

Analyses exploratoires des stratégies de migration et d'hivernage de la Tourterelle des bois (*Streptopelia turtur*)

– Comparaison entre populations européennes et nord-africaines
Basée sur un workflow d'analyse de données ARGOS –



Assia El Abied

Travail réalisé sous la direction de
Raphaël MUSSEAU (PhD)
BioSphère Environnement
Octobre 2026

Analyses exploratoires des stratégies de migration et d'hivernage de la Tourterelle des bois (*Streptopelia turtur*). Comparaison entre populations européennes et nord-africaine basée sur un workflow d'analyse de données ARGOS

Assia El Abied

Travail réalisé sous la direction de
Raphaël MUSSEAU (PhD)
BioSphère Environnement
Octobre 2026

Citation recommandée du document : El Abied, A. (2026). *Analyses exploratoires des stratégies de migration et d'hivernage de la Tourterelle des bois (Streptopelia turtur) : comparaison entre populations européennes et nord-africaines basée sur un workflow d'analyse de données ARGOS*. Rapport de stage réalisé sous la direction de Raphaël Musseau. BioSphère Environnement & Institut Scientifique de Rabat, 70 p.

Résumé

La Tourterelle des bois (*Streptopelia turtur*) est une espèce migratrice transsaharienne en fort déclin, dont la conservation nécessite une meilleure compréhension des stratégies de migration et d'hivernage. Ce travail avait pour objectif de comparer les déplacements d'individus nichant en France et au Maroc, tout en développant un workflow reproductible de traitement et d'analyse de données de télémétrie satellitaire ARGOS. Une synthèse préalable des connaissances a également permis de replacer cette comparaison dans le contexte taxonomique et écologique des populations européennes et nord-africaines.

Les localisations ARGOS ont été filtrées afin de ne conserver que les classes de meilleure qualité, puis structurées pour caractériser la traversée du Sahara et identifier les secteurs successivement utilisés au Sahel. Les analyses ont porté sur 12 individus ayant réalisé une migration transsaharienne complète, dont neuf équipés en France et trois au Maroc. Les conditions de végétation associées aux déplacements intra-hivernaux ont été évaluées à partir du NDVI issu du produit satellitaire MODIS MOD13Q1.

Les individus marocains ont franchi la latitude 30° N en moyenne neuf jours avant les individus français, respectivement autour du 12 et du 21 septembre. Cette différence était statistiquement significative ($p = 0,040$). Une avance de sept jours était encore observée au passage de 20° N, mais sans différence significative ($p = 0,263$). La durée de traversée entre 30° N et 20° N ne différait pas entre les deux groupes ($p = 0,864$), suggérant un décalage du calendrier migratoire plutôt qu'une différence de vitesse de progression dans le Sahara.

Au total, 30 secteurs hivernaux (clusters) ont été identifiés pour 11 suivis individu-hiver. Les oiseaux ont utilisé en moyenne 2,73 secteurs successifs par hivernage, confirmant le caractère mobile de leur séjour sahélien. Parmi les 19 transitions intra-hivernales documentées, 14, soit 73,7 %, conduisaient vers un secteur présentant un NDVI supérieur à celui du secteur quitté. Le NDVI moyen passait ainsi de 0,256 au départ à 0,377 à l'arrivée ; cette différence demeurait significative en ramenant l'analyse à une valeur moyenne par individu ($n = 10$; $p = 0,008$). Une diminution du NDVI pendant le séjour était par ailleurs observée dans huit des neuf secteurs suffisamment documentés. Ces résultats sont compatibles avec l'hypothèse selon laquelle les Tourterelles des bois ajustent leurs déplacements hivernaux aux variations de la végétation sahélienne.

Mots-clés : Tourterelle des bois ; *Streptopelia turtur* ; migration transsaharienne ; télémétrie ARGOS ; hivernage sahélien ; déplacements intra-hivernaux ; NDVI.

Abstract

The European Turtle Dove (*Streptopelia turtur*) is a trans-Saharan migratory species undergoing severe decline, whose conservation requires a better understanding of its migration and wintering strategies. This study aimed to compare the movements of individuals breeding in France and Morocco while developing a reproducible workflow for processing and analysing ARGOS satellite telemetry data. A preliminary review of current knowledge also placed this comparison within the taxonomic and ecological context of European and North African populations.

ARGOS locations were filtered to retain only the most accurate location classes and then processed to characterise the Sahara crossing and identify the areas successively used in the Sahel. Analyses included 12 individuals that completed a full trans-Saharan migration, nine tagged in France and three in Morocco. Vegetation conditions associated with intra-winter movements were assessed using NDVI data derived from the MODIS MOD13Q1 satellite product.

Moroccan individuals crossed 30° N an average of nine days earlier than French individuals, around 12 and 21 September, respectively. This difference was statistically significant ($p = 0.040$). A seven-day advance was still observed at 20° N, although the difference was not statistically significant ($p = 0.263$). The duration of the crossing between 30° N and 20° N did not differ between the two groups ($p = 0.864$), suggesting a shift in migration timing rather than a difference in the rate of progression across the Sahara.

A total of 30 wintering areas (clusters) were identified across 11 individual wintering periods. Birds used an average of 2.73 successive areas per winter, confirming the mobile nature of their stay in the Sahel. Of the 19 documented intra-winter transitions, 14 (73.7%) led to an area with a higher NDVI than the area that had been left. Mean NDVI thus increased from 0.256 at departure to 0.377 at arrival; this difference remained significant when the analysis was reduced to one mean value per individual ($n = 10$; $p = 0.008$). A decline in NDVI during occupancy was also observed in eight of the nine areas for which sufficient temporal data were available. These findings are consistent with the hypothesis that European Turtle Doves adjust their winter movements in response to changes in Sahelian vegetation conditions.

Keywords: European Turtle Dove; *Streptopelia turtur*; trans-Saharan migration; ARGOS telemetry; Sahelian wintering; intra-winter movements; NDVI.

REMERCIEMENTS

Au terme de ce travail, je souhaite remercier les institutions et les personnes qui ont contribué à sa réalisation.

Je remercie tout particulièrement le Dr Raphaël Musseau, directeur de l'association BioSphère Environnement, pour avoir assuré la direction scientifique de ce travail. Sa confiance, ses conseils et son accompagnement tout au long du projet ont été essentiels à la conduite de cette étude et à l'élaboration de ce mémoire.

Je remercie également le Dr Sidi Imad Cherkaoui, de l'Institut Scientifique de Rabat, pour m'avoir permis de réaliser ce stage au sein de l'Institut.

Enfin, j'adresse mes remerciements à toutes les personnes qui, par leurs échanges, leurs conseils ou leur soutien logistique et moral, ont contribué à l'aboutissement de ce travail.

SOMMAIRE

I - INTRODUCTION	1
II - LA TOURTERELLE DES BOIS : TAXONOMIE ET ÉCOLOGIE COMPARÉE.....	2
II.1 - Taxonomie et différenciation des sous-espèces de la Tourterelle des bois	2
II.1.1 - Position taxonomique et phylogénétique	2
II.1.2 - Caractérisation des sous-espèces reconnues	3
II.1.3 - Portée et limites de la classification sous-spécifique	6
II.2 - Écologie comparée des populations reproductrices d'Europe occidentale et du Maghreb	7
II.2.1 - Distribution et état des populations	7
II.2.2 - Migration prénuptiale et arrivée sur les sites de reproduction	8
II.2.3 - Chronologie de la reproduction	8
II.2.4 - Migration postnuptiale	9
II.2.5 - Hivernage	10
II.2.6 - Reproduction et causes d'échec	11
II.2.7 - Habitats de reproduction	12
II.2.8 - Régime alimentaire et habitats d'alimentation	13
II.2.9 - Pressions et facteurs limitants.....	14
II.2.10 - Limites de la comparaison.....	15
II.2.11 - Conclusion	15

III - MATERIEL ET METHODES.....	18
III.1 - Capture et déploiement des balises.....	18
III.2 - Nettoyage des données	19
III.3 - Identification des individus migrateurs	20
III.4 - Analyse de la traversée saharienne	21
III.5 - Conditions d'hivernage.....	22
IV - RESULTATS	25
IV.1 - Traversée du Sahara : phénologie et durée	25
IV.2 - Conditions d'hivernage.....	28
V - DISCUSSION.....	30
V.1 - Traversée du Sahara : phénologie et durée	30
V.2 - Conditions d'hivernage.....	31
VI - CONCLUSION.....	32
VII - REFERENCES	34
VIII - ANNEXES	37

Liste des figures

Figure 1 : Carte des oiseaux ayant franchi le Sahara.....	26
Figure 2 : Clusters définis en région sahélienne.....	28
Figure 3 : Comparaison du NDVI entre secteurs de départ et d'arrivée pour 19 transitions intra-hivernales concernant dix Tourterelles des bois.....	29

Liste des tableaux

Tableau 1: Synthèse des individus ayant franchi le Sahara	25
---	----

Liste des annexes

Annexe 1 : Script R de nettoyage et filtrage des données ARGOS	35
Annexe 2 : Script R de clustering spatio-temporel des haltes migratoires	40
Annexe 3 : Script R d'analyse de la traversée saharienne	48
Annexe 4 : Script R de délimitation des zones d'hivernage	52
Annexe 5 : Script R d'extraction NDVI.....	60

I - INTRODUCTION

La compréhension des stratégies de migration et d'hivernage constitue un enjeu central en biologie de la conservation pour les espèces migratrices à large distribution. Chez la Tourterelle des bois (*Streptopelia turtur*), espèce en fort déclin à l'échelle européenne (Hanane, 2017), les facteurs limitants ne se répartissent pas uniquement sur les sites de reproduction, mais s'expriment tout au long du cycle annuel, en particulier lors de la migration transsaharienne et sur les zones d'hivernage au Sahel (Eraud et al., 2013 ; Lormée et al., 2016).

Dans ce contexte, l'analyse comparative de populations se reproduisant dans des régions géographiques distinctes constitue un levier essentiel pour mieux comprendre les mécanismes écologiques sous-jacents. Les individus se reproduisant en Europe de l'Ouest (en France) et ceux se reproduisant en Afrique du Nord (Maroc) occupent des contextes environnementaux différents en période de reproduction (Hanane & Baamal, 2011), mais sont susceptibles de converger vers des zones communes lors de la migration et de l'hivernage (Eraud et al., 2013).

Cette étude vise ainsi à comparer les stratégies spatiales et temporelles de ces deux ensembles de populations, en s'appuyant sur des données de télémétrie ARGOS. Deux dimensions majeures sont examinées : (i) la traversée du Sahara, qui constitue une contrainte écologique forte et potentiellement sélective, et (ii) l'organisation spatiale des individus en zone sahélienne durant l'hivernage.

L'intérêt de cette approche réside dans sa capacité à identifier le degré de connectivité migratoire entre populations. Si des individus issus d'origines différentes utilisent les mêmes couloirs migratoires ou les mêmes secteurs d'hivernage, cela implique une exposition commune aux pressions environnementales et anthropiques (dégradation des habitats sahéliens, pression de chasse, variabilité climatique). À l'inverse, des stratégies différenciées pourraient indiquer une structuration spatiale plus marquée des populations et des facteurs de risque distincts.

Dans une perspective de conservation, ces éléments sont déterminants. Ils permettent de mieux cibler les zones géographiques prioritaires pour la protection de l'espèce, d'identifier les périodes critiques du cycle annuel et d'évaluer dans quelle mesure les actions de gestion doivent être pensées à l'échelle locale, régionale ou intercontinentale.

Ainsi, en comparant des individus se reproduisant en France et au Maroc, ce travail vise à préciser le rôle du Sahara comme filtre écologique, à caractériser l'utilisation des habitats sahéliens, et à mieux comprendre les niveaux de partage des risques entre populations, dans un contexte de déclin généralisé de l'espèce.

II - LA TOURTERELLE DES BOIS : TAXONOMIE ET ÉCOLOGIE COMPARÉE

II.1 - Taxonomie et différenciation des sous-espèces de la Tourterelle des bois

II.1.1 - Position taxonomique et phylogénétique

La Tourterelle des bois *Streptopelia turtur* appartient à l'ordre des Columbiformes, à la famille des Columbidae et au genre *Streptopelia*. Les premières phylogénies moléculaires consacrées aux genres *Streptopelia* et *Columba* ont placé *S. turtur* dans un groupe comprenant notamment *S. orientalis*, *S. tranquebarica*, *S. bitorquata* et *S. hypopyrrha* (Johnson et al., 2001). Ces analyses ont également montré que le genre *Streptopelia*, dans son acception historique, ne constituait pas un ensemble monophylétique. Elles ont notamment contribué au transfert de *S. chinensis* et de *S. senegalensis* vers le genre *Spilopelia* (Johnson et al., 2001). L'analyse ultérieure du mitogénome complet de *S. orientalis* a conforté la distinction phylogénétique entre les lignées actuellement rattachées aux genres *Streptopelia* et *Spilopelia* (Huang et al., 2016).

