique** ..... **24**
9.1 Limites des modèles et des données ..... 24
9.2 Apport et limites de l'IA générative pour les séries temporelles agricoles ..... 25

**10. Conclusion** ..... **25**

**Bibliographie** ..... **25**

**Annexe A — Prompts générés avec les chiffres de l'étude** ..... **27**

**Annexe B — Réponses intégrales du LLM** ..... **27**

**Annexe C — Concordance R / Python (socle classique)** ..... **30**

**Annexe D — Reproductibilité** ..... **30**

# 1. Introduction

## 1.1 Contexte et justification du choix

**Justification de la thématique (exigée par le sujet) :** la Tunisie importe une grande partie du blé qu'elle consomme ; chaque flambée du prix mondial, comme celle de mars 2022 après l'invasion de l'Ukraine, pèse directement sur le budget de l'État (subvention du pain) et sur la sécurité alimentaire du pays.

**Ordre de grandeur.** En 2024, l'Office des Céréales a importé 2,617 millions de tonnes de céréales, dont environ 1,412 million de tonnes de blé tendre et 601 000 tonnes de blé dur, pour une valeur proche de 2,28 milliards de dinars (chiffres de l'Office repris par Webmanagercenter, 29 août 2026). La presse économique tunisienne situe le taux de dépendance aux importations à 85 % pour le blé tendre, celui du pain, et à 33 % pour le blé dur (African Manager, 9 janvier 2024, qui ne cite pas sa source primaire : chiffre à lire comme un ordre de grandeur).

Nous retenons l'**option A** du sujet (données financières Yahoo Finance) appliquée au domaine agricole. Les contrats à terme (futures) du Chicago Board of Trade (CBOT) servent de prix de référence mondiaux : un contrat à terme est un engagement à acheter ou vendre une quantité de céréales à une date future, à un prix fixé aujourd'hui. Les trois céréales forment un panier d'actifs corrélés : le maïs et le blé se substituent dans l'alimentation animale, et le maïs et le soja se partagent les mêmes terres aux États-Unis.

## 1.2 Problématique et questions de recherche

*Comment se transmettent les chocs entre les prix du blé, du maïs et du soja, leur volatilité est-elle prévisible, et un modèle d'apprentissage profond prévoit-il mieux le prix du blé à 5-10 séances qu'un modèle économétrique ?*

1. Les prix des trois céréales sont-ils non stationnaires, et leurs rendements stationnaires ?
2. Le rendement du maïs ou du soja aide-t-il à prévoir celui du blé (causalité de Granger) ?
3. Existe-t-il des relations d'équilibre de long terme (cointégration) entre les trois prix ?
4. La volatilité du blé forme-t-elle des grappes autour des chocs connus (2010, 2012, 2022, 2023) ?
5. LSTM et GRU battent-ils significativement la marche aléatoire et l'ARIMA (test de Diebold-Mariano) ?

## 1.3 Démarche et organisation du code

L'étude suit l'ordre du sujet : préparation (section 2), exploration (3), modèles classiques (4), modèles avancés (5), évaluation (6), extensions (7), IA générative (8) et discussion critique (9). Le notebook R Markdown `projet_cereales.Rmd` exécute toute l'étude de bout en bout. Conformément au sujet, le socle classique est aussi implémenté en Python (`01_socle_classique.ipynb`) ; les réseaux LSTM et GRU (section 5) sont implémentés **en Python uniquement** (`02_lstm_gru.ipynb`, PyTorch), appelé automatiquement depuis le notebook R.

# 2. Données et préparation

## 2.1 Source et variables

**Tableau 1 — Variables utilisées (Yahoo Finance, prix de clôture journaliers)**

| Variable | Symbole | Unité | Rôle |
|---|---|---|---|
| Blé (CBOT) | ZW=F | cents US / boisseau | Cible principale |

| Variable | Symbole | Unité | Rôle |
|---|---|---|---|
| Maïs (CBOT) | ZC=F | cents US / boisseau | Substitut, variable du VAR |
| Soja (CBOT) | ZS=F | cents US / boisseau | Variable du VAR |
| Indice du dollar | DX-Y.NYB | indice | Exogène (extension) |
| Pétrole WTI | CL=F | dollars / baril | Exogène (extension) |

Les données ont été téléchargées une seule fois avec `quantmod::getSymbols` (période demandée : 1er janvier 2010 – 31 décembre 2025 ; dernière séance disponible : 30 décembre 2025) puis enregistrées dans `data/cereales_yahoo.csv`, ce qui rend le notebook indépendant d'Internet. Un boisseau de blé pèse environ 27,2 kg : 510 cents par boisseau correspondent à 5,10 × 36,74 ≈ 187 dollars par tonne.

## 2.2 Transformations du fournisseur et nettoyage

- **Série continue :** Yahoo enchaîne le contrat le plus proche puis le suivant. Au changement de contrat (« roll »), le prix peut sauter sans événement de marché (voir 3.4).
- **Prix ajusté :** un contrat à terme ne verse pas de dividende ; le prix ajusté est égal au prix de clôture.
- **Calendrier :** le fichier brut contient 4 863 lignes, dont environ 840 sans cotation (week-ends, jours fériés). On garde les 4 022 séances où le blé cote ; pour le dollar et le pétrole, une valeur manquante est remplacée par la veille (marché fermé = prix inchangé).
- **Prix négatif :** le pétrole WTI a clôturé à −37,63 $ le 20 avril 2020 (contrat arrivant à échéance sans acheteur). Un log étant impossible, la valeur est remplacée par celle de la veille ; ce choix est documenté dans le code.

## 2.3 Indicateurs dérivés et découpage

Trois indicateurs techniques sont construits : le **rendement logarithmique** r = 100 × ln(P(t) / P(t−1)), les **moyennes mobiles** à 50 et 200 séances et la **volatilité réalisée** sur 21 séances, annualisée (× √252). Exemple : un blé qui passe de 550 à 561 cents donne un rendement de 100 × ln(561/550) ≈ +1,98 %.

**Tableau 2 — Découpage chronologique 90 / 10 (jamais aléatoire)**

| Échantillon | Début | Fin | Séances |
|---|---|---|---|
| Apprentissage | 04/01/2010 | 22/05/2024 | 3 619 |
| Test | 23/05/2024 | 30/12/2025 | 403 |

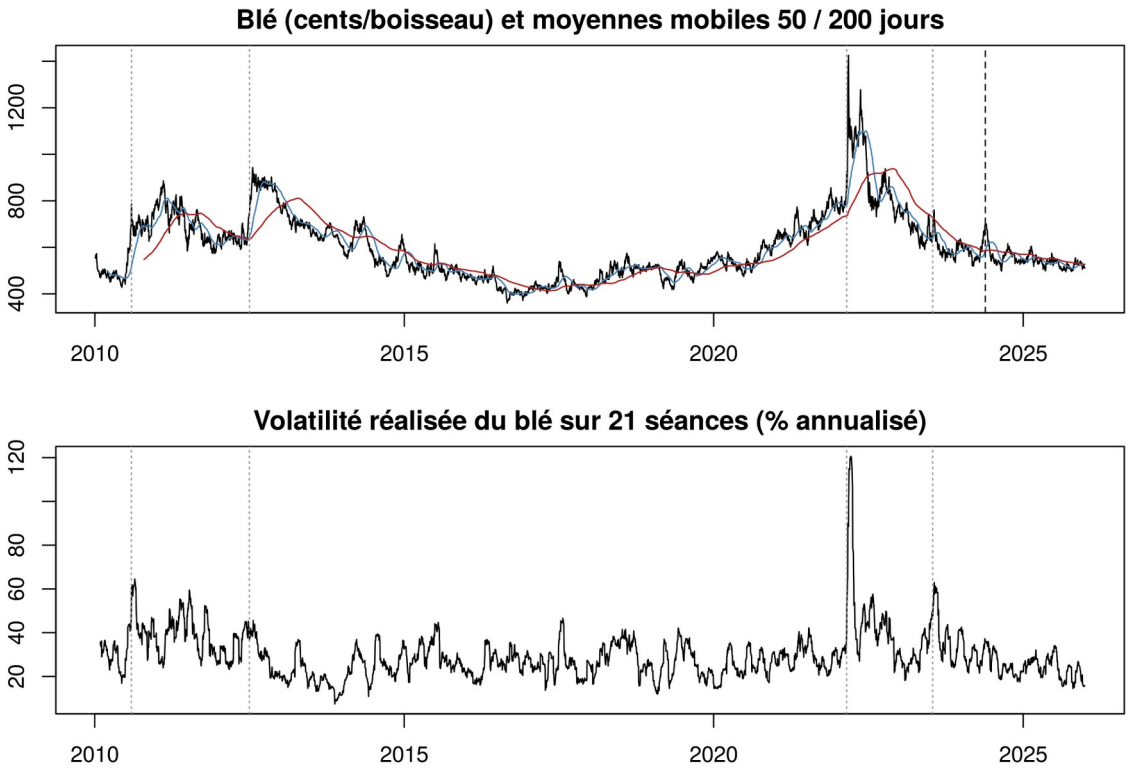



*Figure 1 — Prix du blé avec moyennes mobiles 50 et 200 jours (haut) et volatilité réalisée sur 21 séances (bas). Pointillés noirs : début du test.*

# 3. Analyse exploratoire

## 3.1 Visualisation

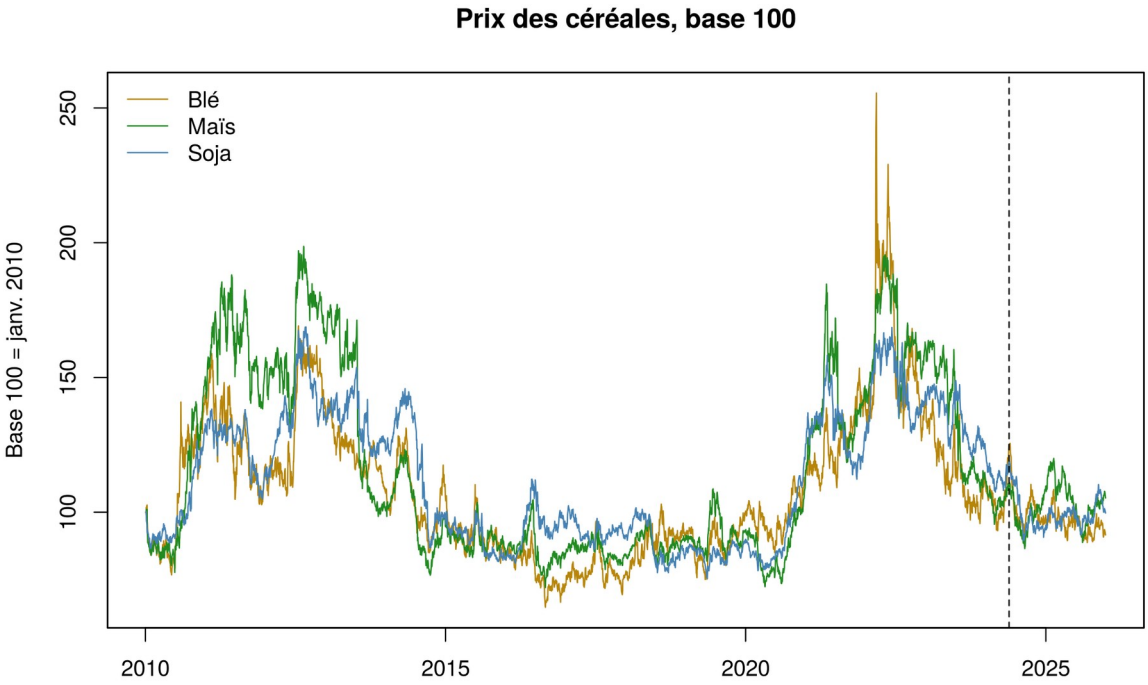



*Figure 2 — Prix des trois céréales en base 100 (janvier 2010).*

Les trois prix évoluent ensemble sur le long terme : hausse en 2010-2012 (sécheresse en Russie puis aux États-Unis), baisse et stabilisation entre 2014 et 2020, puis flambée en 2021-2022. Le blé atteint son maximum de 1 425 cents le 7 mars 2022, deux semaines après l'invasion de l'Ukraine, avant de redescendre à 511 cents fin 2025. Aucune tendance stable ni saisonnalité n'est visible à l'œil.