La classification taxonomique conventionnelle reconnaît quatre sous-espèces de Tourterelle des bois : *Streptopelia turtur turtur* (Linnaeus, 1758), *S. t. arenicola* Hartert, 1894, *S. t. hoggara* Geyr von Schweppenburg, 1916 et *S. t. rufescens* (C. L. Brehm, 1845) (Baptista et al., 2025). Le nom *S. t. isabellina*, historiquement employé pour certains oiseaux de la vallée ou du delta du Nil, est aujourd'hui généralement traité comme un synonyme de *S. t. rufescens* et non comme une cinquième sous-espèce (Baptista et al., 2025).

Cette subdivision sous-spécifique repose historiquement sur la distribution géographique et sur des différences moyennes de coloration, de taille et de proportions corporelles. Ces caractères ont toutefois été définis à partir de séries de spécimens d'ampleur et de provenance variables. Leur valeur diagnostique peut en outre être affectée par le sexe, l'âge, l'usure du plumage, la saison, les conditions de préparation des spécimens et l'existence de variations géographiques continues.

Les études morphométriques et génétiques contemporaines conduisent ainsi à distinguer la reconnaissance nomenclaturale des sous-espèces de la démonstration de leur indépendance évolutive. En particulier, les formes *S. t. turtur* et *S. t. arenicola* ne correspondent ni à deux ensembles morphométriques nettement séparés ni à deux lignées mitochondriales ou nucléaires distinctes (Prakas et al., 2021 ; Domínguez et al., 2026 ; Russ et al., 2026). La classification en quatre sous-espèces demeure donc en usage, mais sa signification biologique doit être évaluée séparément pour chacune des formes reconnues.

II.1.2 - Caractérisation des sous-espèces reconnues

a - Streptopelia turtur turtur

La sous-espèce nominale occupe principalement l'Europe et s'étend vers l'Asie occidentale. Elle est également présente dans certaines îles de l'Atlantique oriental, bien que l'attribution sous-spécifique de plusieurs populations insulaires ait varié selon les classifications. Les populations continentales sont majoritairement migratrices à longue distance et rejoignent l'Afrique subsaharienne en empruntant les voies occidentale, centrale ou orientale de migration transsaharienne (Marx et al., 2016 ; Baptista et al., 2025).

La forme nominale constitue le référentiel morphologique auquel les autres sous-espèces ont historiquement été comparées. Son plumage adulte associe une tête et un cou gris bleuté, un manteau brun, des scapulaires et des couvertures alaires rousses ou châtaines marqués d'un centre noir, ainsi que les plages latérales noires et blanches caractéristiques du cou (Baptista et al., 2025). Ces caractères ne sont cependant pas parfaitement constants. Leur expression varie notamment selon l'âge, l'usure du plumage et, pour certains traits, le sexe.

Une variabilité biométrique importante existe également au sein de la forme nominale. En Grande-Bretagne, l'aile des mâles adultes était en moyenne environ 4 mm plus longue que celle des femelles, et celle des adultes plus longues que celle des juvéniles (Browne et Aebischer, 2003). Des différences de longueur alaire entre les sexes et les classes d'âge ont également été observées parmi les migrants capturés à Anticythère (Dimaki et Alivizatos, 2015). En Espagne, les adultes présentaient une aile plus longue mais relativement plus arrondie que celle des juvéniles (Cabodevilla et al., 2018). Ces variations liées au sexe, à l'âge et à la forme de l'aile doivent être contrôlées avant d'interpréter une différence biométrique entre régions comme un caractère sous-spécifique.

b - Streptopelia turtur arenicola

Streptopelia turtur arenicola est traditionnellement attribuée aux populations reproductrices du nord-ouest de l'Afrique et à un ensemble de populations s'étendant vers le Moyen-Orient, l'Iran, l'Asie centrale et l'ouest de la Chine. Certaines classifications lui rattachent également les populations des îles Baléares (Baptista et al., 2025). En Afrique du Nord-Ouest, les oiseaux sont principalement migrateurs, bien que des observations hivernales et des indices de résidence ou de migration partielle aient été rapportés localement (Hanane, 2017 ; Chedad et al., 2020). La description traditionnelle d'*arenicola* repose sur une coloration moyenne plus claire et plus sableuse que celle de la forme nominale, ainsi que sur des dimensions corporelles légèrement inférieures. Les parties supérieures sont généralement décrites comme moins intensément rousses et les tonalités grisâtres de la tête et du dessous comme plus claires (Baptista et al., 2025). Ces différences correspondent cependant à des tendances moyennes : aucun caractère de plumage ne permet d'attribuer avec certitude tous les individus à *arenicola* ou à *turtur*.

Cette absence de diagnostic individuel est confirmée par les données morphométriques récentes. À partir d'oiseaux provenant notamment d'Europe continentale, de Grande-Bretagne, d'Afrique du Nord, des Baléares et des Canaries, Domínguez et al. (2026) n'ont mis en évidence aucune séparation morphométrique robuste entre les individus attribués à *turtur* et ceux attribués à *arenicola*. Les longueurs de l'aile, de la queue et du tarse se chevauchent largement. Certaines populations, notamment marocaines, présentent des valeurs moyennes inférieures pour plusieurs caractères, mais ces variations s'inscrivent dans un continuum géographique et ne permettent pas une assignation sous-spécifique fiable (Domínguez et al., 2026).

Les données génétiques ne soutiennent pas davantage l'existence de deux lignées distinctes. Les populations marocaines attribuées à *arenicola* ne forment ni un groupe mitochondrial exclusif ni un ensemble nucléaire séparé des populations européennes de *turtur* (Prakas et al., 2021 ; Domínguez et al., 2026 ; Russ et al., 2026). La distinction entre les deux formes est donc maintenue dans les classifications usuelles, mais elle apparaît faiblement étayée par les caractères morphométriques et moléculaires actuellement disponibles.

c - *Streptopelia turtur hoggara*

La sous-espèce *S. t. hoggara* est associée aux populations à distribution restreinte du Sahara central, notamment dans les massifs du Hoggar et de l'Aïr (Baptista et al., 2025). Sa répartition discontinue autour des massifs montagneux, des vallées et des oasis a historiquement conduit à l'hypothèse d'un certain isolement géographique.

La reconnaissance d'*hoggara* repose principalement sur sa description taxonomique historique et sur sa distribution. Les données publiées ne permettent pas de caractériser quantitativement et avec robustesse les différences de taille ou de plumage qui la sépareraient d'*arenicola* ou de *rufescens*. De même, son degré de résidence ou de mobilité demeure insuffisamment documenté pour permettre une comparaison précise avec les autres populations nord-africaines.

Aucun des principaux travaux moléculaires récents consacrés à la structure des populations de Tourterelle des bois n'a inclus d'échantillons attribués avec certitude à *hoggara*. Sa différenciation mitochondriale ou nucléaire par rapport aux autres formes demeure donc inconnue. Son maintien comme sous-espèce repose actuellement davantage sur la nomenclature historique et sur son aire géographique particulière que sur une démonstration morphométrique ou génétique moderne.

d - *Streptopelia turtur rufescens*

Streptopelia turtur rufescens est traditionnellement signalée dans la vallée du Nil, principalement en Égypte et dans le nord du Soudan, ainsi que dans certaines oasis voisines (Baptista et al., 2025). Les populations concernées sont généralement considérées comme résidentes ou partiellement migratrices, mais l'ampleur de leurs déplacements saisonniers reste insuffisamment documentée.

Cette forme est traditionnellement décrite comme plus petite et plus vivement colorée que la sous-espèce nominale, avec des tonalités brun-roux ou orangées plus développées et une proportion plus faible de gris dans le plumage (Baptista et al., 2025). Ces caractères constituent le principal fondement de sa reconnaissance. Leur valeur diagnostique n'a cependant pas été évaluée à partir d'un échantillonnage morphométrique étendu, représentatif de l'ensemble de son aire et standardisé selon l'âge et le sexe.

Russ et al. (2026) ont intégré à leurs analyses un nombre limité d'oiseaux égyptiens attribués à *rufescens*. Les microsattellites nucléaires n'ont pas permis d'identifier un groupe égyptien distinct, tandis que les séquences mitochondriales du cytochrome *b* n'ont révélé qu'une faible différenciation par rapport aux populations européennes. Ces résultats montrent que la singularité

phénotypique attribuée à *rufescens* ne s'accompagne pas, dans les marqueurs étudiés, d'une divergence génétique profonde.

Ils ne suffisent cependant pas à invalider cette sous-espèce. La taille réduite de l'échantillon égyptien et la résolution limitée des microsatellites et du seul cytochrome *b* ne permettent pas d'exclure une différenciation portant sur d'autres régions du génome. Des différences de coloration peuvent notamment dépendre d'un nombre restreint de loci et persister malgré un flux génique affectant le reste du génome. Le statut de *rufescens* doit donc être considéré comme incertain plutôt que formellement invalidé.

II.1.3 - Portée et limites de la classification sous-spécifique

Les quatre sous-espèces ne bénéficient pas du même niveau de justification biologique. La distinction entre *turtur* et *arenicola* a été directement testée à partir de données morphométriques et moléculaires couvrant plusieurs populations du bassin méditerranéen occidental. Aucun caractère morphométrique diagnostique, aucune lignée mitochondriale exclusive et aucune structuration nucléaire concordante n'ont été mis en évidence (Prakas et al., 2021 ; Domínguez et al., 2026 ; Russ et al., 2026). Au regard des populations et des marqueurs étudiés, leur séparation taxonomique apparaît donc faiblement étayée.

Le statut de *rufescens* est moins résolu. Sa différenciation phénotypique traditionnelle paraît plus prononcée, mais elle n'est pas associée, dans l'échantillon actuellement disponible, à une divergence nucléaire ou mitochondriale importante (Russ et al., 2026). Celui d'*hoggara* reste pratiquement non testé, faute de données morphométriques standardisées et d'échantillons moléculaires provenant avec certitude de son aire saharienne.

Plusieurs mécanismes peuvent expliquer la faible concordance entre la géographie, le phénotype et les marqueurs génétiques : une divergence récente, le tri incomplet de lignées mitochondriales ancestrales, les expansions démographiques postglaciaires et un flux de gènes entre populations reproductrices. Une variation clinale de la taille ou de la coloration le long des gradients environnementaux pourrait également avoir été subdivisée artificiellement en catégories discrètes (Calderón et al., 2016 ; Prakas et al., 2021 ; Domínguez et al., 2026 ; Russ et al., 2026).

En définitive, les quatre sous-espèces doivent être présentées comme les unités de la classification conventionnelle actuellement utilisée, et non comme quatre lignées évolutives dont l'indépendance aurait été démontrée. Une révision taxonomique robuste nécessiterait un échantillonnage couvrant l'ensemble de l'aire de répartition, en particulier le Sahara central, la vallée du Nil, le Moyen-Orient et l'Asie centrale. Elle devrait associer des données de génomique

nucléaire, d'ADN mitochondrial, de morphométrie standardisée et de colorimétrie objective du plumage, tout en contrôlant les effets du sexe, de l'âge, de la saison et de l'usure des plumes.

II.2 - Écologie comparée des populations reproductrices d'Europe occidentale et du Maghreb

II.2.1 - Distribution et état des populations

En Europe occidentale, les principaux effectifs reproducteurs se trouvent dans la péninsule Ibérique et en France. L'espèce est devenue beaucoup plus localisée à la limite nord-occidentale de son aire, notamment dans le sud et l'est de l'Angleterre. Les suivis à long terme montrent un déclin généralisé, accompagné dans plusieurs régions d'une contraction de l'aire de reproduction (Fisher et al., 2018). Dans le Pays basque espagnol, la Tourterelle des bois n'a été détectée que dans deux des quatre secteurs biogéographiques prospectés et son abondance variait fortement selon le contexte paysager et la structure du couvert arboré (Sáenz de Buruaga et al., 2012). Ces résultats illustrent une distribution régionale fragmentée, mais ne peuvent être extrapolés à l'ensemble de la péninsule Ibérique ou de l'Europe occidentale.

Au Maghreb, la Tourterelle des bois est une migratrice reproductrice largement répartie au Maroc, en Algérie et en Tunisie. Elle fréquente des milieux agricoles, forestiers, riverains, oasiens et périurbains. Certaines plaines irriguées marocaines accueillent localement des effectifs importants. Dans le Tadla, les densités ont été estimées à $2,96 \text{ nids} \cdot \text{ha}^{-1}$ dans les orangeries et à $0,93 \text{ nid} \cdot \text{ha}^{-1}$ dans les oliveraies. Leur extrapolation aux superficies régionales de ces deux catégories de vergers a conduit à une estimation minimale de 58 969 couples reproducteurs, avec un intervalle de confiance à 95 % compris entre 48 550 et 69 353 couples (Hanane et Besnard, 2014). Dans deux parcelles intensivement prospectées, les densités brutes atteignaient $16 \text{ nids} \cdot \text{ha}^{-1}$ dans une oliveraie et $45 \text{ nids} \cdot \text{ha}^{-1}$ dans une orangerie (Hanane et Baâmal, 2011). Ces valeurs correspondent à des habitats localement très favorables et ne sont directement comparables ni aux estimations régionales tenant compte de la détectabilité ni aux densités de territoires ou d'oiseaux chanteurs généralement employées en Europe.

Le principal contraste entre les deux régions tient ainsi moins à l'abondance observée qu'à la qualité des données disponibles. Le déclin européen est solidement documenté, tandis qu'aucune série standardisée à long terme ne permet actuellement d'établir la tendance générale des populations maghrébines. Les fortes concentrations relevées dans certains vergers ne démontrent donc pas leur stabilité et pourraient aussi traduire une agrégation des oiseaux dans les habitats favorables encore disponibles (Hanane, 2017).

II.2.2 - Migration prénuptiale et arrivée sur les sites de reproduction

Les populations maghrébines rejoignent généralement leurs sites de reproduction plus tôt que les populations européennes. Au Maroc, les premières observations interviennent principalement en mars, avec des mentions dès la fin de février dans certaines régions méridionales, notamment dans la vallée du Souss (Hanane, 2017). À Béni Mellal, les premières arrivées annuelles ont été relevées en moyenne le 20 mars, contre le 28 mars à Midelt, entre 1 400 et 1 650 m d'altitude, soit un décalage moyen de huit jours (Mansouri et al., 2020).

En Europe, la migration prénuptiale se déroule principalement en avril et en mai (Marx et al., 2016). Les oiseaux empruntant la voie occidentale traversent l'Afrique du Nord et la péninsule Ibérique avant de rejoindre leurs sites reproducteurs européens. Les populations du sud-ouest de l'Europe s'installent généralement plus tôt que celles des régions centrales ou septentrionales. Les dix oiseaux suivis jusqu'à leur site reproducteur par Schumm et al. (2021), principalement en Italie, en Bulgarie et en Slovaquie, sont arrivés entre le 1er mai et le 20 juin, avec une date médiane au 18 mai. Cet échantillon ne représente cependant pas l'ensemble des populations d'Europe occidentale.

La comparaison doit être interprétée prudemment, car les dates marocaines correspondent aux premiers oiseaux détectés dans une région, tandis que la télémétrie permet d'identifier l'arrivée individuelle sur le site reproducteur final. En outre, le passage prénuptial des populations européennes à travers le Maghreb se superpose à l'installation des reproducteurs nord-africains. Les observations effectuées en mars et en avril peuvent donc concerner à la fois des reproducteurs locaux et des migrateurs poursuivant leur déplacement vers l'Europe.

II.2.3 - Chronologie de la reproduction

Dans les plaines marocaines, les premiers nids peuvent être construits au cours de la première moitié d'avril et les premières pontes apparaissent généralement à partir de la mi-avril (Hanane et Baâmal, 2011 ; Hanane, 2017). À Midelt, l'installation et la reproduction sont retardées par rapport à Béni Mellal et la période de présence est plus courte. Cette contraction du calendrier est interprétée comme une réponse aux conditions thermiques plus rigoureuses des habitats montagnards (Mansouri et al., 2020).

En Europe occidentale, les premières pontes sont généralement déposées à partir de mai et la reproduction peut se prolonger jusqu'en août ou septembre. Les couples peuvent entreprendre des pontes de remplacement après un échec et, lorsque les conditions le permettent, plusieurs tentatives réussies au cours d'une même saison (Browne et Aebischer, 2004 ; Fisher et al., 2018).

La durée de la période favorable influence donc directement le nombre de tentatives et la productivité annuelle.

En Grande-Bretagne, la comparaison entre les années 1960 et la période 1998–2000 a mis en évidence un raccourcissement de la saison reproductrice. À la fin des années 1990, les couples produisaient environ deux fois moins de pontes et de jeunes par saison, sans diminution équivalente de la réussite propre à chaque tentative (Browne et Aebischer, 2004). Le raccourcissement de la période favorable à la reproduction constitue ainsi l'un des mécanismes susceptibles d'avoir contribué au déclin britannique.

La reproduction généralement plus précoce au Maghreb est cohérente avec la progression saisonnière des conditions favorables du sud vers le nord et des basses vers les hautes altitudes. Elle ne démontre pas, à elle seule, une différence adaptative héritée entre *S. t. arenicola* et *S. t. turtur*.

II.2.4 - Migration postnuptiale

En Europe, les oiseaux suivis par télémétrie quittaient leur site de reproduction entre le 31 juillet et le 14 septembre, avec une date médiane au 30 août. La migration automnale, haltes comprises, durait de 7 à 64 jours, avec une médiane de 25,1 jours. Plusieurs individus effectuaient des haltes prolongées en Europe après leur départ du territoire reproducteur, notamment dans des régions méridionales, avant de poursuivre leur migration vers l'Afrique (Schumm et al., 2021). Les populations reproductrices d'Europe occidentale empruntent principalement une voie passant par la France, la péninsule Ibérique et le Maroc avant la traversée du Sahara (Marx et al., 2016).

Au Maroc, les derniers individus étaient observés en moyenne le 28 septembre à Midelt et le 11 ou le 12 octobre à Béni Mellal, soit un décalage d'environ treize jours (Mansouri et al., 2020). La période de présence est donc plus courte en altitude qu'en plaine, ce qui est compatible avec une fenêtre climatique et trophique plus contrainte dans les habitats montagnards.

Ces calendriers ne peuvent cependant pas être comparés directement. Les données européennes correspondent au départ individuel du site reproducteur, tandis que les dates marocaines reposent sur les derniers oiseaux encore observés dans une région. Ceux-ci peuvent comprendre des reproducteurs locaux, des individus effectuant une halte postnuptiale et des migrants européens en transit. La persistance d'observations jusqu'en octobre à Béni Mellal ne signifie donc pas que les reproducteurs locaux demeurent sur leur territoire jusqu'à cette date. En l'absence de suivi individuel des oiseaux nichant au Maroc, ces différentes composantes ne peuvent pas être dissociées.

II.2.5 - Hivernage

Les populations migratrices reproductrices d'Europe occidentale et du Maghreb hivernent majoritairement dans la bande sahélienne, approximativement entre 10° N et 20° N (Fisher et al., 2018 ; Hanane, 2017). Les connaissances disponibles ne permettent toutefois pas de délimiter des quartiers d'hivernage propres aux populations européennes et nord-africaines, ni de quantifier précisément leur degré de mélange dans les secteurs fréquentés.

Les suivis satellitaires apportent des informations plus précises sur l'hivernage des populations européennes. Requena et al. (2025) ont analysé les données de 14 individus équipés sur leurs sites de reproduction en France et au Royaume-Uni et ayant atteint leurs quartiers d'hivernage au Sénégal, en Gambie, au Mali et en Mauritanie. Leur distribution était associée à la proximité de l'eau et à une mosaïque de forêts ouvertes, de formations arbustives, de prairies naturelles et de terres cultivées, caractéristique des plaines inondables sahéliennes et soudano-sahéliennes comprenant des ripisylves à *Acacia nilotica*.

L'hivernage ne correspond pas nécessairement à une occupation stationnaire d'un même secteur. Dans cette étude, 11 des 14 individus ont déplacé leur domaine vital vers le sud entre le début et la fin de l'hiver, rejoignant des secteurs ayant reçu davantage de précipitations au cours de l'été précédent et présentant une gamme plus large de valeurs de NDVI. Ces résultats suggèrent une utilisation saisonnière de plusieurs secteurs d'hivernage, potentiellement liée à l'évolution des ressources disponibles (Requena et al., 2025).

Les populations européennes et maghrébines pourraient ainsi être exposées, au moins en partie, aux mêmes conditions environnementales pendant l'hiver : disponibilité des graines et des points d'eau, variabilité des précipitations, transformation des habitats et pression de prélèvement. Cette exposition commune reste toutefois à documenter par des comparaisons directes entre oiseaux d'origines reproductrices différentes. L'absence de différenciation génétique nette entre populations reproductrices ne constitue pas, à elle seule, une preuve du partage des mêmes quartiers d'hivernage.

La migration transsaharienne n'est par ailleurs pas systématique. Des observations hivernales régulières ont été rapportées dans le Sahara septentrional algérien, suggérant l'existence locale d'individus résidents, partiellement migrants ou hivernant au nord du Sahel (Chedad et al., 2020). Ces observations ne permettent toutefois ni d'identifier l'origine reproductrice des oiseaux concernés ni d'évaluer leur proportion dans les populations maghrébines. Elles indiquent néanmoins que les stratégies de déplacement sont plus diversifiées qu'une opposition stricte entre reproduction paléarctique et hivernage sahélien.

Si des différences de calendrier de reproduction sont documentées entre les populations européennes et maghrébines, les éventuelles différences de destination et de stratégie d'hivernage restent moins bien connues. Avant les suivis présentés dans ce mémoire, les données individuelles disponibles sur les reproducteurs maghrébins étaient insuffisantes pour caractériser leurs voies migratoires, leurs principaux quartiers d'hivernage et leur connectivité migratoire avec les populations européennes.

II.2.6 - Reproduction et causes d'échec

En Grande-Bretagne, la probabilité estimée qu'une tentative de reproduction atteigne l'éclosion était de 0,53, et celle qu'une tentative ayant atteint l'éclosion produise ensuite au moins un jeune à l'envol était de 0,65. La probabilité globale qu'une tentative produise au moins un jeune à l'envol était ainsi estimée à environ 0,35 (Browne et Aebischer, 2004). Dans l'ouest de la France, le suivi de 86 nids au cours de trois saisons a fourni une estimation du succès de nidification de 38,6 % et une productivité saisonnière moyenne de 2,40 jeunes par couple. Les modèles démographiques indiquaient cependant que cette productivité ne suffisait pas à compenser les niveaux estimés de mortalité et à assurer le maintien de la population étudiée (Eraud et Lormée, 2023).

Dans les orangeries et les oliveries du Tadla, le succès apparent, défini comme la proportion de nids ayant produit au moins un jeune à l'envol, atteignait environ 48 %, pour une production moyenne de 1,04 jeune à l'envol par tentative. Malgré une densité de nids beaucoup plus élevée dans l'orangerie, la taille des pontes et les performances reproductrices étaient comparables entre les deux types de vergers (Hanane et Baâmal, 2011). Une forte concentration de nids ne constitue donc pas nécessairement un indicateur de meilleure qualité reproductive.

Sur 893 nids suivis à Midelt et à Béni Mellal, le nombre de jeunes à l'envol représentait respectivement 57 et 60,15 % du nombre d'œufs pondus. Les causes d'échec différaient entre les deux secteurs : la prédation et les contraintes climatiques occupaient une place importante à Midelt, tandis que les abandons associés aux perturbations humaines prédominaient davantage à Béni Mellal (Mansouri et al., 2020).

Ces résultats ne permettent pas de conclure à une différence générale de performance reproductive entre l'Europe et le Maghreb. Les études utilisent des indicateurs différents : probabilité de survie des nids, proportion apparente de nids réussis, nombre de jeunes par tentative, rendement entre l'œuf et l'envol ou productivité annuelle par couple. Le succès apparent de 48 % observé au Tadla ne peut donc pas être directement comparé aux estimations européennes tenant compte de la durée d'exposition des nids. De même, une production de 1,04

jeune par tentative au Maroc ne peut être opposée aux 2,40 jeunes par couple et par saison mesurés en France.

Les données européennes montrent en outre que la productivité annuelle ne dépend pas uniquement de la réussite de chaque nid, mais également de la durée de la saison et du nombre de tentatives entreprises par les couples (Browne et Aebischer, 2004). Or, l'absence de suivi individuel au Maghreb empêche généralement de distinguer les pontes successives et les pontes de remplacement réalisées par un même couple. Les causes d'échec semblent également dépendre davantage du contexte local — structure de la végétation, communauté de prédateurs, climat et activités humaines — que de l'appartenance sous-spécifique présumée.

II.2.7 - Habitats de reproduction

En Europe occidentale comme au Maghreb, la Tourterelle des bois dépend de paysages associant des formations ligneuses pour la nidification, des surfaces ouvertes offrant des graines accessibles au sol et des points d'eau. Elle n'est donc strictement ni forestière ni agricole, mais caractéristique de mosaïques paysagères dans lesquelles ces ressources complémentaires sont disponibles à une distance compatible avec les déplacements des reproducteurs (Fisher et al., 2018 ; Carboneras et al., 2022).

En Europe, les supports de nidification comprennent notamment les haies, les fourrés, les ripisylves et les boisements ouverts. Dans le Pays basque espagnol, les abondances les plus élevées ont été observées dans les ripisylves et les formations ouvertes à chênes sempervirents. L'espèce devenait rare lorsque le couvert arboré dépassait approximativement 40 % du paysage étudié (Sáenz de Buruaga et al., 2012). Cette valeur décrit une relation régionale et ne constitue pas un seuil écologique généralisable. Une couverture arborée importante peut devenir défavorable lorsqu'elle réduit la superficie ou l'accessibilité des milieux ouverts utilisés pour l'alimentation. Inversement, les paysages agricoles peuvent rester favorables lorsqu'ils conservent des haies, des bosquets, des marges peu intensivement gérées et des points d'eau (Fisher et al., 2018 ; Carboneras et al., 2022).

Au Maghreb, les habitats étudiés comprennent principalement des vergers, des palmeraies et des formations forestières. Dans le Tadla, les orangeries accueilleraient davantage de nids que les oliveraies, sans présenter un succès reproducteur supérieur. Leur valeur dépend vraisemblablement de la structure du feuillage et de la disponibilité des supports, mais également du calendrier des travaux agricoles, du maintien de la végétation herbacée, de l'accès à l'eau et du niveau de perturbation pendant la reproduction (Hanane et Baâmal, 2011 ; Hanane, 2017).