**Tableau 3 — Statistiques des rendements journaliers (%) sur l'apprentissage**

| Céréale | Moyenne | Écart-type | Asymétrie | Kurtosis | Min | Max | Jarque-Bera p |
|---|---|---|---|---|---|---|---|
| Blé | 0,006 | 2,03 | 0,39 | 7,16 | −11,30 | 19,70 | < 0,001 |
| Maïs | 0,003 | 1,78 | −1,73 | 26,26 | −26,86 | 8,62 | < 0,001 |
| Soja | 0,005 | 1,38 | −0,83 | 9,35 | −11,09 | 6,43 | < 0,001 |

Les rendements ont une moyenne quasi nulle et des **queues épaisses** : un kurtosis de 7,2 pour le blé, contre 3 pour une loi normale, signifie que les variations extrêmes sont beaucoup plus fréquentes que ne le prévoit la loi normale. Le test de Jarque-Bera rejette la normalité pour les trois séries. Les corrélations des rendements sont positives : 0,57 entre blé et maïs, 0,53 entre maïs et soja, 0,37 entre blé et soja.

## 3.2 ACF et PACF

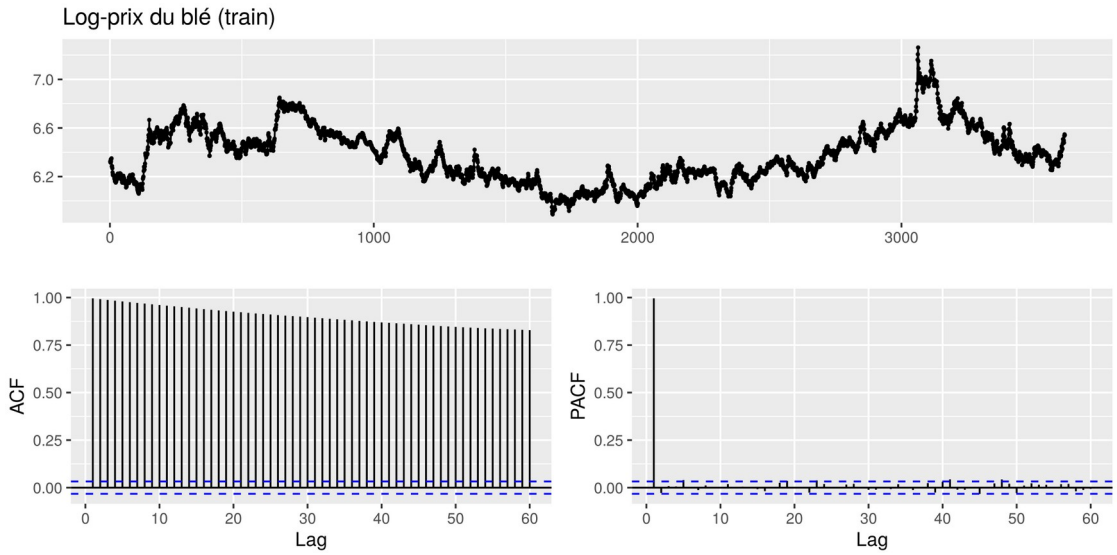



*Figure 3 — Log-prix du blé (apprentissage) : série, ACF et PACF.*

**Série brute (log-prix) :** l'ACF décroît très lentement et reste proche de 1 sur 60 retards, et la PACF a un seul pic au retard 1, proche de 1. C'est la signature d'une **marche aléatoire** : le prix d'aujourd'hui est essentiellement celui d'hier plus un choc.

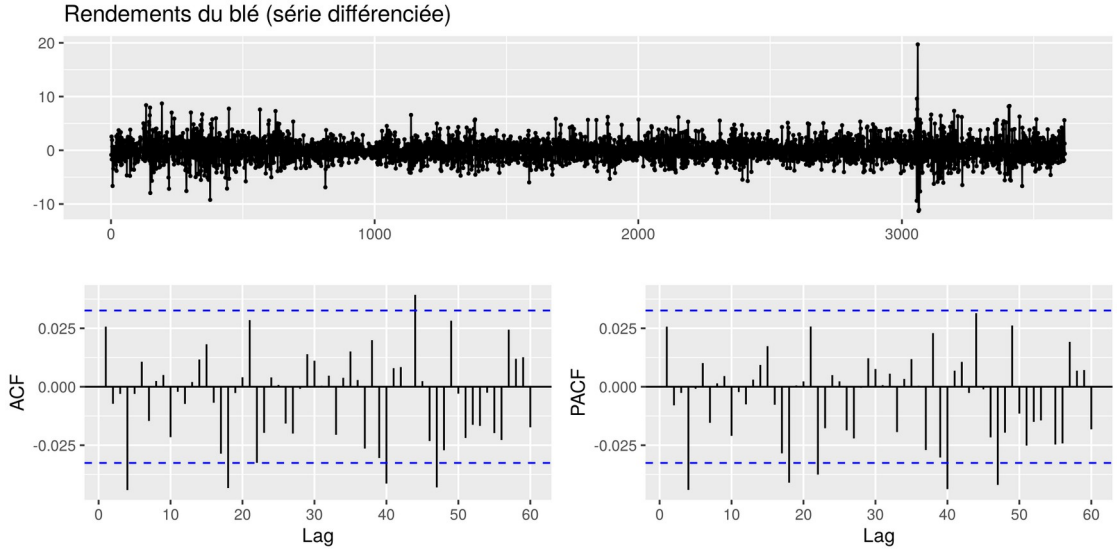



*Figure 4 — Rendements du blé (série différenciée) : série, ACF et PACF.*

**Série différenciée (rendements) :** presque toutes les autocorrélations sont dans la bande de confiance ±1,96/√n = ±0,033. Seule exception notable : le retard 4 (ACF = −0,044, PACF = −0,044). Les rendements passés n'apportent donc presque aucune information sur les rendements futurs, ce qui annonce une faible prévisibilité du prix.

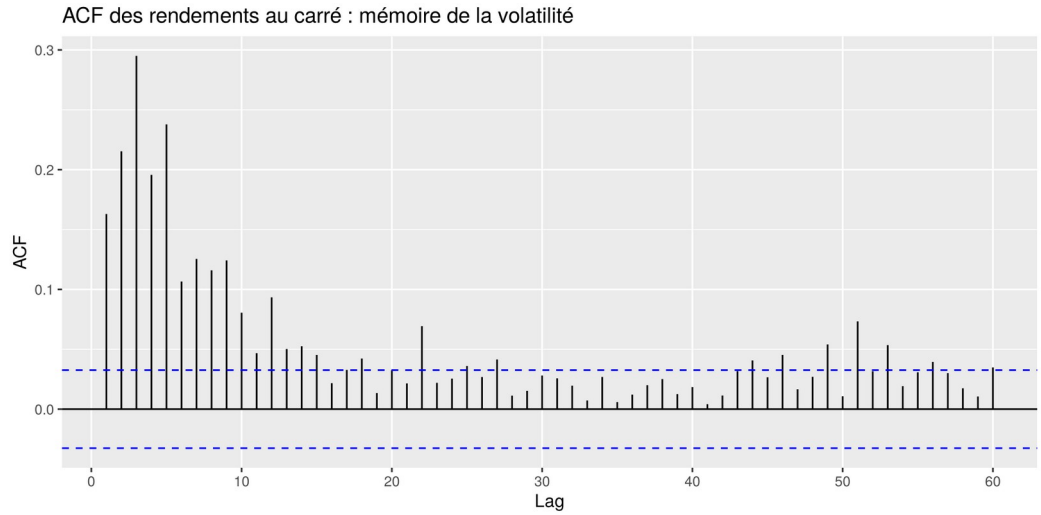



*Figure 5 — ACF des rendements au carré du blé.*

**Rendements au carré :** au contraire, l'ACF des carrés est positive et significative sur des dizaines de retards. Une grosse variation est suivie d'autres grosses variations : c'est la **mémoire de la volatilité**, qui justifie un modèle GARCH (section 4.4).

## 3.3 Décomposition STL

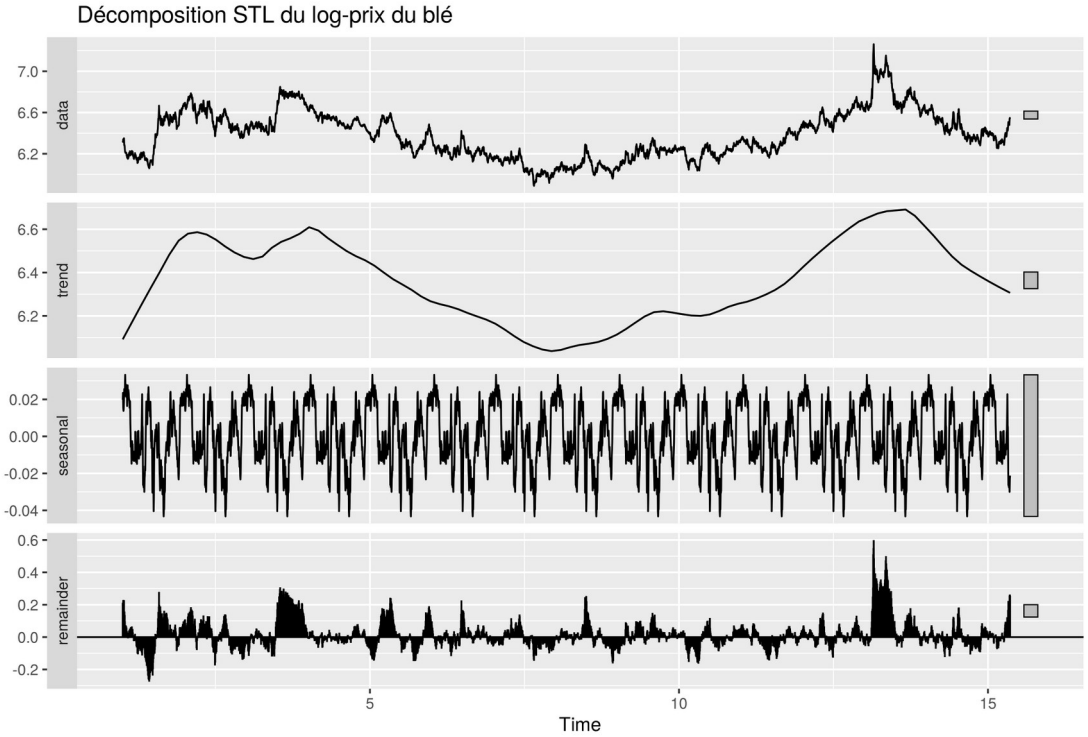



*Figure 6 — Décomposition STL du log-prix du blé (période saisonnière : 252 séances).*

La décomposition STL avec une période d'un an (252 séances) teste l'hypothèse d'un cycle lié aux récoltes. La **force de la saisonnalité** (Hyndman et Athanasopoulos), égale à 1 − Var(reste) / Var(saison + reste), vaut **0,00** : la composante saisonnière n'explique rien. La force de la tendance vaut 0,85 : l'essentiel des mouvements est une tendance stochastique de long terme. Conclusion : un modèle SARIMA n'est pas justifié, et nous utilisons des modèles non saisonniers, comme le permet le sujet (« si saisonnalité »).

## 3.4 Valeurs aberrantes

Avec la règle 3 × IQR (écart interquartile) sur les rendements du blé, 10 séances sur 4 021 sont aberrantes.

**Tableau 4 — Rendements aberrants du blé (règle 3 × IQR) et explication**

| Date | Rendement (%) | Explication probable |
|---|---|---|
| 03/03/2022 | +19,70 | Invasion de l'Ukraine (exportations de la mer Noire bloquées) |
| 08/03/2022 | −11,30 | Retournement après le pic du 7 mars 2022 |
| 10/03/2022 | −11,07 | Même épisode |
| 28/02/2022 | +9,61 | Même épisode |
| 25/02/2022 | −9,39 | Même épisode |
| 30/06/2011 | −9,22 | Rapports USDA du 30 juin 2011 : surfaces de maïs et stocks de blé supérieurs aux attentes |
| 08/10/2010 | +8,71 | Rapport USDA du 8 octobre 2010 : récolte de maïs revue en baisse, blé en hausse limite |
| 15/07/2010 | +8,39 | Sécheresse russe (plus forte hausse en 19 mois) ; date de changement de contrat possible |
| 24/07/2023 | +8,25 | Frappes sur les ports ukrainiens après la fin de l'accord |
| 19/07/2023 | +8,16 | Fin de l'accord céréalier de la mer Noire (17/07/2023) |

**Vérification dans la presse.** Les trois dates de 2010 et 2011 ont été recoupées avec la presse de l'époque (références en bibliographie). Le 15 juillet 2010, Bloomberg titre sur la plus forte hausse du blé en 19 mois, attribuée à la sécheresse en Russie. Le 8 octobre 2010, le rapport de récolte de l'USDA abaisse l'estimation de la récolte américaine de maïs (12,66 milliards de boisseaux contre 13,16 en septembre) : maïs, soja et blé clôturent en hausse limite, le blé de décembre gagnant 60 cents, soit +9,1 %, ce qui correspond bien au rendement logarithmique de +8,71 % du tableau 4. Le 30 juin 2011, les rapports Acreage et Grain Stocks de l'USDA annoncent des surfaces de maïs (92,3 millions d'acres contre 90,8 attendus) et des stocks de blé (861 millions de boisseaux contre 826 attendus) supérieurs aux attentes : le maïs chute et entraîne les autres céréales.