Dans les palmeraies de Biskra, les nids étaient préférentiellement associés aux palmiers les plus développés et à un couvert local relativement dense, soulignant l'importance de la structure verticale et de l'environnement immédiat des arbres (Saâd et al., 2022). Dans la forêt marocaine de la Maâmora, la présence de l'espèce variait avec le couvert arboré et la configuration des infrastructures humaines, selon des relations différentes entre territoires cynégétiques aménagés et non aménagés. Ces associations pourraient notamment refléter la distribution des ressources alimentaires et des points d'eau, mais ces mécanismes n'ont pas été directement testés (Hanane et al., 2023).

La différence entre les deux régions porte donc surtout sur la nature des systèmes arborés étudiés, et non sur leurs fonctions écologiques fondamentales. Dans les deux cas, la présence d'arbres ne suffit pas à définir un habitat favorable : la qualité dépend de la structure du couvert, de la disponibilité des surfaces d'alimentation et de l'eau, ainsi que de l'intensité des perturbations. Les résultats actuels ne démontrent pas une plasticité écologique intrinsèquement supérieure des populations maghrébines.

II.2.8 - Régime alimentaire et habitats d'alimentation

Le régime essentiellement granivore de la Tourterelle des bois est bien documenté en Europe. L'espèce consomme principalement au sol des graines de plantes sauvages et cultivées, dont la contribution varie selon la région, la saison et la disponibilité locale des ressources (Fisher et al., 2018 ; Carboneras et al., 2022).

Dans la Sierra de Andújar, dans le sud de l'Espagne, l'analyse des contenus digestifs de 222 oiseaux a permis d'identifier les graines de 30 espèces de plantes sauvages. Malgré cette diversité, un nombre limité d'espèces représentait l'essentiel du nombre et du volume des graines consommées, notamment *Echium plantagineum* et *Amaranthus deflexus*. Les adultes contenaient davantage de graines sauvages que les juvéniles, tandis que les éléments animaux, principalement des gastéropodes, n'étaient présents que chez 2,7 % des individus (Gutiérrez-Galán et Alonso, 2016). Cette consommation marginale ne remet pas en cause la caractérisation générale du régime comme essentiellement granivore.

Au Maghreb, les données quantitatives sont beaucoup plus rares. La fréquentation importante des agroécosystèmes suggère l'utilisation de graines cultivées et de plantes spontanées, mais la composition du régime a rarement été étudiée directement. L'occupation des vergers ne signifie pas que les fruits ou les arbres cultivés constituent les ressources alimentaires principales. Ces derniers servent avant tout de supports de nidification, tandis que l'alimentation peut être assurée par la végétation des inter-rangs, les friches, les parcelles voisines et les cultures céréalières.

Aucune différence intrinsèque de régime alimentaire ne peut donc être établie entre les populations européennes et maghrébines. Le contraste le plus manifeste concerne le niveau de connaissance : le régime est documenté par des analyses directes en Europe, mais principalement inféré à partir des habitats fréquentés au Maghreb.

II.2.9 - Pressions et facteurs limitants

En Europe occidentale, le déclin est principalement associé à la transformation des paysages agricoles et à la dégradation des habitats de reproduction. L'usage d'herbicides, la simplification des rotations, la disparition des chaumes et la gestion intensive des marges ont réduit l'abondance et l'accessibilité des graines de plantes adventices. La suppression des haies, des fourrés et des lisières a parallèlement diminué la disponibilité des supports de nidification (Fisher et al., 2018 ; Carboneras et al., 2022).

La Tourterelle des bois dépend toutefois d'un ensemble spatialement organisé de ressources. La dissociation entre les formations ligneuses, les surfaces d'alimentation et les points d'eau peut réduire la qualité fonctionnelle d'un territoire même lorsque chacun de ces éléments subsiste séparément (Carboneras et al., 2022). Les prélèvements cynégétiques peuvent également affecter la survie et les populations reproductrices situées en amont des voies migratoires. Leur incidence dépend du niveau de prélèvement, de la productivité et de l'état démographique des populations. Les conditions rencontrées pendant la migration et dans les quartiers d'hivernage sahéliens participent enfin à une dynamique qui doit être appréhendée sur l'ensemble du cycle annuel (Fisher et al., 2018).

Au Maghreb, les facteurs limitants sont beaucoup moins bien quantifiés. Dans les vergers marocains, la récolte des fruits, la taille des arbres et la fréquentation humaine peuvent provoquer l'abandon ou la destruction des nids (Hanane et Baâmal, 2011 ; Hanane, 2017). Dans les régions d'altitude, les basses températures et les épisodes météorologiques défavorables peuvent contribuer à la non-éclosion des œufs et à la mortalité des poussins, mais les observations de Midelt ne permettent pas de généraliser cette contrainte à l'ensemble du Maghreb (Mansouri et al., 2020).

La chasse, la transformation des paysages agricoles, les changements dans les pratiques d'irrigation, l'emploi de pesticides et les modifications climatiques constituent également des pressions potentielles en Afrique du Nord. Leur contribution respective à la dynamique des populations demeure néanmoins impossible à hiérarchiser en l'absence de séries démographiques standardisées et d'études ciblées (Hanane, 2017).

Le contraste entre l'Europe occidentale et le Maghreb concerne donc avant tout le degré de démonstration. En Europe, plusieurs mécanismes du déclin sont étayés par des séries démographiques et des études écologiques convergentes. Au Maghreb, les pressions sont principalement décrites à l'échelle locale et leur influence sur la dynamique globale des populations reste inconnue.

II.2.10 - Limites de la comparaison

La comparaison entre les populations européennes et maghrébines est limitée par l'hétérogénéité des méthodes et des échelles d'étude. Les travaux européens reposent sur des suivis standardisés d'oiseaux chanteurs, des atlas, des programmes de baguage, des reprises cynégétiques, la télémétrie et des études locales de la reproduction ou de l'alimentation. Les données maghrébines proviennent surtout de quelques agroécosystèmes et formations forestières du Maroc et du nord de l'Algérie.

Les unités de mesure diffèrent également : territoires, couples, oiseaux chanteurs, nids détectés ou probabilités de présence. Le succès reproducteur peut désigner une proportion brute de nids ayant produit un jeune, une probabilité de survie tenant compte de la durée d'exposition, un nombre de jeunes par tentative, un rendement par œuf ou une productivité annuelle par couple. Les différences numériques ne doivent donc pas être interprétées comme des différences biologiques sans tenir compte de la définition des indicateurs.

Enfin, la latitude, l'altitude, le climat, la structure des paysages et les pratiques agricoles diffèrent simultanément entre les sites étudiés. Leur influence ne peut être séparée d'un éventuel effet de l'appartenance sous-spécifique. Une comparaison démonstrative nécessiterait l'application simultanée de protocoles identiques dans plusieurs populations situées des deux côtés de la Méditerranée.

II.2.11 - Conclusion

Les populations reproductrices de Tourterelle des bois d'Europe occidentale et du Maghreb présentent des différences géographiques de phénologie, d'utilisation des habitats et de causes locales d'échec reproducteur. L'installation et la reproduction commencent généralement plus tôt au Maghreb, conformément au gradient climatique latitudinal, tandis que l'altitude retarde l'arrivée et raccourcit la période de présence dans les montagnes marocaines. En revanche, aucune différence claire n'est démontrée dans les stratégies d'hivernage : les populations migratrices des

deux régions semblent rejoindre majoritairement le Sahel, tandis qu'une fraction non quantifiée des oiseaux peut hiverner ou demeurer localement en Afrique du Nord.

Les besoins écologiques fondamentaux apparaissent comparables dans les deux régions. La reproduction nécessite l'association de supports ligneux, de surfaces ouvertes riches en graines et de points d'eau accessibles. Les supports utilisés diffèrent selon les paysages — haies, ripisylves et boisements ouverts en Europe ; vergers, palmeraies et formations forestières au Maghreb — mais remplissent des fonctions analogues. Cette diversité témoigne de la capacité de l'espèce à exploiter différents systèmes arborés naturels ou cultivés, sans démontrer une plasticité intrinsèquement supérieure des populations maghrébines.

Les performances reproductrices publiées ne permettent pas davantage de mettre en évidence une différence générale entre les deux ensembles. Les indicateurs employés ne sont pas directement comparables et l'absence de suivis individuels au Maghreb empêche d'estimer le nombre annuel de tentatives et la productivité par couple. Les causes d'échec semblent surtout dépendre du contexte local : la prédation est importante dans plusieurs populations européennes et dans les sites marocains d'altitude, tandis que les perturbations agricoles peuvent provoquer de nombreux abandons dans certains vergers de plaine.

La différence la mieux établie concerne finalement l'état des connaissances. En Europe occidentale, le déclin démographique est clairement documenté et associé à la dégradation des mosaïques d'habitats, à la diminution des ressources alimentaires et, selon les régions, aux prélèvements cynégétiques. Au Maghreb, les informations demeurent trop fragmentaires pour déterminer une tendance démographique générale ou hiérarchiser les facteurs limitants. Les fortes densités observées dans certains vergers ne permettent pas de conclure à la stabilité des populations nord-africaines (Hanane, 2017 ; Fisher et al., 2018).

Les données actuelles ne démontrent donc pas une divergence écologique intrinsèque entre les populations traditionnellement rattachées à *S. t. turtur* et à *S. t. arenicola*. Elles indiquent plutôt des réponses aux conditions climatiques, à la structure des paysages et aux pratiques de gestion. Cette interprétation concorde avec les études morphométriques et génétiques récentes, qui ne mettent pas en évidence de séparation nette entre les deux sous-espèces présumées (Prakas et al., 2021 ; Domínguez et al., 2026 ; Russ et al., 2026).

La priorité consiste désormais à appliquer des protocoles démographiques, reproducteurs et trophiques communs des deux côtés de la Méditerranée. Ces suivis devraient employer les mêmes définitions du succès reproducteur, tenir compte de la détectabilité et mesurer conjointement la structure des habitats, la disponibilité des graines, l'accès à l'eau et les pratiques

agricoles. Au Maghreb, l'étude directe du régime alimentaire, des pesticides et des agents pathogènes, notamment *Trichomonas gallinae*, reste particulièrement nécessaire.

Enfin, l'extension du baguage et de la télémétrie aux reproducteurs maghrébins permettrait d'identifier leurs voies migratoires et leurs quartiers d'hivernage, d'évaluer leur degré de mélange avec les populations européennes dans le Sahel et de quantifier les stratégies de résidence ou d'hivernage en Afrique du Nord. L'association de ces données de déplacement aux paramètres démographiques et génomiques est indispensable pour comprendre la connectivité entre populations et attribuer les pressions rencontrées au cours du cycle annuel. La conservation de la Tourterelle des bois doit ainsi être coordonnée à l'échelle de l'Europe, du Maghreb et des quartiers d'hivernage sahéliens (Hanane, 2017 ; Fisher et al., 2018).

III - MATERIEL ET METHODES

III.1 - Capture et déploiement des balises

Cette étude repose sur des données de suivi satellitaire collectées entre mai 2015 et janvier 2026 chez des Tourterelles des bois équipées de balises ARGOS sur leurs sites de reproduction en France et au Maroc. Les oiseaux ont été capturés à l'aide de cages-pièges avec apâts (mélange blé / maïs concassé).

En France, 17 individus ont été équipés de balises solaires ARGOS Solar PTT-100 (Microwave Telemetry ; masse : 5 g) au cours des saisons de reproduction entre 2015 et 2019. Au Maroc, huit individus ont été équipés en mai 2025 : quatre ont reçu une balise Solar 2G (Microwave Telemetry ; masse : 1,8–2,0 g) et quatre une balise S1H 130A (Lotek ; masse : 2,0–2,2 g), également à alimentation solaire.

Les balises ont été fixées à l'aide d'un harnais de type *wing-loop*, réalisé en téflon tubulaire de 2 mm de largeur. Les deux brins du harnais, passant de part et d'autre de la tête sous la fosse claviculaire, ont été réunis dans une gaine thermorétractable de 15 mm de longueur afin d'ajuster la tension du dispositif.

Au Maroc, les quatre balises Microwave Telemetry ont cessé d'émettre avant le début de la migration et n'ont donc fourni aucune donnée exploitable pour les analyses migratoires et hivernales. Un autre individu, équipé d'une balise Lotek, a probablement été prédaté avant son départ en migration. Les analyses présentées dans ce mémoire portent ainsi sur trois individus marocains.

III.2 - Nettoyage des données

Les analyses reposent sur des données de télémétrie ARGOS acquises sur des individus de Tourterelle des bois (*Streptopelia turtur*) suivis en France et au Maroc. Les données comprennent pour chaque localisation un identifiant individuel, une date et heure d'acquisition, ainsi que des coordonnées géographiques (WGS84).

Script : Les données ARGOS brutes ont été nettoyées afin de produire un jeu de données homogène et directement exploitable pour l'analyse des déplacements. Le traitement a été réalisé sous R à l'aide d'un script dédié.

Seules les localisations associées aux classes ARGOS les plus fiables (LC 3, LC 2 et LC 1) ont été conservées, garantissant un niveau de précision compatible avec une analyse spatiale à l'échelle régionale. Ces classes correspondent à des niveaux d'incertitude croissants, avec une précision nominale généralement inférieure à 250 m pour LC 3, comprise entre 250 et 500 m pour LC 2, et pouvant atteindre 1 500 m pour LC 1. Il convient de noter que ces valeurs restent indicatives et peuvent varier en fonction des conditions de réception et de géométrie des satellites.

Les données ont ensuite été structurées par individu et par satellite, puis ordonnées chronologiquement.

Afin de limiter la redondance des localisations issues d'un même passage satellite, les données ont été regroupées en séquences temporelles continues. Un passage satellite a été défini comme une succession de localisations séparées par un intervalle maximal de 15 minutes. Au-delà de ce seuil, un nouveau passage était considéré.

Pour chaque passage, une seule localisation a été retenue selon une hiérarchie de critères visant à maximiser la qualité de l'information : priorité à la meilleure classe ARGOS, puis à la plus faible incertitude spatiale (approchée par la taille de l'ellipse d'erreur lorsque disponible), et enfin à la localisation la plus récente.

Cette procédure a permis d'obtenir un jeu de données final standardisé, contenant une localisation unique par passage satellite, réduisant fortement la redondance des données tout en conservant l'information la plus fiable. Ce jeu de données nettoyé a ensuite été utilisé comme base pour l'ensemble des analyses spatiales et comportementales.

Voir script en annexe 1.

III.3 - Identification des individus migrants

Les données nettoyées ont été utilisées pour identifier des structures spatio-temporelles correspondant à des zones d'utilisation prolongée au cours de la migration et de l'hivernage. Cette étape a été réalisée sous R à l'aide d'un script dédié de clustering spatio-temporel.

Seuls les individus ayant effectivement franchi le Sahara ont été retenus pour cette analyse. Le franchissement a été défini de manière opérationnelle par le passage successif vers le sud des latitudes 30°N puis 20°N, permettant d'isoler les trajectoires comportant une migration transsaharienne complète.

Les localisations ont ensuite été segmentées temporellement afin de tenir compte des discontinuités de suivi. Une rupture temporelle a été définie lorsque l'intervalle entre deux localisations successives excédait 72 heures, ce qui permet de distinguer des phases de présence continue de périodes de déplacement ou d'absence de données.

Au sein de chaque segment temporel, un clustering spatial séquentiel a été appliqué. Les localisations ont été agrégées en groupes successifs dès lors qu'elles restaient situées dans un rayon maximal de 40 km autour d'un centroïde dynamique recalculé à chaque ajout de point. Ce procédé permet de détecter des noyaux d'utilisation spatiale cohérents au cours du temps.

Les clusters ainsi obtenus ont ensuite été fusionnés à une échelle régionale lorsque leurs centroïdes successifs étaient séparés par une distance inférieure à 20 km, afin de regrouper des structures proches spatialement et appartenant à un même secteur fonctionnel.

Chaque cluster final a été caractérisé par ses propriétés spatio-temporelles : nombre de localisations, dates de début et de fin, durée d'occupation, coordonnées du centroïde et rayon maximal d'utilisation. Un filtrage final a été appliqué afin de ne conserver que les structures comportant au moins 2 localisations et une durée minimale de 1 jour, garantissant leur cohérence écologique.

Les clusters ainsi définis, désignés comme des « spots », correspondent à des zones d'utilisation significative au cours de la migration ou de l'hivernage. Ils constituent la base des analyses ultérieures visant à caractériser l'organisation spatiale des trajectoires et les dynamiques de déplacement à l'échelle régionale.

Voir script en annexe 2.

III.4 - Analyse de la traversée saharienne

La traversée saharienne a été analysée à partir du jeu de données ARGOS nettoyé. L'objectif était de caractériser la chronologie du franchissement du Sahara vers le sud et de comparer les trajectoires selon l'origine des individus.

Dans cette analyse, la traversée saharienne a été définie de manière opérationnelle comme le passage successif vers le sud de deux limites latitudinales : 30°N, correspondant à l'entrée dans l'ensemble saharien, puis 20°N, correspondant à l'entrée dans la zone sahélienne. Pour chaque individu, les localisations ont été ordonnées chronologiquement, puis les franchissements de ces deux seuils ont été détectés lorsqu'une localisation passait d'une latitude supérieure au seuil considéré à une latitude inférieure ou égale à ce seuil.

Pour chaque individu ayant franchi successivement les deux limites, trois variables principales ont été calculées : la date de passage de 30°N, la date de passage de 20°N et la durée de traversée saharienne, définie comme le nombre de jours écoulés entre ces deux franchissements. Les dates de passage ont également été converties en jours juliens afin de permettre les comparaisons entre individus et entre origines.

Les résultats ont ensuite été résumés par origine géographique des individus, en calculant les dates moyennes de franchissement de 30°N et 20°N, ainsi que la durée moyenne de traversée entre ces deux latitudes. Des tests de Wilcoxon ont été utilisés de manière exploratoire pour comparer les individus d'origine française et marocaine concernant la phénologie d'entrée dans le Sahara, l'arrivée dans la zone sahélienne et la durée de traversée.

Voir script en annexe 3.

III.5 - Conditions d'hivernage

Les secteurs d'utilisation hivernale ont été identifiés à partir du jeu de données ARGOS nettoyé, en ne conservant que les périodes de présence au sud de 20°N, utilisées ici pour définir l'entrée dans la zone sahélienne. Pour chaque individu, une période d'hivernage a été définie entre le premier franchissement de 20°N vers le sud et le franchissement ultérieur de cette même latitude vers le nord.

Afin de ne conserver que les hivernages suffisamment documentés, seules les périodes comportant au moins 90 jours de présence au Sahel et au moins 50 localisations ont été retenues. Les localisations associées à ces hivernages valides ont ensuite été utilisées pour identifier les secteurs d'utilisation hivernale.

Les données ont d'abord été segmentées temporellement. Une rupture a été définie lorsque l'intervalle entre deux localisations successives dépassait 14 jours, afin de séparer les périodes de présence continue de celles potentiellement interrompues par une absence prolongée de données ou un changement majeur de secteur.

Au sein de chaque segment temporel, les localisations ont été regroupées par clustering spatial séquentiel. Un noyau local était défini tant que les localisations successives restaient dans un rayon maximal de 50 km autour d'un centroïde dynamique. Lorsque cette distance était dépassée, un nouveau noyau était créé.

Les noyaux locaux successifs ont ensuite été fusionnés à une échelle régionale lorsque leurs centroïdes étaient distants de moins de 100 km, afin de regrouper des zones proches appartenant à un même secteur fonctionnel d'hivernage. Les clusters finaux n'ont été conservés que s'ils comprenaient au moins 5 localisations et une durée minimale d'occupation de 7 jours.

Pour chaque cluster hivernal valide, les dates de début et de fin d'occupation, le nombre de localisations, la durée d'occupation, le centroïde géographique et le rayon maximal d'utilisation ont été calculés. Des centroïdes et polygones convexes ont également été produits afin de permettre la représentation spatiale des secteurs d'hivernage.

Voir script en annexe 4.

Afin de tester si le départ des secteurs d'hivernage est associé à une dégradation des conditions de végétation locale, celles-ci sont caractérisées à l'aide de l'indice de végétation par différence normalisée (NDVI). L'analyse porte sur l'évolution du NDVI pendant l'occupation des secteurs ensuite quittés, ainsi que sur les différences de NDVI entre secteurs de départ et d'arrivée, afin de déterminer si les oiseaux rejoignent des zones plus vertes. Calculé à partir des réflectances dans le rouge et le proche infrarouge, le NDVI présente généralement des valeurs plus élevées lorsque la végétation est plus verte et plus active sur le plan photosynthétique.

Les données sont extraites à partir du produit MODIS MOD13Q1, à une résolution spatiale de 250 m. Ce produit fournit des composites sur des périodes de 16 jours : chaque pixel correspond à une observation sélectionnée durant cette période, et non à une moyenne temporelle. Les dates réelles d'observation peuvent donc différer entre les pixels d'un même composite.

Pour chaque cluster hivernal défini à partir des localisations ARGOS, l'extraction est centrée sur son centroïde. Une fenêtre fixe s'étendant de 25 km dans chaque direction, soit environ 50 × 50 km, permet de caractériser la végétation environnante. Cette fenêtre représente le paysage autour du cluster et non le contour exact de l'espace utilisé par l'oiseau. Sa taille reste constante, y compris en cas de difficulté de téléchargement.

L'extraction associe les valeurs de NDVI, les indicateurs de qualité des pixels et leurs dates réelles d'observation. L'analyse principale retient uniquement les pixels classés de bonne qualité. Les valeurs de remplissage, les valeurs hors de l'intervalle valide et les observations présentant des dates incohérentes sont exclues. Pour chaque composite, le NDVI moyen est calculé sur les pixels retenus dans la fenêtre. Une observation est considérée comme exploitable lorsqu'elle couvre au moins 50 % des pixels de la fenêtre et comprend au moins 30 pixels valides. Une analyse de sensibilité inclut également les pixels de qualité marginale.

Pour chaque transition d'un cluster A vers un cluster B, l'analyse principale sélectionne séparément le composite dont la date nominale est la plus proche de la dernière localisation ARGOS dans A et celui le plus proche de la première localisation dans B. Ces localisations délimitent la présence détectée dans chaque secteur, sans fournir nécessairement les instants exacts de départ et d'arrivée. La différence est calculée comme suit : $\Delta\text{NDVI} = \text{NDVI de la zone d'arrivée} - \text{NDVI de la zone de départ}$. Deux analyses de sensibilité utilisent, respectivement, un composite commun aux deux secteurs et des observations réalisées pendant les périodes d'occupation, au plus tard à la dernière localisation dans A et au plus tôt à la première localisation dans B.

L'analyse statistique repose sur plusieurs approches complémentaires. La distribution des différences de NDVI entre secteurs de départ et d'arrivée est décrite globalement et selon l'origine géographique des individus. Un test de Wilcoxon bilatéral compare les différences appariées à zéro. Un test binomial unilatéral évalue si la proportion de transitions vers une zone plus verte dépasse 50 %. Ces tests par transition sont considérés comme exploratoires, plusieurs transitions pouvant provenir d'un même oiseau. Une analyse complémentaire utilise une seule différence moyenne par individu, testée par Wilcoxon, afin de limiter l'influence des observations répétées. Aucun test comparatif France–Maroc n'est réalisé, les origines et les années de suivi étant confondues dans l'échantillon.

Le changement de NDVI pendant l'occupation des secteurs de départ est calculé entre les premiers et derniers composites exploitables, en conservant uniquement les pixels valides communs aux deux dates et dont les observations ont lieu pendant le séjour. Le calcul nécessite une couverture commune d'au moins 50 % de la fenêtre, au moins 30 pixels et un intervalle médian entre observations d'au moins huit jours. Une pente descriptive est également estimée lorsque la série comporte au moins trois dates exploitables couvrant au moins 16 jours, avec une couverture suffisante en pixels communs à toutes les dates. Ces mesures décrivent l'évolution pendant le séjour et ne correspondent pas nécessairement à une baisse immédiatement avant le départ.

Une analyse complémentaire porte sur les 48 jours précédant la dernière localisation dans chaque secteur. Cette fenêtre peut commencer avant l'arrivée de l'oiseau : elle décrit alors l'évolution récente de la végétation du secteur, plutôt que celle observée uniquement pendant son occupation. Les changements et les pentes temporelles sont présentés de manière descriptive, sans test permettant d'attribuer le déclenchement du départ à une baisse du NDVI.

Les analyses sont conduites sous R avec les packages MODISTools, dplyr, ggplot2 et geosphere. Les extractions, les critères de qualité et les motifs d'exclusion sont consignés dans des tables d'audit. Les transitions ou séries temporelles insuffisamment documentées sont conservées dans ces tables, mais exclues des calculs correspondants.

Voir script en annexe 5.

IV - RESULTATS

IV.1 - Traversée du Sahara : phénologie et durée

Au total, 12 individus ayant franchi le Sahara ont été retenus dans le jeu de données, dont 9 individus d'origine française et 3 individus d'origine marocaine. Dans le cadre de cette étude, le franchissement du Sahara est défini comme le passage de la limite latitudinale des 20°N, utilisée ici comme proxy opérationnel de la transition entre les zones tempérées nord-africaines et les écosystèmes sahéliens. Les suivis couvrent trois périodes distinctes pour les individus d'origine française (2015, 2016 et 2019) et une période plus récente pour les individus d'origine marocaine (2025).

Les durées de suivi varient de 115 jours à 727 jours, traduisant des différences importantes dans la couverture temporelle des trajectoires. Le nombre de localisations par individu est compris entre 271 et 1 593 points, assurant un niveau de résolution spatiale suffisant pour l'analyse des déplacements et de l'utilisation de l'espace sahélien.

Ces 12 trajectoires constituent l'ensemble des individus ayant effectivement franchi cette limite latitudinale et pour lesquels les données sont suffisamment complètes pour permettre l'identification des périodes d'hivernage et l'analyse fine de la structuration spatiale en clusters.