**Traitement retenu : conservation.** Ce sont presque tous de vrais chocs de marché, et les supprimer reviendrait à cacher l'information que le GARCH doit modéliser. Cas particulier : la chute de **−26,9 % du maïs le 15 juillet 2013** correspond au passage du contrat de juillet (ancienne récolte, chère) au contrat de septembre (nouvelle récolte). C'est un artefact de la série continue Yahoo, et non un événement de marché. Il est conservé par souci de simplicité mais explique le kurtosis très élevé du maïs (26,3) ; c'est une limite discutée en section 9.

## 3.5 Stationnarité : ADF et KPSS

**Tableau 5 — Tests de stationnarité (apprentissage, seuil 5 %)**

| Série | ADF stat. | ADF v.c. 5 % | KPSS stat. | KPSS v.c. 5 % | Conclusion |
|---|---|---|---|---|---|
| log-prix Blé | −2,72 | −2,86 | 1,706 | 0,463 | Non stationnaire |
| log-prix Maïs | −1,95 | −2,86 | 1,684 | 0,463 | Non stationnaire |
| log-prix Soja | −1,97 | −2,86 | 1,594 | 0,463 | Non stationnaire |
| rendement Blé | −42,29 | −2,86 | 0,045 | 0,463 | Stationnaire |
| rendement Maïs | −44,05 | −2,86 | 0,075 | 0,463 | Stationnaire |
| rendement Soja | −42,60 | −2,86 | 0,070 | 0,463 | Stationnaire |

ADF (H0 : racine unitaire) et KPSS (H0 : stationnarité) ont des hypothèses nulles opposées et se complètent. Pour les log-prix, ADF ne rejette pas H0 (statistique moins négative que −2,86) et KPSS rejette H0 (1,7 > 0,463) : les deux concluent à la non-stationnarité. Pour les rendements, c'est l'inverse : les deux concluent à la stationnarité. La fonction `ndiffs` confirme **d = 1** pour les trois céréales : les log-prix sont intégrés d'ordre 1, I(1).

# 4. Modélisation classique

## 4.1 ARIMA

**Identification manuelle.** D'après 3.2, seule l'autocorrélation au retard 4 sort de la bande de confiance. Nous estimons donc un AR « creux » sur les rendements, c'est-à-dire un ARIMA(4,1,0) où seul φ4 est libre (φ1 = φ2 = φ3 = 0). **Recherche automatique :** `auto.arima` (recherche exhaustive, p et q ≤ 4) choisit aussi un ARIMA(4,1,0), mais avec les quatre coefficients libres. Les deux approches convergent, et le modèle manuel est plus parcimonieux. La marche aléatoire ARIMA(0,1,0) sert de baseline.

**Tableau 6 — Comparaison des modèles ARIMA (apprentissage)**

| Modèle | Coefficients (écart-type) | AICc | Ljung-Box p | Jarque-Bera p | ARCH-LM p |
|---|---|---|---|---|---|
| Manuel (4,1,0) creux | φ4 = −0,044 (0,017) | −17 938,1 | 0,576 | < 0,001 | < 0,001 |
| auto.arima (4,1,0) | φ1 = 0,026 ; φ2 = −0,008 ; φ3 = −0,002 ; φ4 = −0,044 | −17 934,7 | 0,593 | < 0,001 | < 0,001 |
| Marche aléatoire (0,1,0) | — | −17 933,0 | 0,213 | < 0,001 | < 0,001 |

**Diagnostic des résidus.** Le test de Ljung-Box à 20 retards ne rejette pas l'indépendance des résidus (p ≈ 0,58), donc la dynamique linéaire est bien captée. La normalité est rejetée (Jarque-Bera et Shapiro-Wilk, p < 0,001) à cause des queues épaisses, ce qui rend les intervalles de prévision gaussiens trop étroits lors des crises. Surtout, le test ARCH-LM rejette massivement l'homoscédasticité (p ≈ $10^{-112}$) : la variance des résidus change dans le temps, ce qui impose la modélisation GARCH de la section 4.4. Le coefficient φ4 est significatif (−0,044, soit 2,7 écarts-types), mais son effet économique est minime : un rendement de +1 % il y a 4 séances réduit le rendement attendu d'aujourd'hui de 0,04 %.

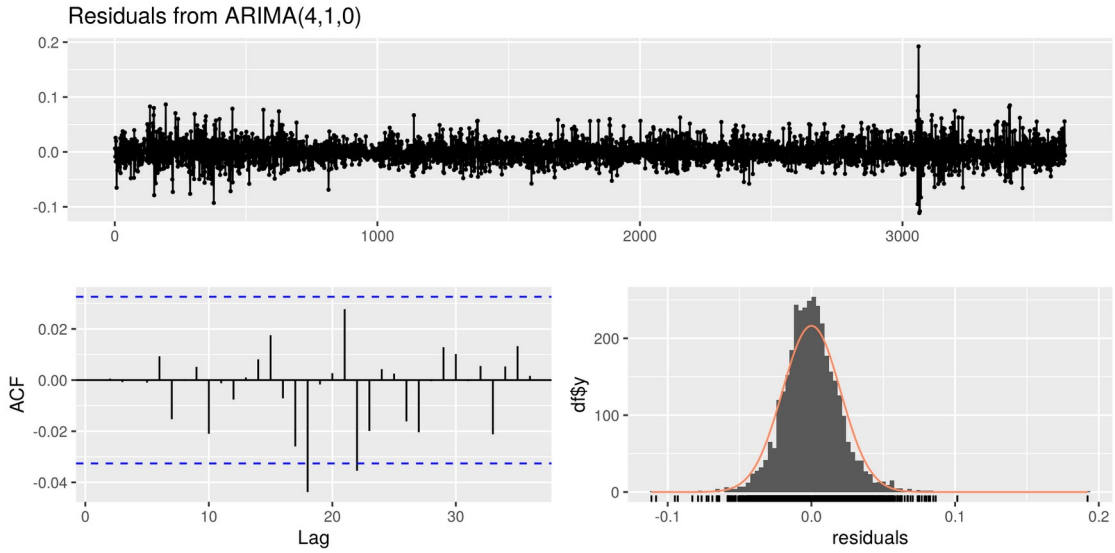



*Figure 7 — Résidus de l'ARIMA(4,1,0) automatique : chronologie, ACF et histogramme.*

## 4.2 VAR et causalité de Granger

Le VAR (Vector AutoRegression) modélise les trois rendements ensemble : chaque rendement dépend des rendements passés des trois céréales. Les quatre critères (AIC, HQ, BIC et FPE) choisissent **un seul retard (p = 1)**. Le VAR est stable : la plus grande racine vaut 0,039, bien en dessous de 1.

**Diagnostic :** le test du portemanteau rejette légèrement l'absence d'autocorrélation résiduelle (p = 0,014) ; il reste un peu de dynamique non captée, sans doute l'effet au retard 4 vu en 4.1. Les tests ARCH multivarié et Jarque-Bera multivarié rejettent l'homoscédasticité et la normalité (p < 0,001), comme pour l'ARIMA.

**Tableau 7 — Causalité de Granger entre rendements (tests par paires, 1 retard)**

| Cause | Effet | F | p-value | Conclusion (5 %) |
|---|---|---|---|---|
| Maïs | Blé | 1,190 | 0,276 | Pas de causalité |
| Blé | Maïs | 0,001 | 0,981 | Pas de causalité |
| Soja | Blé | 1,140 | 0,286 | Pas de causalité |
| Blé | Soja | 0,117 | 0,732 | Pas de causalité |
| Soja | Maïs | 2,812 | 0,094 | Pas de causalité (significatif à 10 %) |
| Maïs | Soja | 0,042 | 0,838 | Pas de causalité |

**Aucune causalité de Granger n'est significative à 5 %.** Le rendement d'hier du maïs ou du soja n'aide pas à prévoir celui du blé aujourd'hui. Cela ne veut pas dire que les marchés sont indépendants : la corrélation **le même jour** est forte (0,57 entre blé et maïs). Les chocs se transmettent donc dans la journée, trop vite pour être visibles avec des données journalières. C'est ce qu'on attend d'un marché liquide et efficient. Les tests joints dans le VAR (une céréale vers les deux autres) donnent la même conclusion (p = 0,89 ; 0,18 ; 0,16).

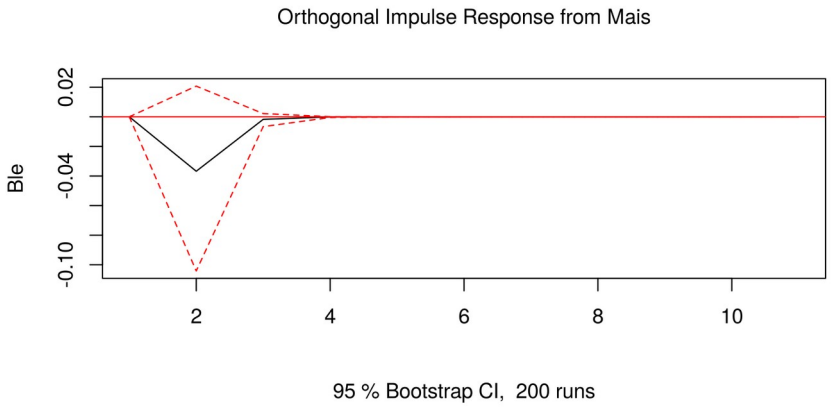



*Figure 8 — Réponse impulsionnelle du blé à un choc sur le maïs (intervalle bootstrap à 95 %).*

La réponse impulsionnelle orthogonalisée confirme ce résultat : un choc sur le maïs fait réagir le blé surtout à l'instant 0 (corrélation contemporaine), puis l'effet disparaît dès la séance suivante.

## 4.3 Cointégration et VECM

Deux séries I(1) peuvent partager une **relation d'équilibre de long terme** : chacune dérive, mais leur écart reste stable. Le test de Johansen est appliqué aux trois log-prix (constante dans la relation de cointégration, K = 2 retards en niveau).

**Tableau 8 — Test de Johansen sur les log-prix (apprentissage)**

| H0 | Trace | V.c. 5 % | Valeur propre max. | V.c. 5 % | Décision |
|---|---|---|---|---|---|
| r = 0 | 49,83 | 34,91 | 29,29 | 22,00 | Rejet |
| r ≤ 1 | 20,54 | 19,96 | 16,63 | 15,67 | Rejet |
| r ≤ 2 | 3,91 | 9,24 | 3,91 | 9,24 | Non rejet |

Les deux statistiques concluent à un **rang de cointégration r = 2** : il existe deux relations de long terme entre les trois prix, donc une seule tendance stochastique commune (3 − 2 = 1). Le test de Phillips-Ouliaris confirme la cointégration du couple blé–maïs (p ≤ 0,01). Les relations estimées (vecteurs β, normalisés) sont :

$$\ln(\text{Blé}) \approx 1{,}005 \times \ln(\text{Soja}) - 0{,}706$$
$$\ln(\text{Maïs}) \approx 1{,}306 \times \ln(\text{Soja}) - 3{,}062$$

**Interprétation de la première relation :** à long terme, le prix du blé suit celui du soja avec une élasticité proche de 1 : une hausse durable de 10 % du soja s'accompagne d'une hausse d'environ 10 % du blé. **Vitesse d'ajustement :** le coefficient α du blé vaut −0,0092 (p = 0,005). Négatif et significatif, il signifie que lorsque le blé est « trop cher » par rapport à cet équilibre, 0,92 % de l'écart est corrigé chaque séance. La demi-vie de l'écart vaut ln(0,5) / ln(1 − 0,0092) ≈ **75 séances**, soit environ 3,5 mois. Le maïs s'ajuste aussi (α = +0,006, p = 0,036), le soja non (p = 0,24) : le soja joue le rôle de prix « directeur » de la tendance commune.

## 4.4 Volatilité conditionnelle : ARCH et GARCH

Les tests ARCH-LM sur les résidus de l'ARIMA et sur les rendements du blé rejettent massivement l'absence d'effet ARCH (p ≈ $10^{-111}$). Cinq modèles sont estimés sur les rendements du blé, tous avec une loi de Student pour tenir compte des queues épaisses.

**Tableau 9 — Comparaison des modèles de volatilité (rendements du blé, apprentissage)**

| Modèle | Log-vraisemblance | AIC | BIC | Persistance | Ljung-Box $z^2$ p | ARCH-LM z p |
|---|---|---|---|---|---|---|
| GARCH(1,1) | −7 421,4 | 4,1053 | 4,1138 | 0,986 | 0,291 | 0,191 |
| GARCH(2,1) | −7 421,3 | 4,1058 | 4,1160 | 0,986 | 0,292 | 0,192 |
| GARCH(1,2) | −7 421,3 | 4,1058 | 4,1160 | 0,986 | 0,294 | 0,192 |
| GJR-GARCH(1,1) | −7 421,4 | 4,1058 | 4,1160 | 0,986 | 0,299 | 0,201 |
| EGARCH(1,1) | −7 416,5 | 4,1031 | 4,1133 | 0,982 | 0,081 | 0,026 |

AIC et BIC sont donnés par observation, comme dans rugarch. Les ordres supérieurs (2,1) et (1,2) n'améliorent pas la vraisemblance et ont un AIC et un BIC plus élevés : le GARCH(1,1) suffit. L'EGARCH a l'AIC et le BIC les plus faibles, mais de très peu (AIC 4,1031 contre 4,1053 ; BIC 4,1133 contre 4,1138), et ses résidus standardisés **échouent** au test ARCH-LM (p = 0,026). Le test de biais de signe (sign bias) ne détecte qu'une asymétrie faible (test joint p = 0,060). Nous retenons donc le **GARCH(1,1)**, qui passe tous les diagnostics : un modèle simple bien diagnostiqué plutôt qu'un modèle complexe mal diagnostiqué.

**Tableau 10 — Coefficients du GARCH(1,1) à loi de Student**

| Paramètre | Estimation | Écart-type | p-value | Interprétation |
|---|---|---|---|---|
| μ (moyenne) | −0,038 | 0,029 | 0,180 | Rendement moyen nul |
| ω (constante) | 0,058 | 0,023 | 0,011 | Variance plancher |
| α1 (choc récent) | 0,058 | 0,011 | < 0,001 | Réaction aux chocs de la veille |

| Paramètre | Estimation | Écart-type | p-value | Interprétation |
|---|---|---|---|---|
| β1 (mémoire) | 0,929 | 0,015 | < 0,001 | Persistance de la volatilité |
| ν (degrés de liberté) | 9,22 | 1,27 | < 0,001 | Queues épaisses confirmées |

**Lecture :** α1 + β1 = 0,986, donc un choc de volatilité est très persistant : sa demi-vie vaut ln(0,5) / ln(0,986) ≈ **51 séances** (environ 2,5 mois). Après une crise, le marché du blé reste agité plusieurs mois. Les résidus standardisés redeviennent homoscédastiques (Ljung-Box sur $z^2$ : p = 0,29 ; ARCH-LM : p = 0,19), ce qui valide le modèle.

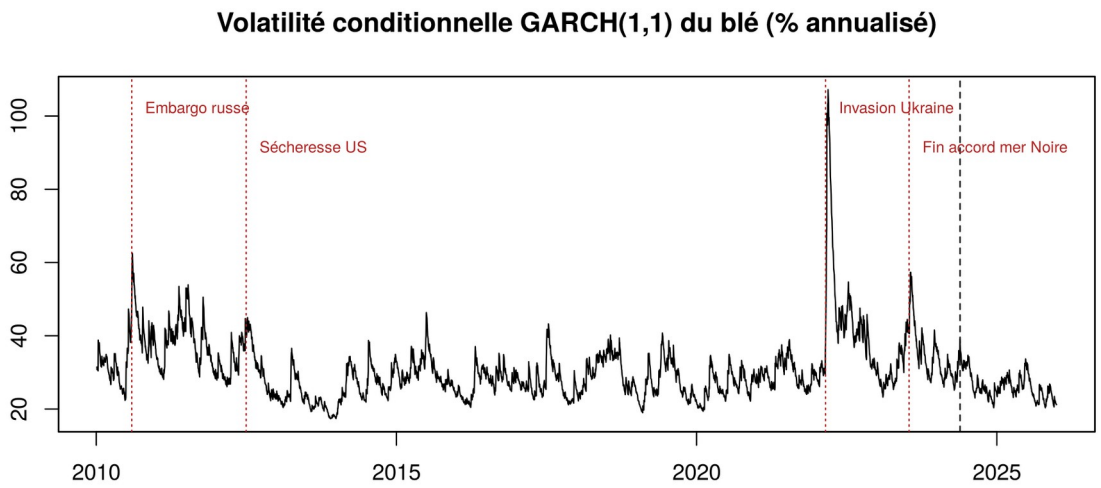



*Figure 9 — Volatilité conditionnelle GARCH(1,1) du blé, annualisée, avec les événements connus. Pointillés noirs : début du test.*

**Tableau 11 — Pics de volatilité conditionnelle (un par mois) et événements associés**

| Date | Volatilité annualisée | Événement |
|---|---|---|
| 11/03/2022 | 107 % | Invasion de l'Ukraine (24/02/2022) : blocage des ports de la mer Noire |
| 01/04/2022 | 74 % | Même crise, volatilité encore élevée (persistance) |
| 09/08/2010 | 63 % | Embargo russe sur les exportations de blé (annoncé le 05/08/2010) |
| 25/07/2023 | 57 % | Fin de l'accord céréalier de la mer Noire (17/07/2023) |
| 01/08/2023 | 56 % | Même épisode, volatilité encore élevée une semaine plus tard (persistance) |
| 13/07/2022 | 55 % | Négociations de l'accord de la mer Noire (signé le 22/07/2022) |

La volatilité de base du blé se situe autour de 25-30 % par an. Les trois plus grands pics correspondent exactement aux trois grands chocs géopolitiques de la période. La sécheresse américaine de 2012 n'apparaît que comme un pic modéré (environ 45 %), car elle a surtout touché le maïs et le soja.

# 5. Modélisation avancée

> **Langage :** la section 5 est implémentée en Python uniquement (PyTorch 2.15, notebook `02_lstm_gru.ipynb`), comme l'autorise la section 8.1 du sujet. Le notebook R l'exécute automatiquement et lit ses prévisions.

## 5.1 LSTM et GRU

**Mise en forme des données.** Le réseau reçoit les rendements des L dernières séances des trois céréales (fenêtre glissante de taille L × 3) et prédit directement le **rendement cumulé du blé sur 1 à 10 séances** (10 sorties, stratégie directe multi-horizon). Prédire l'écart au dernier prix connu, plutôt que le niveau du prix, évite au réseau d'extrapoler des niveaux de prix jamais vus, et la marche aléatoire correspond exactement à « prédire 0 ». Les entrées sont normalisées avec l'écart-type du **train uniquement**, pour éviter toute fuite d'information.

**Tableau 12 — Hyperparamètres et justification**

| Hyperparamètre | Valeur | Justification |
|---|---|---|
| Taille de fenêtre L | 20, 60, 120 (comparés) | 1 mois, 3 mois, 6 mois de bourse |
| Couches / neurones | 1 couche, 32 neurones | Environ 3 000 fenêtres : un réseau plus gros surapprendrait |
| Dropout | 0,2 | Régularisation |
| Optimiseur | Adam, lr = 0,001, lots de 64 | Réglage standard |
| Arrêt anticipé | Patience 8 époques (60 max.) | Arrêt dès que la perte de validation remonte |
| Graines | 3 graines, prévisions moyennées | Mesure de la stabilité |

**Validation croisée temporelle (walk-forward).** Cinq plis à fenêtre d'apprentissage croissante, chacun validé sur les 250 séances suivantes (environ un an), en respectant l'ordre du temps (jamais de k-fold classique).

**Tableau 13 — Validation walk-forward (RMSE du rendement cumulé à 1-10 séances, en points de %)**

| Modèle | L | RMSE moyen | Écart-type (5 plis) | Marche aléatoire |
|---|---|---|---|---|
| LSTM | 20 | 4,979 | 1,572 | 4,987 |
| LSTM | 60 | 4,966 | 1,551 | 4,987 |
| LSTM | 120 | 4,961 | 1,534 | 4,987 |
| GRU | 20 | 4,981 | 1,569 | 4,987 |
| GRU | 60 | 4,981 | 1,569 | 4,987 |
| GRU | 120 | 4,986 | 1,581 | 4,987 |

Les gains par rapport à la marche aléatoire sont minuscules (0,5 % au mieux, pour le LSTM avec L = 120) et bien plus petits que la variabilité entre plis (écart-type de 1,5). On retient L = 120 pour le LSTM et L = 20 pour le GRU. L'écart entre les 3 graines est faible (0,04 et 0,07 point de %) : les réseaux convergent vers des prévisions proches de 0.

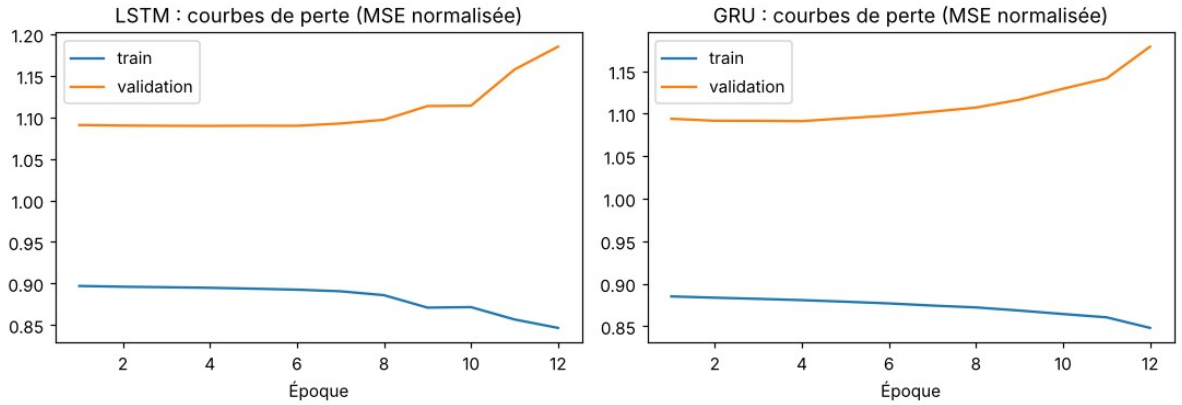



*Figure 10 — Courbes de perte train / validation des modèles finaux.*

**Lecture des courbes de perte :** la perte de validation est minimale dès les premières époques, puis remonte alors que la perte d'entraînement continue de baisser. C'est un **surapprentissage immédiat** : le réseau apprend le bruit du passé, car il n'y a presque pas de signal à apprendre. L'arrêt anticipé conserve les poids des premières époques.

### 5.2 Approche hybride ARIMA + LSTM

Les résidus de l'ARIMA(4,1,0), c'est-à-dire la partie non expliquée par le modèle linéaire, sont modélisés par un LSTM univarié (L = 60) qui prévoit les 10 résidus suivants. La prévision hybride ajoute à la prévision ARIMA une correction égale à la somme des résidus prévus, pondérée par les poids ψ de l'ARIMA intégré (effet cumulé d'un choc sur le niveau futur du prix).

### 5.3 Modèle additionnel : Prophet

Prophet (Meta) décompose la série en tendance par morceaux et saisonnalité annuelle. Comme il doit être réestimé à chaque nouvelle origine (environ 4 secondes par estimation), il est évalué sur 10 origines réparties dans le test et non sur les 394.

## 6. Prévision et évaluation

### 6.1 Protocole

**Évaluation par origine glissante :** pour chacune des 394 séances o du test (de la dernière séance du train à la 10e avant la fin), chaque modèle prévoit les 10 séances suivantes en n'utilisant que l'information disponible en o, avec les paramètres estimés sur le train (aucune réestimation). On obtient 394 × 10 prévisions par modèle, bien plus fiables qu'une seule prévision. Les métriques sont calculées sur les **prix** en cents par boisseau, plus parlants pour un décideur que les log-prix.