Tableau 1: Synthèse des individus ayant franchi le Sahara

ID	Origine	Date début	Date fin	Durée (jours)	Nb points
146588	France	14/05/2015	16/02/2016	278	743
146585	France	18/05/2015	13/05/2017	727	1275
146593	France	20/05/2015	24/04/2016	339	715
146583	France	28/05/2015	16/09/2016	477	1361
146581	France	28/05/2015	18/03/2017	660	1543
158001	France	30/05/2016	16/05/2017	351	1300
64899	France	04/06/2019	27/09/2019	115	271
64887	France	18/06/2019	09/10/2020	479	1161
64882	France	18/06/2019	09/10/2020	479	1223
283694	Maroc	20/05/2025	28/03/2026	312	1593
283695	Maroc	21/05/2025	26/04/2026	340	1174
283696	Maroc	21/05/2025	12/01/2026	236	1231

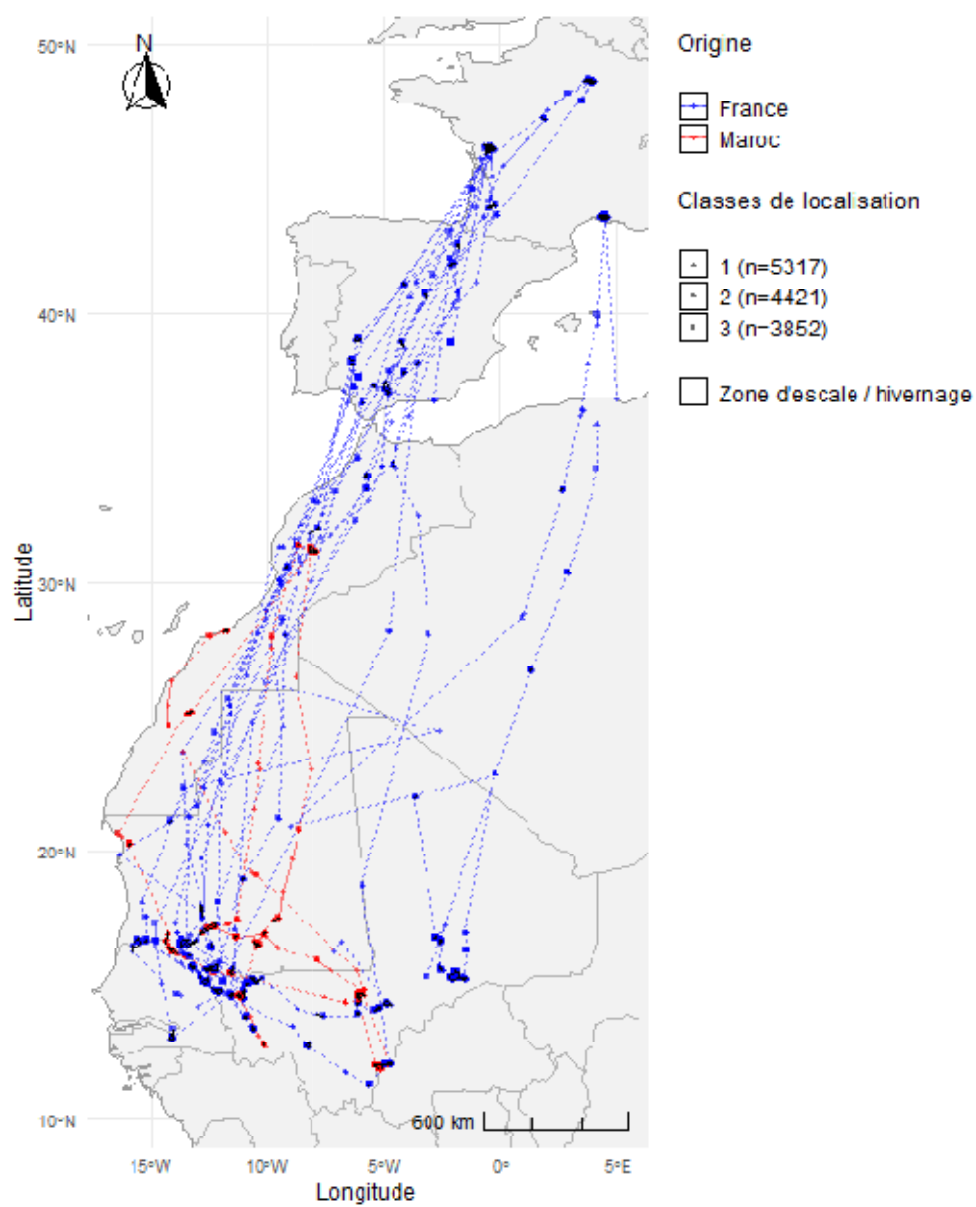


Figure 1 : Carte des oiseaux ayant franchi le Sahara.

Les individus marocains franchissent 30°N plus précocement que les individus français. En moyenne, ce passage intervient au jour julien 255 ($\pm 3,2$ jours), soit autour du 12 septembre, contre 264 ($\pm 5,7$ jours), soit autour du 21 septembre, pour les individus français. Cet écart d'environ 9 jours est statistiquement significatif ($p = 0,040$).

Au passage de 20°N, les individus marocains restent également plus précoces. Ils franchissent ce seuil en moyenne au jour julien 261, soit autour du 18 septembre, contre 268, soit autour du 25 septembre, pour les individus français. L'écart est donc du même ordre de grandeur (≈ 7 jours), mais n'est pas statistiquement significatif ($p = 0,263$). Ce résultat traduit une tendance similaire à celle observée à 30°N, mais qui ne peut pas être confirmée statistiquement, probablement en raison du faible effectif marocain.

La durée de traversée saharienne (intervalle entre 30°N et 20°N) ne diffère pas entre les deux groupes ($p = 0,864$), indiquant des vitesses de progression comparables au sein du Sahara.

Les résultats mettent en évidence un décalage temporel global de la migration, les individus marocains initiant et poursuivant la traversée saharienne plus précocement que les individus français. En revanche, la durée de traversée reste similaire, ce qui suggère que les différences observées relèvent du calendrier migratoire plutôt que de la vitesse de déplacement dans le Sahara.

IV.2 - Conditions d'hivernage

L'analyse a permis d'identifier 30 clusters hivernaux sahéliens valides, dont 22 pour les individus d'origine française et 8 pour les individus d'origine marocaine. Ces clusters correspondent à 11 suivis individu-hiver exploitables. En moyenne, les oiseaux ont utilisé $2,73 \pm 0,33$ clusters par hivernage (IC95 % : 1,99–3,47), avec une médiane de 3 clusters et une variation allant de 1 à 4 clusters selon les individus-hivers. Ce nombre moyen était très proche entre origines, avec $2,75 \pm 0,41$ clusters par hivernage pour les individus d'origine française et $2,67 \pm 0,67$ clusters pour les individus d'origine marocaine.

Les changements successifs de clusters ont représenté 19 transitions intra-hivernales. La distance moyenne entre deux clusters sahéliens successifs était de 336 ± 40 km (moyenne $\pm$ SE), indiquant des déplacements régionaux marqués au cours de l'hivernage. Ces distances étaient en moyenne plus faibles chez les individus d'origine française (299 ± 40 km, $n = 14$ transitions) que chez les individus d'origine marocaine (440 ± 97 km, $n = 5$ transitions), mais cette différence doit être interprétée avec prudence compte tenu du faible nombre de transitions disponibles, en particulier pour les oiseaux marocains.

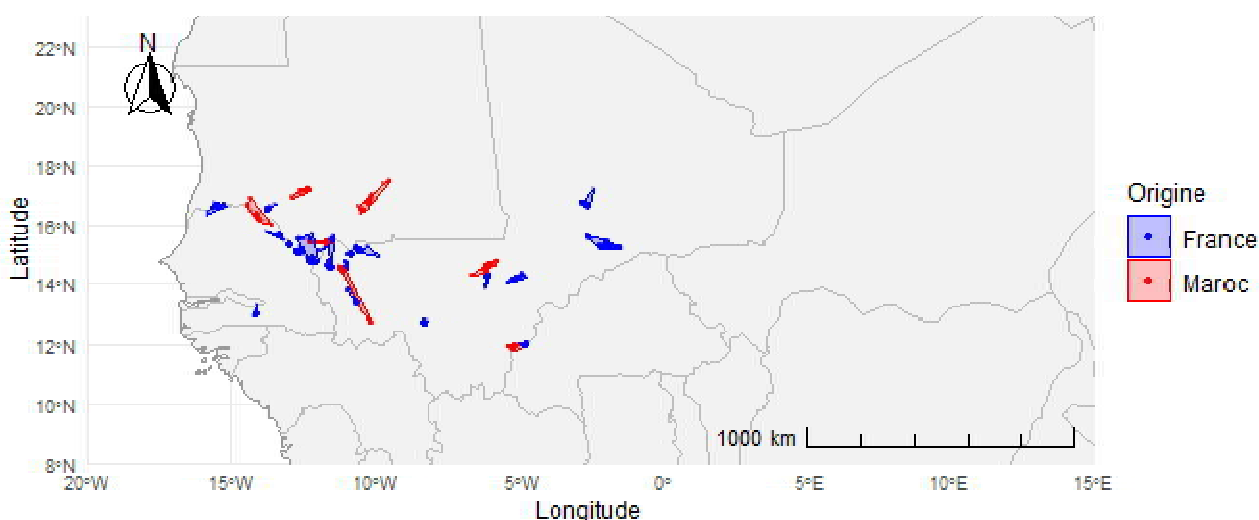


Figure 2 : Clusters définis en région sahélienne.

L'analyse des transitions intra-hivernales entre clusters successifs montre que 14 des 19 transitions documentées (73,7 %), concernant dix individus, s'effectuent vers un secteur présentant un NDVI supérieur à celui du cluster quitté. Le NDVI moyen associé au départ est de 0,256 dans les secteurs d'origine, contre 0,377 dans les secteurs d'arrivée, soit une différence moyenne de +0,121. Cette différence est significativement supérieure à zéro au test de Wilcoxon bilatéral ($p = 0,0018$). Le test binomial unilatéral indique également une proportion de transitions vers un secteur plus vert supérieure à 50 % ($p = 0,0318$). Ces tests par transition restent exploratoires, car ils ne tiennent pas compte des transitions répétées chez un même individu. Toutefois, la différence demeure significative lorsque l'analyse repose sur une seule différence moyenne par oiseau ($n = 10$; test de Wilcoxon bilatéral : $p = 0,008$). Ces résultats suggèrent que les déplacements intra-hivernaux conduisent généralement les individus vers des secteurs présentant une végétation plus verte.

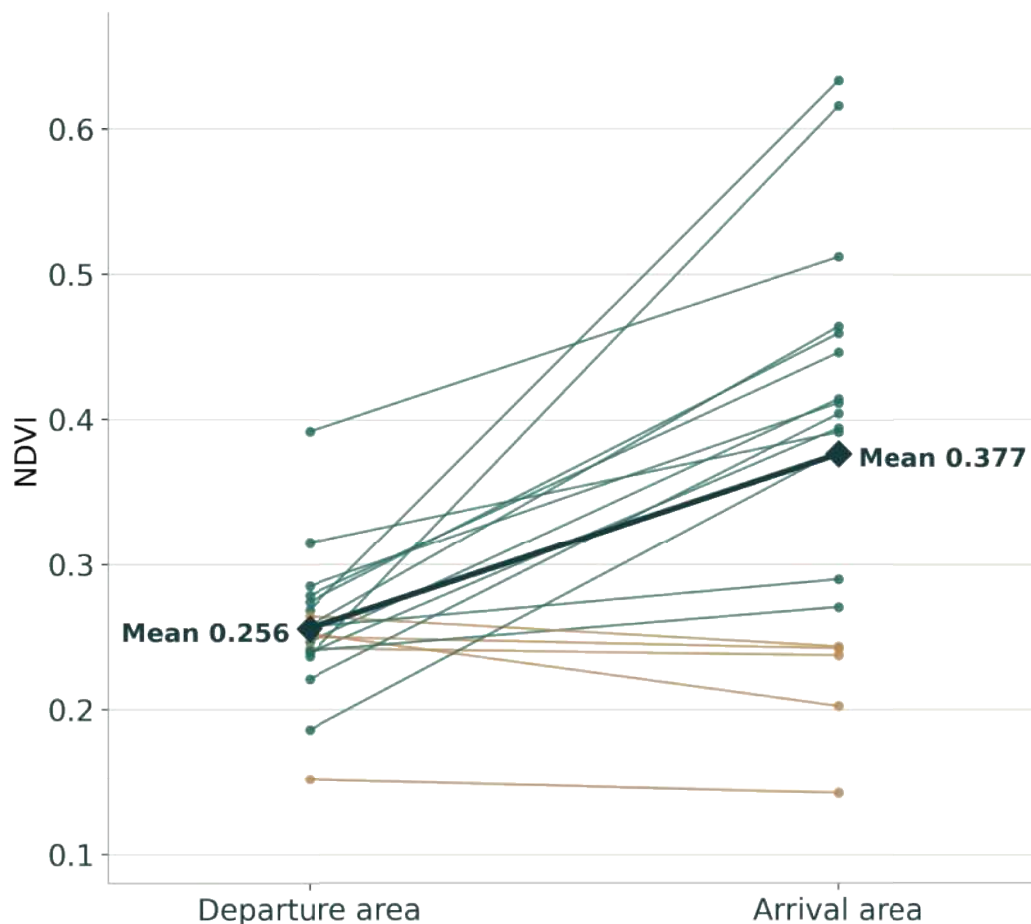


Figure 3 : Comparaison du NDVI entre secteurs de départ et d'arrivée pour 19 transitions intra-hivernales concernant dix Tourterelles des bois. Chaque ligne représente une transition. Les lignes vertes indiquent un NDVI supérieur à l'arrivée et les lignes ocre un NDVI inférieur. Les losanges représentent les moyennes au départ (0,256) et à l'arrivée (0,377).