### 6.2 Tableau récapitulatif

**Tableau 14 — Tableau récapitulatif unique : prévision du prix du blé à 1-10 séances sur le test**

| Modèle | MSE | RMSE | MAE | MAPE (%) | DM p (h=1) | DM p (h=10) | Interprétabilité | Temps |
|---|---|---|---|---|---|---|---|---|
| VECM (r = 2) | 349,2 | 18,69 | 13,68 | 2,50 | 0,058 | 0,170 | Moyenne | < 1 s |
| ARIMA auto (4,1,0) | 370,6 | 19,25 | 13,94 | 2,54 | 0,262 | 0,933 | Forte | 1,5 s |
| ARIMA manuel (4,1,0) | 370,7 | 19,25 | 13,94 | 2,54 | 0,317 | 0,922 | Forte | < 1 s |
| Marche aléatoire | 371,6 | 19,28 | 13,98 | 2,55 | — | — | Très forte | 0 |
| GRU (L = 20) | 372,0 | 19,29 | 13,98 | 2,55 | 0,300 | 0,668 | Faible | ~ 3 min |

| Modèle | MSE | RMSE | MAE | MAPE (%) | DM p (h=1) | DM p (h=10) | Interprétabilité | Temps |
|---|---|---|---|---|---|---|---|---|
| VAR(1) | 372,8 | 19,31 | 14,01 | 2,56 | 0,598 | 0,308 | Moyenne | < 1 s |
| Hybride ARIMA + LSTM | 373,1 | 19,32 | 14,00 | 2,56 | 0,483 | 0,473 | Moyenne | ~ 1 min |
| LSTM (L = 120) | 373,3 | 19,32 | 13,99 | 2,56 | 0,156 | 0,309 | Faible | ~ 3 min |
| Prophet (10 origines) | 3 543,4 | 59,53 | 54,05 | 9,54 | n.c. | n.c. | Forte | 36 s |

*DM p : p-value du test de Diebold-Mariano contre la marche aléatoire (H0 : même précision). Temps : estimation sur le train (validation croisée incluse pour les réseaux). n.c. : non calculé (10 origines seulement).*

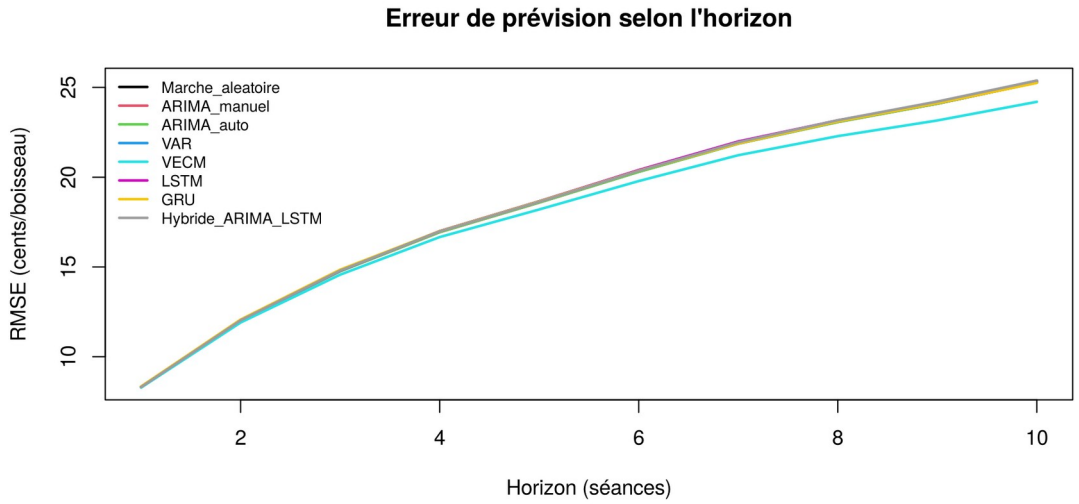



*Figure 11 — RMSE selon l'horizon de prévision (1 à 10 séances).*

**Lecture.** Tous les modèles sauf Prophet ont une erreur moyenne d'environ 2,5 % du prix (MAPE). L'erreur croît avec l'horizon, d'environ 8,3 cents à 1 séance à environ 25 cents à 10 séances, comme pour une marche aléatoire dont l'incertitude croît avec √h. Le **VECM** est le seul modèle à se détacher, d'autant plus que l'horizon s'allonge (24,2 contre 25,3 cents à h = 10) : c'est logique, puisque le rappel vers l'équilibre de long terme agit lentement. Prophet est très mauvais (MAPE de 9,5 %) : il extrapole une tendance et une saisonnalité qui n'existent pas sur un prix de marché.

## 6.3 Comparaison par critères d'information (AIC, AICc, BIC)

**Concept clé — AIC (critère d'Akaike) :** AIC = −2 × log-vraisemblance + 2 × k, où k est le nombre de paramètres. Il récompense l'ajustement aux données mais pénalise la complexité ; **plus l'AIC est bas, meilleur est le modèle**. Le BIC remplace 2 × k par ln(n) × k et pénalise donc davantage les modèles complexes. Règle pratique : une différence d'AIC (ΔAIC) inférieure à 2 signifie que deux modèles sont équivalents ; au-delà de 4, le moins bon est nettement moins soutenu par les données.

**Précaution :** un AIC n'est comparable qu'entre modèles estimés sur la même série et les mêmes observations. Nous faisons donc trois comparaisons séparées. Les LSTM, GRU et Prophet n'ont pas de vraisemblance comparable : ils sont jugés uniquement sur les métriques de prévision (6.2) et le test de Diebold-Mariano (6.4).

**Tableau 15 — (a) Modèles ARIMA sur le log-prix du blé (apprentissage, d = 1)**

| Modèle | k | Log-vraisemblance | AIC | AICc | BIC | ΔAIC |
|---|---|---|---|---|---|---|
| ARIMA(4,1,0) creux (manuel) | 1 | 8 971,0 | −17 938,1 | −17 938,1 | −17 925,7 | 0,0 |
| ARIMA(4,1,0) (auto.arima) | 4 | 8 972,4 | −17 934,7 | −17 934,7 | −17 903,7 | 3,4 |
| ARIMA(0,1,1) | 1 | 8 968,7 | −17 933,4 | −17 933,4 | −17 921,0 | 4,7 |
| ARIMA(1,1,0) | 1 | 8 968,7 | −17 933,4 | −17 933,4 | −17 921,0 | 4,7 |
| ARIMA(0,1,0) marche aléatoire | 0 | 8 967,5 | −17 933,0 | −17 933,0 | −17 926,8 | 5,1 |
| ARIMA(1,1,1) | 2 | 8 969,1 | −17 932,3 | −17 932,3 | −17 913,7 | 5,8 |
| ARIMA(0,1,2) | 2 | 8 968,8 | −17 931,6 | −17 931,6 | −17 913,0 | 6,4 |
| ARIMA(2,1,0) | 2 | 8 968,8 | −17 931,6 | −17 931,6 | −17 913,0 | 6,5 |
| ARIMA(2,1,2) | 4 | 8 969,2 | −17 928,4 | −17 928,4 | −17 897,4 | 9,7 |

L'ARIMA(4,1,0) creux choisi manuellement a le **meilleur AIC et le meilleur AICc** : il capte l'effet du retard 4 avec un seul paramètre. auto.arima, qui estime les quatre coefficients, est pénalisé de 3,4 points pour trois paramètres inutiles. Le BIC, plus sévère, place la marche aléatoire juste derrière le modèle creux (−17 926,8 contre −17 925,7) : l'information prédictive des rendements passés est presque nulle.

**Tableau 16 — (b) Équation du blé des modèles linéaires (mêmes 3 614 rendements, en %)**

| Modèle | k | Log-vraisemblance | AIC | BIC | ΔAIC |
|---|---|---|---|---|---|
| ARIMA(4,1,0) creux (manuel) | 2 | −7 682,3 | 15 368,6 | 15 381,0 | 0,0 |
| VECM (r = 2) — équation du blé | 6 | −7 678,8 | 15 369,6 | 15 406,8 | 1,0 |
| ARIMA(4,1,0) (auto.arima) | 5 | −7 680,9 | 15 371,8 | 15 402,8 | 3,2 |
| VAR(1) — équation du blé | 5 | −7 682,5 | 15 375,0 | 15 406,0 | 6,4 |
| Marche aléatoire avec dérive | 2 | −7 685,8 | 15 375,7 | 15 388,0 | 7,1 |

Pour comparer des modèles univariés et multivariés, on calcule la vraisemblance gaussienne des résidus de **l'équation du blé** seule, sur exactement les mêmes observations. Le VECM a la **meilleure log-vraisemblance** (−7 678,8) : le terme de correction d'erreur apporte une information réelle. Avec 6 paramètres, son AIC n'est qu'à 1,0 point de l'ARIMA creux : selon la règle ΔAIC < 2, **les deux modèles sont équivalents**. Le VAR(1) et la marche aléatoire sont nettement moins bons (ΔAIC > 6). Limite : les coefficients de cointégration β, estimés conjointement pour les trois équations, ne sont pas comptés dans k.

**Tableau 17 — (c) Nombre de retards du VAR (critères calculés par VARselect)**

| Retards p | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| AIC | 2,4896 | 2,4915 | 2,4916 | 2,4905 | 2,4940 | 2,4969 |
| BIC (SC) | 2,5102 | 2,5275 | 2,5430 | 2,5573 | 2,5763 | 2,5947 |

AIC et BIC sont minimaux pour **p = 1**, ce qui confirme le choix d'un VAR(1). Pour les modèles de volatilité (tableau 9), l'EGARCH a l'AIC le plus bas (4,1031 par observation contre 4,1053 pour le GARCH(1,1), soit environ 8 points d'AIC total), mais il échoue au diagnostic ARCH-LM : le GARCH(1,1) reste retenu.

**AIC contre performance hors échantillon.** L'AIC mesure l'ajustement sur l'apprentissage ; le RMSE de 6.2 mesure la prévision sur le test. Les deux classements concordent sur l'essentiel : le VECM et l'ARIMA creux sont en tête, le VAR et la marche aléatoire derrière. Sur le test, le

> VECM passe devant (RMSE de 18,69 contre 19,25), car son rappel vers l'équilibre de long terme agit surtout à 10 séances, alors que l'AIC ne juge que l'ajustement à 1 séance.

## 6.4 Test de Diebold-Mariano

Aucune différence avec la marche aléatoire n'est significative au seuil de 5 %. Le VECM s'en approche à h = 1 (p = 0,058). Entre réseaux, le GRU bat le LSTM de façon significative (p = 0,026), mais l'écart est de 0,03 cent : significatif statistiquement, négligeable en pratique. **Les gains apparents dans les RMSE ne sont donc pas des preuves de supériorité.** Ce résultat est cohérent avec l'hypothèse d'efficience des marchés : sur un marché liquide, l'information est intégrée dans le prix presque immédiatement.

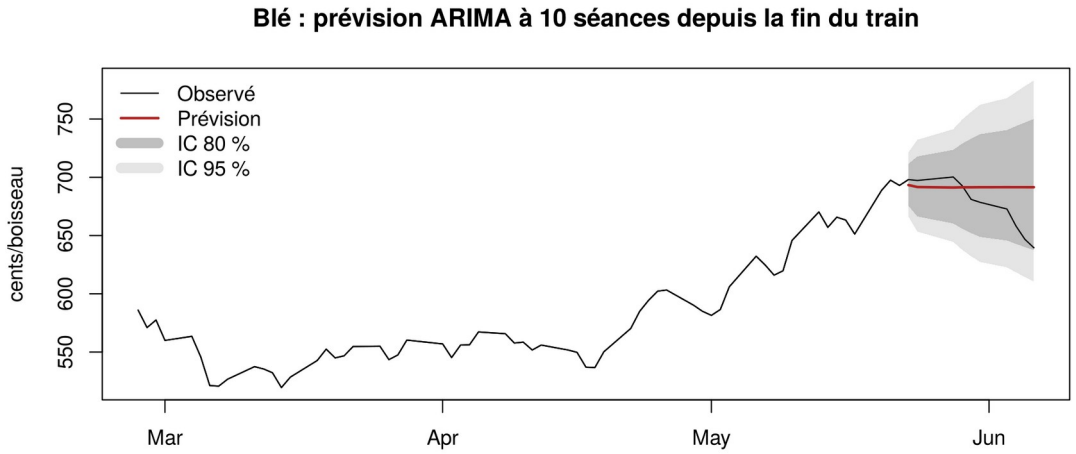



*Figure 12 — Prévision ARIMA à 10 séances depuis la fin du train, avec intervalles de prédiction à 80 % et 95 %.*

## 6.5 Prévision de la volatilité

Puisque le prix est peu prévisible, nous évaluons aussi la prévision de la **variance** du lendemain (volatilité) sur le test. Le GARCH(1,1) est réestimé toutes les 50 séances. Il est comparé à deux baselines : la volatilité historique sur 250 séances et la moyenne mobile exponentielle EWMA de RiskMetrics (λ = 0,94). La fonction de perte QLIKE est la référence pour comparer des prévisions de variance (plus elle est basse, mieux c'est).

**Tableau 18 — Prévision de la variance des rendements du blé à 1 séance (test)**

| Modèle | RMSE variance | QLIKE | DM p vs historique |
|---|---|---|---|
| GARCH(1,1) | 3,356 | 1,829 | 0,077 |
| EWMA RiskMetrics | 3,342 | 1,835 | 0,087 |
| Historique 250 séances | 3,419 | 1,847 | — |

Le GARCH a la meilleure perte QLIKE et bat la volatilité historique avec une p-value de 0,077 (significatif à 10 %, pas à 5 %). Le test couvre une période calme (2024-2025, sans grand choc), ce qui laisse peu d'occasions aux modèles dynamiques de montrer leur avantage.

## 6.6 Choix du meilleur modèle

> **Conclusion de l'évaluation.** Pour **prévoir le prix**, aucun modèle ne bat significativement la marche aléatoire. Pour un gestionnaire, le meilleur compromis est le **VECM** : le plus précis sur le test (−3,1 % de RMSE), la meilleure log-vraisemblance et un AIC équivalent au meilleur ARIMA (ΔAIC = 1,0), interprétable (relations de long terme blé–soja et maïs–soja) et quasi instantané à

> estimer. LSTM, GRU et hybride, plus coûteux et opaques, n'apportent rien de mesurable. Pour **gérer le risque**, le **GARCH(1,1)** est l'outil utile : il prévoit l'incertitude, ce qui sert à dimensionner une couverture ou un stock de sécurité.

# 7. Extensions

## 7.1 Corrélation dynamique blé–maïs (DCC-GARCH)

Un DCC-GARCH(1,1) (loi de Student multivariée) estime une corrélation qui varie dans le temps entre les rendements du blé et du maïs. Les paramètres DCC valent a = 0,009 et b = 0,984 : la corrélation évolue lentement et de manière très persistante.

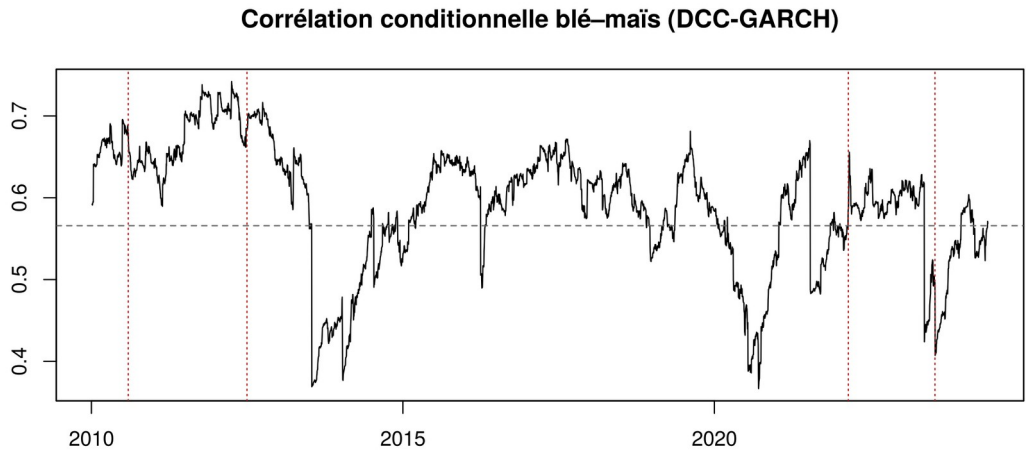



*Figure 13 — Corrélation conditionnelle blé–maïs (DCC-GARCH). Tirets : corrélation statique (0,57).*

La corrélation dynamique varie entre **0,37 et 0,74**, autour d'une corrélation statique de 0,57. Une seule corrélation constante masque donc de grandes différences selon les périodes. Pour un importateur qui voudrait se couvrir sur le blé avec un contrat sur le maïs (couverture croisée), l'efficacité de la couverture changerait fortement dans le temps.

## 7.2 Variables exogènes : dollar et pétrole

Le VAR est élargi aux rendements de l'indice du dollar et du pétrole WTI (5 variables, 1 retard selon le BIC). Ni le dollar (p = 0,69) ni le pétrole (p = 0,38) ne causent le blé au sens de Granger. Le RMSE passe de 19,308 à 19,306 cents, une différence non significative (DM p = 0,91 à h = 1, 0,36 à h = 10). À l'échelle journalière, ces variables n'ajoutent pas d'information prédictive.

## 7.3 Le site web AgriScope

**Présentation.** Les résultats de ce rapport sont aussi présentés sur un site web, AgriScope, à l'adresse https://agri-scope.syris.tn. Le site ne ré-estime aucun modèle : il reprend les résultats du notebook R (prix, tableaux, coefficients du GARCH, prévisions du LSTM et du GRU), si bien que ses chiffres sont ceux de ce rapport. Les données y sont figées au 30 décembre 2025.

**Contenu.** Une page retrace la démarche de l'étude avec ses tableaux et ses figures. Un laboratoire propose cinq expériences sur les données réelles : explorer les séries et leurs corrélations sur une période choisie, comparer les prix avant et après quatre chocs historiques, confronter les modèles avec le test de Diebold-Mariano, tenter de battre les prévisions du LSTM et

du GRU, et chiffrer le risque de prix pour le budget d'un importateur à partir du GARCH. Le rapport et la présentation y sont téléchargeables.

**Assistant.** Un assistant conversationnel, fondé sur un modèle de langage interrogé par l'API Groq, répond aux questions sur l'étude. Il ne peut citer un chiffre qu'après avoir appelé nos fonctions de calcul, et chaque nombre de sa réponse est comparé à leurs résultats. Ce contrôle est seulement numérique : il ne garantit pas la justesse du raisonnement, comme l'illustrent les erreurs de LLM analysées à la section 8.

**Limites.** Le jeu « humain contre machine » et le calcul du risque pour un importateur sont des démonstrateurs pédagogiques : leurs résultats ne figurent pas dans ce rapport et ne constituent pas un conseil d'achat.

# 8. IA générative et LLM

Le notebook R génère automatiquement trois prompts contenant les **vrais chiffres** de l'étude (fichier `outputs/prompts_LLM.md`, reproduit en annexe A) : hypothèses de modèles à partir des statistiques descriptives, vulgarisation pour un responsable de l'Office des Céréales, et recommandations d'achat simulées pour un importateur à partir des prévisions du 30 décembre 2025.

> **Règle suivie :** notre interprétation (8.1) est rédigée **avant** d'interroger le LLM. Prompts et réponses sont reproduits intégralement en annexes A et B.

**Protocole.** LLM interrogé : Claude (Anthropic, modèle Claude Fable 5.2), le 21 novembre 2026. Chaque prompt a été envoyé dans une conversation séparée, sans autre contexte : le LLM n'a vu ni les données, ni le code, ni ce rapport, et n'avait pas accès à Internet. Chaque chiffre de ses réponses a ensuite été recalculé à partir de nos tableaux.

## 8.1 Notre interprétation (rédigée avant le LLM)

Cette synthèse reprend les conclusions des sections 3 à 7, rédigées avant l'envoi des prompts.

- **Prix.** Les log-prix sont I(1) et les rendements n'ont presque aucune mémoire (seul le retard 4 ressort, $\varphi 4 = -0{,}044$). Aucun modèle ne bat significativement la marche aléatoire à 1-10 séances ; seul le VECM s'en détache (−3,1 % de RMSE), avec une p-value de Diebold-Mariano de 0,058 à 1 séance et de 0,17 à 10 séances, à la limite de la significativité.
- **Liens entre céréales.** Aucune causalité de Granger n'est significative à 5 %, alors que la corrélation le même jour atteint 0,57 entre blé et maïs : les chocs se transmettent dans la journée. À long terme, deux relations de cointégration lient les trois prix ; le blé corrige 0,92 % de son écart à l'équilibre par séance (demi-vie de 75 séances).
- **Risque.** La volatilité est très persistante ($\alpha 1 + \beta 1 = 0{,}986$, demi-vie de 51 séances) et ses pics coïncident avec les chocs de 2010, 2022 et 2023. Un GARCH(1,1) à loi de Student suffit et passe tous les diagnostics.
- **Réseaux de neurones.** LSTM, GRU et hybride n'apportent aucun gain mesurable et surapprennent dès les premières époques : il n'y a presque pas de signal à apprendre.
- **Décision.** Le moment d'achat ne peut pas être choisi à partir de l'historique des prix ; ce qui se gère, c'est le risque : achats étalés, couverture, stocks de sécurité.

## 8.2 Génération d'hypothèses de modèles (prompt 1)

**Réponse du LLM (annexe B).** Il propose, dans l'ordre : un ARMA d'ordres faibles sur les rendements (p et q de 0 à 2, choisis par BIC, avec le bruit blanc comme référence) ; un GARCH(1,1) à innovations de Student ; un GJR-GARCH ou un EGARCH pour tester une

asymétrie ; un test de Johansen suivi d'un VECM (rang 1 ou 2) ou, à défaut, d'un VAR avec causalité de Granger ; enfin un DCC-GARCH(1,1) pour les corrélations dynamiques.

**Convergence.** Ces familles sont celles que nous avons testées, et les ordres proposés (GARCH(1,1), VECM de rang 1 ou 2) coïncident avec ceux retenus. Le seul calcul de la réponse, la statistique t de la moyenne (0,006 / (2,029 / √3618) ≈ 0,18), est exact.

**Écarts relevés :**

- La grille ARMA limitée à 2 retards aurait manqué le seul effet détecté dans nos données, au retard 4 (φ4 = −0,044). Le LLM précise lui-même qu'il n'a vu ni ACF ni PACF.
- De l'asymétrie positive des rendements (0,39), il tire l'hypothèse d'un « effet de levier inversé ». Nos estimations ne la confirment pas : le GJR-GARCH a une log-vraisemblance quasi identique à celle du GARCH(1,1) (−7 421,4 dans les deux cas, tableau 9) et le test de biais de signe n'est pas significatif à 5 % (p = 0,060). L'asymétrie de la distribution des rendements et l'asymétrie de la réaction de la volatilité aux chocs sont deux propriétés différentes ; le LLM passe de l'une à l'autre, en la présentant toutefois comme une hypothèse à tester.
- Il avertit qu'une persistance α + β proche de 1 peut venir d'une rupture de régime. La remarque est fondée et nous ne l'avons pas testée (section 9.1).

## 8.3 Vulgarisation pour un décideur (prompt 2)

**Clarté.** La réponse tient en une page et n'emploie aucun terme technique. La marche aléatoire devient la règle la plus simple, « le prix dans dix séances sera celui d'aujourd'hui », et la p-value du VECM (0,17) devient un écart qui « apparaîtrait environ une fois sur six entre deux méthodes de même qualité » : c'est une lecture correcte d'une p-value.

**Exactitude.** Tous les chiffres cités figurent dans le prompt ou s'en déduisent par un calcul juste : 3 % d'erreur en moins pour le VECM (18,69 contre 19,28), une demi-vie d'« une cinquantaine de séances, soit environ deux mois et demi », et la conversion de 19 cents par boisseau en « environ 7 dollars par tonne » (19,28 × 36,74 / 100 = 7,08). La causalité de Granger est bien rendue : les variations passées « n'aident pas à anticiper », sans aucune idée de cause économique, et l'indice soja → maïs (p = 0,094) « n'atteint pas le seuil habituel de fiabilité ».

**Erreur relevée.** Pour un achat décalé de dix séances, le LLM recommande « une marge d'environ 14 dollars par tonne autour du prix du jour (deux fois l'erreur typique) ». Or le RMSE de 19,3 cents est une moyenne sur les horizons 1 à 10 ; à 10 séances exactement, il vaut 25,3 cents (section 6.2). La marge correspondante est de 2 × 25,3 = 50,6 cents, soit 18,6 dollars par tonne, et notre intervalle à 95 % (451 à 579 cents) donne −22 à +25 dollars par tonne. La marge du LLM sous-estime donc le risque d'un quart à 40 %. L'erreur vient en partie de notre prompt, dont l'intitulé « Précision des prévisions à 10 séances » était ambigu : le LLM ne pouvait pas savoir ce que le tableau agrège, mais il n'a pas non plus signalé l'ambiguïté.

**Simplification.** « Surveiller un marché ne donne donc pas d'avance sur l'autre » est exact pour la prévision, mais le texte ne dit pas que les prix bougent ensemble le même jour (corrélation de 0,57), information absente du prompt. Un décideur pourrait en conclure à tort que les trois marchés sont indépendants.