V - DISCUSSION

V.1 - Traversée du Sahara : phénologie et durée

La traversée du Sahara (7 à 9 millions de km²) constitue le principal filtre écologique du système migratoire des oiseaux paléarctiques-africains, imposant des contraintes physiologiques majeures qui façonnent la phénologie et les stratégies de vol (Strandberg et al., 2009). Chez la Tourterelle des bois (*Streptopelia turtur*), les travaux antérieurs s'accordent sur une traversée dominée par le vol nocturne (87,5 % des localisations actives ; Lormée et al., 2016) à une vitesse moyenne de $52,6 \pm 0,5$ km/h, combinée à des haltes diurnes opportunistes dans le désert (Eraud et al., 2013). Cependant, la question de savoir si les variations spatio-temporelles observées lors du transit saharien découlent de performances de vol différenciées ou d'un décalage du calendrier d'initiation propre aux zones de reproduction demeurait jusqu'ici peu documentée par télémétrie comparative.

Nos suivis par satellite apportent un éclairage nouveau en dissociant deux composantes clés du transit saharien : la chronologie d'initiation et la vitesse de vol. Les individus d'origine marocaine ($n = 3$) franchissent la latitude 30°N significativement plus tôt que les individus d'origine française ($n = 9$), avec un passage moyen au jour julien $255 \pm 3,2$ (12 septembre) contre $264 \pm 5,7$ (21 septembre), soit un décalage de 9 jours ($p = 0,040$). Cette précocité des oiseaux marocains se maintient au franchissement des 20°N (jour julien 261 vs 268, soit 7 jours d'écart), bien que ce résultat ne soit pas statistiquement significatif ($p = 0,263$), vraisemblablement en raison du faible effectif. En revanche, la durée de traversée entre 30°N et 20°N ne diffère pas entre les deux origines ($p = 0,864$). Les vitesses de progression au sein du Sahara restent ainsi comparables, démontrant que la divergence observée relève du calendrier d'initiation migratoire et non d'une différence de performance de vol.

La précocité des départs observée chez les individus marocains s'inscrit dans le cadre de l'écologie reproductrice de *S. t. arenicola* en Afrique du Nord. Les travaux de Hanane & Baamal (2011) et Hanane (2017) ont établi que la disponibilité trophique dans les périmètres irrigués (ex. la plaine du Tadla) ainsi que le dessèchement estival rapide des ressources induisent un achèvement précoce du cycle de reproduction, incitant les oiseaux à initier leur migration plus tôt. Nos observations confirment cette plasticité phénologique, dont l'amplitude est comparable aux décalages altitudinaux enregistrés entre plaine et montagne au Maroc (Mansouri et al., 2020). Par ailleurs, l'homogénéité des durées de traversée ($p = 0,864$) confirme que le Sahara opère comme une barrière écologique majeure (Strandberg et al., 2009 ; Eraud et al., 2013), imposant des contraintes physiologiques et comportementales strictes qui dictent la vitesse de progression indépendamment de l'origine géographique. La convergence spatiale détectée au sud de 20°N rejoint également les schémas de connectivité hivernale sahélienne documentés dans la littérature

(Eraud et al., 2013). Ces conclusions doivent néanmoins être interprétées au regard du faible effectif marocain ($n = 3$) et du décalage temporel entre les suivis (2015–2019 en France vs 2025 au Maroc), deux facteurs susceptibles d'intégrer des variations interannuelles ou des effets récents du changement climatique

V.2 - Conditions d'hivernage

Le Sahel constitue une zone d'hivernage majeure pour de nombreuses espèces transsahariennes, mais également un environnement particulièrement variable sur les plans climatique et trophique (Jones, 1995 ; Zwarts et al., 2009). Les modifications des régimes pluviométriques, la dégradation des formations végétales ligneuses, l'intensification agricole ou encore les processus de désertification sont susceptibles d'altérer profondément la disponibilité spatio-temporelle des ressources exploitées par les oiseaux hivernants dans cette région (Wilson & Cresswell, 2006 ; Zwarts et al., 2009).

Dans ce contexte, le recours conjoint à la télémétrie et aux données issues de la télédétection, notamment le NDVI (*Normalized Difference Vegetation Index*), apparaît aujourd'hui comme central pour étudier les relations entre dynamique de la végétation, disponibilité trophique et mobilité des oiseaux migrateurs. Les données de télémétrie montrent notamment l'importance écologique des vallées fluviales, des zones humides sahéliennes et des mosaïques agro-pastorales, qui présentent généralement des valeurs de NDVI plus élevées et plus stables durant la saison sèche et concentrent une part importante des oiseaux hivernants (Walther & Rahbek, 2002 ; Vickery et al., 2014).

Nos résultats suggèrent que les conditions de végétation contribuent à structurer la mobilité intra-hivernale des Tourterelles des bois au Sahel. La majorité des transitions observées conduit vers des secteurs présentant un NDVI plus élevé que celui des secteurs quittés. Par ailleurs, dans huit des neuf secteurs pour lesquels la couverture temporelle permet une comparaison, le NDVI diminue pendant l'occupation. Ces observations sont compatibles avec l'hypothèse selon laquelle la diminution de la verdure de la végétation contribue aux déplacements vers des secteurs plus verts. Une évolution des ressources alimentaires pourrait participer à cette relation, mais le NDVI ne permet pas, à lui seul, de mesurer leur disponibilité pour les oiseaux. L'ensemble suggère une utilisation dynamique du paysage d'hivernage sahélien, dont les mécanismes restent à préciser sur un échantillon plus large.

Ces observations sont cohérentes avec plusieurs travaux récents suggérant que les mouvements intra-tropicaux observés au cours de la période d'hivernage constituent une stratégie adaptative permettant aux individus de suivre les variations spatiales de productivité végétale afin d'optimiser

l'accès aux ressources alimentaires avant la migration pré-nuptiale (Klaassen et al., 2018). Chez le Busard cendré (*Circus pygargus*), Schlaich et al. (2016) ont notamment montré que les individus hivernant au Sahel ajustent leurs déplacements en réponse à la dégradation progressive des conditions écologiques au cours de la saison sèche. Les oiseaux exploitent successivement différentes zones présentant des niveaux de productivité végétale plus favorables, traduisant une stratégie de suivi des ressources directement liée à l'évolution locale de la productivité végétale mesurée par NDVI et aux régimes pluviométriques sahéliens.

VI -CONCLUSION

Cette étude apporte de nouvelles connaissances sur la migration et l'hivernage de la Tourterelle des bois (*Streptopelia turtur*), à partir du suivi satellitaire ARGOS d'individus nichant en France et au Maroc. Le croisement des localisations avec des données environnementales issues de la télédétection a permis de caractériser leur calendrier migratoire, leur utilisation des quartiers d'hivernage et les variations de végétation associées à leurs déplacements. La structuration des données et des procédures d'analyse constitue également une base reproductible pour poursuivre ces travaux.

Dans l'échantillon étudié, les individus marocains franchissent la latitude de 30° N en moyenne neuf jours avant les individus français. En revanche, aucune différence de durée de franchissement entre 30° N et 20° N n'a été détectée. Le décalage observé à l'entrée du Sahara n'est donc pas accompagné d'une différence mesurable de durée de traversée. Les conditions environnementales et le calendrier de reproduction dans les régions d'origine pourraient contribuer à cette précocité, mais les données disponibles ne permettent pas d'en identifier les mécanismes. Ces comparaisons doivent également être interprétées au regard du faible nombre d'individus marocains et des différentes années de suivi.

Les trajectoires révèlent un chevauchement des secteurs sahéliens fréquentés par les oiseaux des deux origines. Ce résultat indique que des individus nichant en France et au Maroc peuvent utiliser des espaces communs pendant l'hivernage, sans permettre de quantifier la connectivité migratoire à l'échelle des populations. L'utilisation moyenne de près de trois secteurs successifs par hivernage souligne par ailleurs le caractère mobile de cette phase du cycle annuel.

Les analyses du NDVI apportent des éléments complémentaires pour comprendre cette mobilité. Dans l'analyse principale, 14 des 19 transitions intra-hivernales documentées, soit 73,7 %, conduisent vers un secteur présentant un NDVI supérieur à celui du secteur quitté. Le NDVI moyen passe de 0,256 dans les secteurs d'origine à 0,377 dans les secteurs de destination. Cette tendance se maintient dans les analyses de sensibilité portant sur la qualité des observations et leur comparabilité temporelle.

L'analyse des séries temporelles met également en évidence une baisse du NDVI avant plusieurs départs. La documentation de cette dynamique reste toutefois incomplète, notamment pour les séjours courts, et les fenêtres précédant les départs peuvent inclure des dates antérieures à l'arrivée des oiseaux. Ces résultats sont compatibles avec l'hypothèse d'un suivi des ressources au cours de la saison sèche : les individus quitteraient des secteurs dont les conditions se dégradent pour rejoindre des secteurs relativement plus favorables. Ils ne démontrent cependant ni que la baisse du NDVI déclenche les départs, ni que les destinations offrent directement davantage de nourriture. Le NDVI renseigne sur l'état de la végétation, mais ne mesure pas la disponibilité des graines consommées par les tourterelles.

Sur le plan de la conservation, ces observations soulignent l'intérêt de préserver un réseau de secteurs complémentaires permettant aux oiseaux de se déplacer au cours de l'hivernage. La protection des plaines inondables, des ripisylves, des points d'eau et des mosaïques de milieux ouverts et agricoles doit ainsi être envisagée à une échelle dépassant celle d'un site isolé. Le chevauchement des espaces fréquentés suggère également que certaines actions menées au Sahel pourraient bénéficier à des oiseaux issus de différentes régions de reproduction.

La poursuite des recherches devra augmenter le nombre d'individus suivis, particulièrement au Maghreb, et permettre des comparaisons entre origines sur des années communes. L'association des suivis satellitaires à des mesures de disponibilité alimentaire, d'accès à l'eau et de pratiques agricoles permettrait de préciser les mécanismes des déplacements intra-hivernaux. Des analyses comparant les secteurs quittés, les destinations et les autres secteurs disponibles aideraient également à distinguer les effets de la saisonnalité d'une sélection active des conditions environnementales.

Cette étude met ainsi en évidence un hivernage dynamique et des déplacements associés aux contrastes spatiaux et aux variations temporelles de la végétation sahéenne. Elle fournit une base pour approfondir les liens entre mobilité et ressources, et pour développer des actions de conservation coordonnées entre l'Europe, le Maghreb et l'Afrique subsaharienne.

VII - REFERENCES

- Baptista, L. F., Trail, P. W., Horblit, H. M., Boesman, P., & Sharpe, C. J. (2025). European Turtle-Dove (*Streptopelia turtur*), version 1.1. In S. M. Billerman (Ed.), *Birds of the World*. Cornell Lab of Ornithology.
- Browne, S. J., & Aebischer, N. J. (2003). Temporal variation in the biometrics of Turtle Doves *Streptopelia turtur* caught in Britain between 1956 and 2000. *Ringing & Migration*, 21(4), 203–208.
- Browne, S. J., & Aebischer, N. J. (2004). Temporal changes in the breeding ecology of European Turtle Doves *Streptopelia turtur* in Britain, and implications for conservation. *Ibis*, 146(1), 125–137.
- Cabodevilla, X., Moreno-Zarate, L., & Arroyo, B. (2018). Differences in wing morphology between juvenile and adult European Turtle Doves *Streptopelia turtur*: Implications for migration and predator escape. *Ibis*, 160(2), 458–463.
- Calderón, L., Campagna, L., Wilke, T., Lormée, H., Eraud, C., Dunn, J. C., Rocha, G., Zehtindjiev, P., Bakaloudis, D. E., Metzger, B., Cecere, J. G., Marx, M., Quillfeldt, P., & Illera, J. C. (2016). Genomic evidence of demographic fluctuations and lack of genetic structure across flyways in a long-distance migrant, the European Turtle Dove. *BMC Evolutionary Biology*, 16(1), Article 237.
- Carboneras, C., Moreno-Zarate, L., & Arroyo, B. (2022). The European Turtle Dove in the ecotone between woodland and farmland: Multi-scale habitat associations and implications for the design of management interventions. *Journal of Ornithology*, 163(2), 339–355.
- Chedad, A., Bendjoudi, D., & Guezoul, O. (2020). New data on the wintering and sedentary life of the European Turtle-dove *Streptopelia turtur* in the Algerian northern Sahara. *Current Trends in Natural Sciences*, 9(17), 65–73.
- Dimaki, M., & Alivizatos, H. (2015). Ringing studies of the Turtle Dove *Streptopelia turtur* during passage through Antikythera Island, southwestern Greece. *Journal of Natural History*, 49(7–8), 419–427.
- Domínguez, J. C., Bota, G., García, J. T., Arroyo, B., Hanane, S., Escandell, R., Sardà-Palomera, F., Ichen, A., & Moreno-Zarate, L. (2026). No subspecies, no borders: Morphometric and mitochondrial DNA evidence in western Mediterranean European Turtle-doves. *Bird Conservation International*, 36, e25.
- Eraud, C., & Lormée, H. (2023). Nesting success and productivity of a declining European Turtle Dove (*Streptopelia turtur*) population in western France. *Ibis*, 165(1).
- Eraud, C., Boutin, J. M., Rivière, M., & Lormée, H. (2013). Migration routes and staging areas of trans-Saharan Turtle Doves appraised from light-level geolocators. *PLoS ONE*, 8(3), e59396.
- Fisher, I., Ashpole, J., Scallan, D., Proud, T., & Carboneras, C. (2018). International Single Species Action Plan for the Conservation of the European Turtle-dove *Streptopelia turtur* (2018–2028). European Commission.
- Gutiérrez-Galán, A., & Alonso, C. (2016). European Turtle Dove *Streptopelia turtur* diet composition in southern Spain: The role of wild seeds in Mediterranean forest areas. *Bird Study*, 63(4), 490–499.
- Hanane, S. (2017). The European Turtle-Dove *Streptopelia turtur* in Northwest Africa: A review of current knowledge and priorities for future research. *Ardeola*, 64(2), 273–287.