## 8.4 Recommandations simulées et leurs limites (prompt 3)

**Stratégie proposée.** Constatant que la prévision est quasi plate (510,5 à 511,0 cents), le LLM en déduit qu'attendre n'apporte « aucun gain espéré, uniquement du risque ». Il propose une approche mixte : sécuriser tout de suite environ la moitié du besoin (achat physique ou contrats à terme), étaler le reste en deux ou trois tranches à dates fixes, et préférer des options d'achat si le

budget est plafonné. Cette logique, gérer le risque plutôt que parier sur le prix, rejoint notre conclusion (8.1).

**Limites citées par le LLM lui-même.** La part de 50 % est « un choix de prudence de ma part, pas un résultat tiré des données » ; les intervalles supposent des chocs à peu près normaux alors que le blé connaît des sauts ; le CBOT cote un blé tendre américain, et le prix payé dépend de l'origine, de la qualité, du fret et du change (risque de base) ; les contrats à terme exigent des appels de marge, de la trésorerie en devises et un mandat légal ; dix séances sont courtes face aux délais d'un appel d'offres ; enfin, « c'est un exercice de simulation, pas un conseil financier ».

**Apport inattendu : une incohérence détectée dans notre tableau.** Le LLM remarque que nos intervalles impliquent environ 2 % d'écart-type par séance, soit près de 32 % annualisés, alors que la volatilité affichée est de 21 à 22 %. Vérification faite, il a raison : l'intervalle vient de l'ARIMA, qui suppose une variance constante égale à sa moyenne de 2010-2024 (écart-type de 2,03 % par séance), tandis que la volatilité vient du GARCH, qui décrit le régime calme de fin 2025. Avec la volatilité du GARCH, l'intervalle à 10 séances serait d'environ 471 à 555 cents, au lieu de 451 à 579.

Limites que nous ajoutons ou précisons :

- Le test de Diebold-Mariano ne montre aucune supériorité sur la marche aléatoire : une prévision de prix ponctuelle (environ 511 cents) n'a pas de valeur pour décider du moment d'acheter.
- L'intervalle à 95 % à 10 séances (451 à 579 cents, soit ±13 %) est trop large en période calme, pour la raison ci-dessus, et sans doute trop étroit en cas de crise, car les résidus ne sont pas normaux.
- Le prix de Chicago n'est pas le prix payé par la Tunisie : il faut ajouter le fret, la prime de qualité et l'origine du blé (souvent mer Noire ou Europe), et le payer en dollars, d'où un risque de change sur le dinar.
- Se couvrir avec des contrats à terme coûte de l'argent (marges, appels de marge, frais) et laisse un risque de base (écart entre prix local et prix de Chicago).
- La part de 50 % et le nombre de tranches ne reposent sur aucun calcul : une règle de couverture demanderait de chiffrer ce que coûte à l'Office un dépassement de budget par rapport à une baisse manquée.
- Le LLM affirme que le rapport USDA de janvier « tombe habituellement dans cette fenêtre » : ce fait, absent du prompt, est plausible mais nous ne l'avons pas vérifié.
- Exercice académique simulé : ce n'est pas un conseil d'investissement.

## 8.5 Comparaison humain / LLM et erreurs détectées

**Tableau 19 — Grille de comparaison entre notre interprétation et celle du LLM**

| Point analysé | Notre interprétation | Interprétation du LLM | Verdict |
|---|---|---|---|
| Prévisibilité du prix | Marche aléatoire non battue (DM) | Aucun modèle ne fait « de façon fiable mieux que la règle la plus simple » ; VECM : « une piste, pas une preuve » | Accord |
| Causalité de Granger | Aucune à 5 % ; liens contemporains | Aucune avance d'un marché sur l'autre ; soja → maïs sous le seuil | Accord ; lien du même jour non mentionné (absent du prompt) |
| Persistance GARCH | Demi-vie de 51 séances | « Une cinquantaine de séances » ; une persistance proche de 1 peut cacher une rupture de régime | Accord ; réserve utile, non testée par nous |
| Recommandation | Pas de timing possible ; couverture du risque | Sécuriser la moitié, étaler le reste, options si budget | Accord sur la logique ; marge sous- |

| Point analysé | Notre interprétation | Interprétation du LLM | Verdict |
|---|---|---|---|
| | | plafonné ; marge de 14 $/t | estimée (18,6 $/t) ; part de 50 % arbitraire |
| Chiffres cités | Ceux de l'étude uniquement | Chiffres du prompt ou déduits par calcul ; conversions exactes | Aucune invention ; un fait externe non vérifié (date du rapport USDA) |
| Choix des modèles | AR creux au retard 4, GARCH(1,1), VECM de rang 2 | ARMA d'ordre ≤ 2, GARCH(1,1) Student, GJR ou EGARCH, VECM de rang 1 ou 2, DCC | Accord sur les familles ; retard 4 manqué ; effet de levier non confirmé |

**Bilan de la recherche d'erreurs.** Chaque chiffre des trois réponses a été recalculé à partir de nos tableaux. Aucune p-value ni aucun chiffre n'a été inventé ; la causalité de Granger n'est pas présentée comme une causalité économique ; aucun réseau de neurones n'est déclaré « meilleur » ; la conversion des cents par boisseau en dollars par tonne est exacte. Les écarts relevés sont d'une autre nature :

- une marge de sécurité sous-estimée, par lecture d'un RMSE moyen comme un RMSE à 10 séances (prompt 2) ;
- une hypothèse d'effet de levier inversé tirée de l'asymétrie, non confirmée par le GJR-GARCH (prompt 1) ;
- une grille ARMA trop courte pour voir le retard 4 (prompt 1) ;
- une part de 50 % sans fondement quantitatif, que le LLM reconnaît lui-même (prompt 3) ;
- un fait extérieur au prompt, non vérifié : la date du rapport USDA de janvier (prompt 3).

À l'inverse, le LLM a relevé deux faiblesses de notre propre étude : l'incohérence entre l'intervalle de l'ARIMA et la volatilité du GARCH (8.4), et le risque d'une persistance surestimée en présence de ruptures (8.2). **Limite de l'exercice :** un seul LLM et une seule réponse par prompt ; une autre exécution donnerait un texte différent, et ces constats ne se généralisent pas à d'autres modèles.

# 9. Discussion critique

## 9.1 Limites des modèles et des données

- **Séries continues Yahoo :** les changements de contrat créent des sauts artificiels (−26,9 % sur le maïs le 15/07/2013). Ils gonflent le kurtosis du maïs et peuvent biaiser le VAR et le DCC. Une série « ajustée du roll » (rapport ou différence entre contrats) serait plus propre, mais Yahoo ne la fournit pas.
- **Hypothèses violées :** la normalité des résidus est rejetée partout, donc les intervalles gaussiens de l'ARIMA sont trop étroits en crise. Le test du portemanteau du VAR ($p = 0{,}014$) montre une petite autocorrélation résiduelle.
- **Intervalles de prévision :** ceux de l'ARIMA supposent une variance constante (2,03 % par séance, soit 32 % annualisés), alors que le GARCH prévoit 21 à 22 % fin 2025. Ils sont donc trop larges en période calme et trop étroits en crise ; des intervalles construits sur la variance du GARCH seraient plus cohérents (incohérence relevée par le LLM, section 8.4).
- **Stabilité :** l'échantillon contient des régimes très différents (2010-2012 agité, 2014-2020 calme, 2022 extrême). Des paramètres fixes sur 14 ans peuvent mal représenter le régime actuel ; un test de rupture (Chow, Bai-Perron) serait une suite naturelle. Une persistance GARCH proche de 1 (0,986) peut d'ailleurs refléter ces changements de régime plutôt qu'une mémoire réelle de la volatilité : la demi-vie de 51 séances est peut-être surestimée (remarque du LLM, section 8.2).

- **Période de test calme :** le test (mai 2024 à décembre 2025) ne contient aucun choc majeur. Les conclusions sur les crises reposent donc sur l'échantillon d'apprentissage.
- **Réseaux de neurones :** environ 3 000 fenêtres seulement et un signal quasi nul ; leurs résultats dépendent de la graine (écart de 0,04 à 0,07 point). Ils n'ont pas d'intervalle de prévision natif.
- **Concordance R / Python :** les coefficients GARCH, l'ordre du VAR et la statistique de Johansen sont identiques dans les deux langages (annexe C). L'ordre ARIMA n'est pas une vérification indépendante : le notebook Python reprend l'ordre trouvé par auto.arima en R et ne refait pas la recherche automatique. Les statistiques ADF diffèrent, car le nombre de retards choisi automatiquement n'est pas le même, et le test de Granger maïs → blé diffère car R teste la paire seule ($p = 0,28$) et Python la conditionne au soja ($p = 0,09$). La conclusion est la même dans les deux cas : pas de causalité au seuil de 5 %.

### 9.2 Apport et limites de l'IA générative pour les séries temporelles agricoles

Les LLM sont utiles pour reformuler des résultats techniques à destination de décideurs, proposer des pistes de modèles et aider à rédiger le code. Ils ne peuvent cependant ni lire les données ni garantir l'exactitude d'un chiffre, ont tendance à transformer une corrélation ou une causalité de Granger en causalité réelle, et connaissent mal le contexte local (prix réellement payés par la Tunisie, logistique, subventions). Ils doivent donc rester un outil d'assistance, vérifié à chaque étape par les tests statistiques et par l'expertise du domaine.

**Retour d'expérience (section 8).** Gain de temps : les trois réponses sont arrivées en une à deux minutes chacune. Gain de clarté : la vulgarisation du prompt 2 est utilisable telle quelle une fois la marge de sécurité corrigée, et le LLM a repéré deux faiblesses réelles de notre étude (8.5). Ce qu'il n'a pas su faire : voir le retard 4 sans l'ACF, trancher une hypothèse (l'effet de levier) qu'il ne peut que proposer, et remarquer l'ambiguïté de notre propre prompt. Aucun de ses chiffres n'a été repris sans être recalculé à partir de nos tableaux : le LLM s'est montré un bon relecteur et un bon vulgarisateur, pas une source de résultats.

## 10. Conclusion

Les réponses aux cinq questions de recherche sont les suivantes : (1) les log-prix des trois céréales sont I(1) et leurs rendements stationnaires ; (2) aucune causalité de Granger n'est significative à 5 %, les chocs se transmettant le jour même ; (3) il existe deux relations de cointégration, dont une liant le blé au soja avec une élasticité proche de 1 et une demi-vie d'ajustement de 75 séances ; (4) la volatilité du blé est très persistante (demi-vie de 51 séances) et ses pics correspondent aux chocs de 2010, 2022 et 2023 ; (5) ni le LSTM ni le GRU ne battent la marche aléatoire.

Pour la Tunisie, le message est double. Il est illusoire de vouloir « deviner » le bon moment d'acheter à partir de l'historique des prix. En revanche, le risque (la volatilité) est mesurable et prévisible, ce qui justifie des outils de gestion du risque : couverture, stocks de sécurité, achats étalés. Prolongements possibles : intégrer des données fondamentales (rapports USDA, météo des zones de production), des données textuelles d'actualité (détection d'événements par embeddings), et le prix réellement payé par l'Office des Céréales.

## Bibliographie


- Hyndman, R. J. et Athanasopoulos, G. (2021). *Forecasting: Principles and Practice*, 3e édition. OTexts.

- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59(6), 1551-1580.
- Engle, R. (2002). Dynamic conditional correlation. *Journal of Business & Economic Statistics*, 20(3), 339-350.
- Diebold, F. X. et Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253-263.
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175.
- Hochreiter, S. et Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780.
- Sources de presse (repérées par recherche web le 21 novembre 2026). Tunisie : Hamza, A. (29 août 2026), « Tunisie : 200 000 tonnes de céréales achetées, une dépendance aux importations toujours structurelle », Webmanagercenter ; African Manager (9 janvier 2024), « La Tunisie mange son blé vert en importations ». Chocs de 2010-2011 : Bloomberg (15 juillet 2010), « Wheat Surges Most in 19 Months as Drought Hurts Russia Crop » ; Bloomberg (8 octobre 2010), « Corn, Soybean, Wheat Futures Surge in Chicago After Crop Report » ; Fortune (8 octobre 2010), « Corn goes nuts » ; The Western Producer (2011), « Bearish USDA report caps bad month for crop futures ».
- Packages R : forecast, urca, vars, rugarch, rmgarch, quantmod, tseries, prophet. Python : statsmodels, arch, PyTorch.

# Annexe A — Prompts générés avec les chiffres de l'étude

## Prompt 1 : hypothèses de modèles

```
Voici les statistiques des rendements logarithmiques journaliers (en %) du contrat à terme sur le
blé CBOT (ZW=F), du 2010-01-05 au 2024-05-22 (3618 observations) : moyenne = 0.006, écart-type =
2.029, asymétrie = 0.391, kurtosis = 7.16. Tests de stationnarité :

               ADF_stat ADF_vc5 KPSS_stat KPSS_vc5       Conclusion
log-prix Ble     -2.715   -2.86     1.706    0.463 Non stationnaire
log-prix Mais    -1.951   -2.86     1.684    0.463 Non stationnaire
log-prix Soja    -1.974   -2.86     1.594    0.463 Non stationnaire
rendement Ble   -42.291   -2.86     0.045    0.463     Stationnaire
rendement Mais  -44.051   -2.86     0.075    0.463     Stationnaire
rendement Soja  -42.599   -2.86      0.07    0.463     Stationnaire

Quelles familles de modèles de séries temporelles me conseilles-tu de tester, avec quels ordres, et
pourquoi ?
```

## Prompt 2 : vulgarisation

```
Explique en une page, à un responsable de l'Office des Céréales tunisien qui n'est pas statisticien,
les résultats suivants d'une étude des prix mondiaux du blé.

Précision des prévisions à 10 séances (prix en cents par boisseau) :
                        RMSE   MAPE DM_vs_marche_h10
Marche_aleatoire    19.2764 2.5530               NA
ARIMA_manuel        19.2525 2.5446           0.9222
ARIMA_auto          19.2515 2.5448           0.9327
VAR                 19.3077 2.5593           0.3077
VECM                18.6880 2.5000           0.1704
LSTM                19.3208 2.5554           0.3087
GRU                 19.2868 2.5529           0.6684
Hybride_ARIMA_LSTM  19.3170 2.5572           0.4729
Prophet_10_origines 59.5261 9.5370               NA

Causalité de Granger entre rendements :
  Cause Effet     F      p
1  Mais   Ble 1.190 0.2755
2   Ble  Mais 0.001 0.9811
3  Soja   Ble 1.140 0.2858
4   Ble  Soja 0.117 0.7321
5  Soja  Mais 2.812 0.0937
6  Mais  Soja 0.042 0.8379

GARCH(1,1) : persistance = 0.9864, demi-vie = 50.6 séances.
```

## Prompt 3 : recommandations simulées

```
Voici la prévision du prix du blé CBOT (cents par boisseau) pour les 10 séances suivant le 2025-12-
30, avec intervalle à 95 % et volatilité prévue :

   Seance  Prix IC95_bas IC95_haut Volatilite_annualisee
1       1 510.5    490.6     531.2                 20.7
2       2 510.7    482.4     540.6                 20.9
3       3 511.0    476.5     547.9                 21.0
4       4 511.1    471.4     554.0                 21.2
5       5 511.1    467.3     558.9                 21.4
6       6 511.1    463.6     563.3                 21.6
7       7 511.0    460.2     567.5                 21.7
8       8 511.0    457.1     571.4                 21.9
9       9 511.0    454.1     575.1                 22.0
10     10 511.0    451.3     578.7                 22.2

Propose une stratégie d'achat simulée pour un importateur public de blé (acheter maintenant, étaler
les achats, ou se couvrir avec des contrats à terme). Liste ensuite toi-même les limites et les
risques de ces recommandations.
```

# Annexe B — Réponses intégrales du LLM

**Modèle et date.** Claude (Anthropic), modèle Claude Fable 5.2, interrogé le 21 novembre 2026. Chaque prompt de l'annexe A a été envoyé dans une conversation séparée, sans accès aux

données, au code, à ce rapport ni à Internet. Deux consignes encadraient chaque prompt : avant, répondre directement, sans utiliser d'outil ; après, une consigne de forme, « Réponds en texte simple (paragraphes et listes à tirets, sans tableau, sans titres ni gras), en 400 mots maximum. » Les réponses sont reproduites sans modification.

### Réponse au prompt 1 : hypothèses de modèles

Tes tests sont nets : les log-prix sont I(1) et les rendements I(0) (ADF rejette, KPSS ne rejette pas). On modélise donc les rendements, soit d = 1 sur les log-prix, sans différencier davantage. Je testerais, dans cet ordre :

- ARMA(p,q) sur les rendements, avec p et q de 0 à 2, choisis par BIC. Attends-toi à des ordres très faibles : la moyenne n'est pas significative ($t \approx 0{,}006 / (2{,}029/\sqrt{3618}) \approx 0{,}18$) et les rendements de contrats à terme sont en général presque non autocorrélés. Garde le bruit blanc, ARMA(0,0), comme référence.
- GARCH(1,1) avec innovations de Student. La kurtosis de 7,16 est bien au-delà de la valeur normale, qu'elle soit brute ou en excès : elle signale des queues épaisses et, probablement, des grappes de volatilité. Confirme-le par un test ARCH-LM et un Ljung-Box sur les rendements au carré. L'ordre (1,1) suffit presque toujours : ne monte à (1,2) ou (2,1) que si les résidus standardisés au carré restent autocorrélés.
- GJR-GARCH(1,1) ou EGARCH(1,1), avec une Student asymétrique. L'asymétrie positive (0,391) indique des hausses extrêmes plus marquées que les baisses. Sur les matières premières agricoles, on observe souvent un effet de levier inversé, où les chocs haussiers (craintes de pénurie) accroissent davantage la volatilité. Le signe du coefficient d'asymétrie tranchera.
- En multivarié, puisque les trois log-prix sont I(1) : test de cointégration de Johansen. S'il y a cointégration, un VECM (rang 1 ou 2, un à deux retards selon BIC) ; sinon un VAR sur les rendements, avec causalité de Granger. Un DCC-GARCH(1,1) complète pour les corrélations dynamiques entre blé, maïs et soja.

Deux précautions :

- La période inclut 2010-2012 (sécheresses) et 2022 (guerre en Ukraine). Si $\alpha + \beta$ est très proche de 1, soupçonne une rupture de régime avant de conclure à une volatilité quasi intégrée.
- Départage les modèles par BIC et par prévision hors échantillon, et vérifie les résidus standardisés (Ljung-Box sur les niveaux et sur les carrés).

Je n'ai vu ni ACF/PACF ni test ARCH : ces ordres sont des points de départ, à confirmer sur tes données.

### Réponse au prompt 2 : vulgarisation

À dix séances, aucun des modèles testés ne fait de façon fiable mieux que la règle la plus simple : « le prix dans dix séances sera celui d'aujourd'hui ». La direction du marché à court terme reste imprévisible ; son niveau de risque, lui, ne l'est pas.

Précision des prévisions :

- La règle simple se trompe typiquement de 19 cents par boisseau, environ 7 dollars par tonne, soit en moyenne 2,5 % du prix. Des écarts deux fois plus grands restent possibles.
- ARIMA, VAR et les réseaux de neurones (LSTM, GRU, hybride) obtiennent entre 19,25 et 19,32 : des différences sans portée pratique.
- Seul le VECM, qui relie le blé aux autres prix sur le long terme, fait un peu mieux : 18,7, soit 3 % d'erreur en moins. Mais un écart de cette taille apparaîtrait environ une fois sur six entre deux méthodes de même qualité : c'est une piste, pas une preuve.
- Prophet est à écarter : erreur plus de trois fois supérieure (9,5 %), même s'il n'a été évalué que sur dix dates.

Liens entre blé, maïs et soja : les variations passées du maïs ou du soja n'aident pas à anticiper celles du blé, ni l'inverse. Le seul indice, un léger effet du soja sur le maïs, n'atteint pas le seuil habituel de fiabilité. Surveiller un marché ne donne donc pas d'avance sur l'autre.

Risque : l'agitation du marché est très persistante. Après un choc, il faut une cinquantaine de séances, soit environ deux mois et demi, pour que le surcroît de fluctuations diminue de moitié. Un marché nerveux le reste longtemps, un marché calme aussi.

Conséquences pour l'Office :

- Ne pas caler les achats sur une prévision de prix à deux semaines : rien ne permet de dire si attendre sera payant.
- Raisonner en marge de risque : pour un achat décalé de dix séances, une marge d'environ 14 dollars par tonne autour du prix du jour (deux fois l'erreur typique) couvre la plupart des cas en période normale.
- Après un choc, élargir cette marge pendant plusieurs mois ; étaler les achats ou se couvrir prend alors tout son sens.

Limites : ces résultats concernent le prix mondial en dollars à dix séances. Ils ne disent rien du fret, du taux de change du dinar ni des horizons plus longs, et l'avantage du VECM reste à confirmer.

### Réponse au prompt 3 : recommandations simulées

La prévision est quasi plate (510,5 à 511,0) : le modèle ne donne aucun signal de direction, seulement une incertitude qui s'élargit. Attendre n'apporte donc aucun gain espéré, uniquement du risque. La décision relève de la gestion du risque, pas d'un pari sur le prix.

Stratégie simulée proposée : une approche mixte.

- Sécuriser tout de suite environ la moitié du besoin de la période, par achat physique si un appel d'offres est prêt, sinon par achat de contrats à terme sur l'échéance la plus proche de la livraison.
- Étaler le reste en deux ou trois tranches égales, à dates fixes sur les dix séances, sans chercher le meilleur jour.
- Si le budget est plafonné, préférer des options d'achat aux contrats à terme sur la part restante : le prix maximal est borné et une baisse reste profitable, au prix d'une prime.

Justification : pour un importateur public, une hausse subie (dépassement budgétaire, sécurité d'approvisionnement) coûte plus cher qu'une baisse manquée. Or la borne haute à dix séances atteint 578,7, soit environ +13 %. Tout acheter aujourd'hui concentrerait en revanche le risque sur un seul prix.

Limites et risques :

- La part de 50 % est un choix de prudence de ma part, pas un résultat tiré des données.
- Incohérence à vérifier : les intervalles impliquent environ 2 % d'écart-type par séance, soit près de 32 % annualisés, alors que la volatilité affichée est de 21 à 22 %. Avec 21 %, la fourchette à dix séances serait plutôt d'environ 471 à 555.
- Les intervalles supposent des chocs à peu près normaux. Le blé connaît des sauts : météo, mer Noire, restrictions à l'exportation, rapports USDA (celui de janvier tombe habituellement dans cette fenêtre).
- Risque de base : le CBOT cote du blé tendre américain. Le prix réellement payé dépend de l'origine, de la qualité, des primes, du fret et du taux de change.
- Les contrats à terme exigent des appels de marge, de la trésorerie en devises et un mandat légal. Une couverture perdante en cas de baisse est politiquement visible.
- Dix séances, c'est court face aux délais réels d'un appel d'offres et d'un affrètement.
- Je n'ai vu ni le modèle, ni ses données, ni un test rétrospectif : c'est un exercice de simulation, pas un conseil financier.

# Annexe C — Concordance R / Python (socle classique)

**Tableau 20 — Principaux résultats du socle classique dans les deux langages**

| Indicateur | R | Python |
|---|---|---|
| ADF stat log-prix blé | -2.715 | -2.568 |
| ADF stat rendement blé | -42.291 | -31.367 |
| ARIMA auto ordre | 4,1,0 | 4,1,0 (repris de R) |
| Ordre VAR (BIC) | 1 | 1 |
| Granger Mais->Ble p | 0.2755 | 0.0891 |
| Johansen trace r=0 | 49.83 | 49.77 |
| GARCH alpha1 | 0.0576 | 0.0576 |
| GARCH beta1 | 0.9288 | 0.9287 |
| GARCH shape | 9.22 | 9.23 |

# Annexe D — Reproductibilité

- **Ordre d'exécution :** ouvrir `notebook/projet_cereales.Rmd` dans RStudio et cliquer sur « Knit ». Le notebook appelle automatiquement `python/02_lstm_gru.ipynb` (Jupyter requis). Exécuter ensuite `python/01_socle_classique.ipynb` pour la version Python du socle classique.
- **Graines :** set.seed(2026) en R ; numpy et torch avec la graine 2026 (plus 2027 et 2028 pour la stabilité des réseaux).
- **Versions :** R 4.3.3 (forecast 8.21.1, urca 1.3.3, vars 1.6.1, rugarch 1.5.6, rmgarch 1.4.3, quantmod 0.4.26, tseries 0.10.55, prophet 1.0) ; Python 3.13 (PyTorch 2.15.0, statsmodels 0.15.0, arch 9.0.0, pandas 3.0.5). `sessionInfo()` est imprimé à la fin du notebook.
- **Données :** `data/cereales_yahoo.csv` (téléchargé le 15/11/2026 avec quantmod) ; aucun chemin absolu dans le code.