- Hanane, S., & Baâmal, L. (2011). Are Moroccan fruit orchards suitable breeding habitats for Turtle Doves *Streptopelia turtur*? *Bird Study*, 58(1), 57–67.
- Hanane, S., & Besnard, A. (2014). Are nest-detection probability methods relevant for estimating Turtle Dove breeding populations? A case study in Moroccan agroecosystems. *European Journal of Wildlife Research*, 60(4), 673–680.
- Hanane, S., Bouamama, M., Bounous, A., & Mihoubi, H. (2023). Contrasting occurrence patterns in the European Turtle Dove (*Streptopelia turtur*) in managed and unmanaged hunting forest estates: Does human presence matter? *Biologia*, 78(8), 2107–2117.
- Huang, Z., Tu, F., & Murphy, R. W. (2016). Analysis of the complete mitogenome of Oriental Turtle Dove (*Streptopelia orientalis*) and implications for species divergence. *Biochemical Systematics and Ecology*, 65, 209–213.
- Johnson, K. P., de Kort, S., Dinwoodey, K., Mateman, A. C., ten Cate, C., Lessells, C. M., & Clayton, D. H. (2001). A molecular phylogeny of the dove genera *Streptopelia* and *Columba*. *The Auk*, 118(4), 874–887.
- Jones, P. (1995). Migration strategies of Palearctic passerines in Africa. *Israel Journal of Zoology*, 41(3), 393–406.
- Klaassen, R. H. G., Schlaich, A. E., Bouten, W., Both, C., & Koks, B. J. (2018). Intra-tropical movements as a beneficial strategy for Palearctic migratory birds. *Royal Society Open Science*, 5(2), Article 171675.
- Lormée, H., Boutin, J. M., & Eraud, C. (2016). Turtle Dove *Streptopelia turtur* migration routes and wintering areas revealed using satellite telemetry. *Bird Study*, 63(3), 425–429.
- Mansouri, I., Mounir, M., Squalli, W., Elhanafi, L., Dakki, M., & El Ghadraoui, L. (2020). Migratory dates, breeding phenology, and reproductive success of European Turtle Doves between lowlands and highest breeding habitats in North Africa. *International Journal of Zoology*, 2020, Article 8816577.
- Marx, M., Korner-Nievergelt, F., & Quillfeldt, P. (2016). Analysis of ring recoveries of European Turtle Doves *Streptopelia turtur*: Flyways, migration timing and origin areas of hunted birds. *Acta Ornithologica*, 51(1), 55–70.
- Prakas, P., Butkauskas, D., Švažas, S., Bea, A., Yanenko, V., Ragauskas, A., & Vaitkuvienė, D. (2021). The genetic diversity and structure of the European Turtle Dove *Streptopelia turtur*. *Animals*, 11(5), Article 1283.
- Requena, S., Lormée, H., Beresford, A. E., Buchanan, G. M., Eraud, C., Orsman, C. J., Rivière, M., Vickery, J. A., & Mallord, J. W. (2025). Riparian forests and open landscapes in the West African Sahel are key wintering habitats for the threatened European Turtle-dove (**Streptopelia turtur**). *Ibis*, 167(4), 895–911.
- Russ, H., Schumm, Y. R., Gerz, U., Metzger, B. J., Rocha, G., Mader, V., Sultanov, E., Mustafayev Mirzaga, I., Hering, J., Fünfstück, H.-J., Illera, J. C., Masello, J. F., & Quillfeldt, P. (2026). Towards more comprehensive population genetics of European Turtle Doves: Microsatellite markers and cytochrome b from across the distribution range. *European Journal of Wildlife Research*, 72(3), Article 47.
- Saâd, N., Hanane, S., Farhi, K., Matallah, N., & Khemis, M. D. E. H. (2022). Tree characteristics, microhabitat and edge effect in plantations govern European Turtle-dove *Streptopelia turtur* nest habitat selection at the edge of Sahara: Implications for conservation of a vulnerable species. *Bird Conservation International*, 32(3), 322–336.

- Sáenz de Buruaga, M., Onrubia, A., Fernández-García, J. M., Campos, M. Á., Canales, F., & Unamuno, J. M. (2012). Breeding habitat use and conservation status of the Turtle Dove *Streptopelia turtur* in northern Spain. *Ardeola*, 59(2), 291–300.
- Schlaich, A. E., Klaassen, R. H. G., Bouten, W., Both, C., & Koks, B. J. (2016). How individual Montagu's Harriers cope with Moreau's Paradox during the Sahelian winter. *Journal of Animal Ecology*, 85(6), 1491–1501.
- Schumm, Y. R., Metzger, B., Neuling, E., Austad, M., Galea, N., Barbara, N., & Quillfeldt, P. (2021). Year-round spatial distribution and migration phenology of a rapidly declining trans-Saharan migrant: Evidence of winter movements and breeding site fidelity in European Turtle Doves. *Behavioral Ecology and Sociobiology*, 75(11), Article 152.
- Strandberg, R., Klaassen, R. H. G., Hake, M., & Alerstam, T. (2009). How hazardous is the Sahara Desert crossing for migratory birds? Indications from satellite tracking of raptors. *Biology Letters*, 6(3), 297–300.
- Vickery, J. A., Ewing, S. R., Smith, K. W., Pain, D. J., Bairlein, F., Škorpilová, J., & Gregory, R. D. (2014). Defining the key wintering habitats in the Sahel for declining African-Eurasian migrants using expert assessment. *Bird Conservation International*, 24(4), 477–491.
- Walther, B. A., & Rahbek, C. (2002). Habitat use by Palearctic migrant passerine birds in West Africa. *Ibis*, 144(2), 210–221.
- Wilson, J. M., & Cresswell, W. (2006). How robust are Palearctic migrants to habitat loss and degradation in the Sahel? *Ibis*, 148(4), 789–800.
- Zwarts, L., Bijlsma, R. G., van der Kamp, J., & Wymenga, E. (2009). *Living on the edge: Wetlands and birds in a changing Sahel*. KNNV Publishing.

VIII - ANNEXES

ANNEXE 1 : Script R de nettoyage et filtrage des données ARGOS

```
# =====  
#  
# SCRIPT: ARGOS_CLEANING  
#  
# AUTHOR: Raphaël Musseau – BioSphère Environnement  
# YEAR: 2026  
# LICENSE: MIT  
#  
# DESCRIPTION GÉNÉRALE  
# Ce script nettoie des données ARGOS et conserve une seule  
# localisation par passage satellite opérationnel.  
#  
# -----  
# DÉFINITION OPÉRATIONNELLE D'UN PASSAGE SATELLITE  
#  
# Pour un même individu (ID) et un même satellite (Sat.),  
# un passage est défini comme une séquence continue de  
# localisations ordonnées dans le temps.  
#  
# - Les localisations sont triées chronologiquement.  
# - L'intervalle de temps entre deux localisations successives  
#   est calculé (dt, en minutes).  
# - Si  $dt \leq 15$  minutes → même passage  
# - Si  $dt > 15$  minutes → nouveau passage  
#  
# Le paramètre clé est gap_min = 15 minutes  
# Chaque passage est identifié par un identifiant interne : ep_id  
#  
# -----  
# FILTRAGE QUALITÉ DES LOCALISATIONS  
#  
# Seules les classes ARGOS les plus fiables sont conservées :  
#  
# LC 3, 2, 1  
#  
# Les classes suivantes sont exclues :  
#  
# LC 0, A, B, C  
#  
# Justification :  
# - ces classes présentent une précision insuffisante  
# - elles génèrent des localisations aberrantes (ex : points en mer)  
# - leur inclusion dégrade fortement la lisibilité des trajectoires  
#  
# -----  
# SÉLECTION D'UNE LOCALISATION PAR PASSAGE  
#  
# Pour chaque passage (ID + Sat. + ep_id), une seule localisation  
# est conservée selon les critères hiérarchiques suivants :  
#  
# 1) meilleure classe ARGOS ( $3 > 2 > 1$ )  
# 2) plus petite ellipse (demi-grand axe), si disponible  
# 3) message le plus récent (Date Msg., sinon Date de loc.)  
#  
# -----  
# SORTIE  
#  
# Le script produit :
```

```

#
# - d_final : data.frame nettoyé (1 localisation / passage)
# - argos_clean.csv : fichier exporté dans le dossier source
#
# -----
# REMARQUE MÉTHODOLOGIQUE IMPORTANTE
#
# La définition des passages repose uniquement sur la continuité
# temporelle (gap de 15 minutes) et le satellite.
#
# Le champ "Loc idx" n'est pas utilisé, car il est absent ou
# non exploitable dans certains jeux de données (notamment France).
#
# Cette approche garantit une méthode homogène entre individus et jeux de données.
#
# =====

## PACKAGES ----
suppressPackageStartupMessages({
  library(dplyr)
  library(lubridate)
})

## 1) PARAMÈTRES & SOURCE DES DONNÉES ----
csv_path <- "C:/Users/Utilisateur/Desktop/ARGOS FRANCE
MAROC/DATA_ARGOS_FRANCE_MAROC.csv"

# Nouveau passage si dt > gap_min (minutes), même individu + même satellite
gap_min <- 15

# Ordre de qualité ARGOS (meilleur -> moins bon)
lc_order <- c("3", "2", "1")
lc_rank_map <- setNames(seq_along(lc_order), lc_order)

## 2) FONCTIONS UTILITAIRES ----
to_num <- function(x) {
  suppressWarnings(as.numeric(gsub(",", ".", as.character(x))))
}

to_time <- function(x) {
  x <- trimws(as.character(x))
  x[x == ""] <- NA_character_
  x <- sub(".*$", "", x)
  x <- gsub("T", " ", x, fixed = TRUE)

  suppressWarnings(as.POSIXct(
    x,
    tz = "UTC",
    tryFormats = c(
      "%Y-%m-%d %H:%M:%S",
      "%Y-%m-%d %H:%M",
      "%d/%m/%Y %H:%M:%S",
      "%d/%m/%Y %H:%M"
    )
  ))
}

## 3) LECTURE DU CSV SOURCE ----
raw <- read.csv(
  csv_path,
  sep = ";",
  stringsAsFactors = FALSE,
  check.names = FALSE,

```

```

fileEncoding = "UTF-8"
)

required <- c("ID", "Date de loc.", "Longitude", "Latitude", "Qualite loc.", "Sat.")
if (!all(required %in% names(raw))) {
  stop(
    "Colonnes requises manquantes. Trouvées :\n",
    paste(names(raw), collapse = " | ")
  )
}

has_smaj <- "Demi-grand axe" %in% names(raw)
has_date_msg <- "Date Msg" %in% names(raw)
has_origine <- "Origine" %in% names(raw)

## 4) PRÉPARATION & NORMALISATION ----
d <- tibble(
  id = trimws(as.character(raw[["ID"]])),
  `Sat.` = toupper(trimws(as.character(raw[["Sat."]]))),
  date_loc = to_time(raw[["Date de loc."]]),
  lon = to_num(raw[["Longitude"]]),
  lat = to_num(raw[["Latitude"]]),
  lc = toupper(trimws(as.character(raw[["Qualite loc."]]))),
  Origine = if (has_origine) trimws(as.character(raw[["Origine"]])) else NA_character_
)

if (has_smaj) {
  smaj_raw <- to_num(raw[["Demi-grand axe"]])
  d$smaj_km <- if (isTRUE(stats::median(smaj_raw, na.rm = TRUE) > 50)) smaj_raw / 1000 else smaj_raw
} else {
  d$smaj_km <- NA_real_
}

if (has_date_msg) {
  d$date_msg <- to_time(raw[["Date Msg"]])
} else {
  d$date_msg <- as.POSIXct(NA, tz = "UTC")
}

# Normalisation simple du champ Origine
if (has_origine) {
  d <- d %>%
    mutate(
      Origine = case_when(
        Origine == "France" ~ "France",
        Origine == "Maroc" ~ "Maroc",
        TRUE ~ Origine
      )
    )
}

# Signalement des classes ARGOS non retenues
excluded_lc <- setdiff(sort(unique(d$lc)), lc_order)
if (length(excluded_lc) > 0) {
  cat(
    "\nClasses ARGOS exclues du cleaning :",
    paste(excluded_lc, collapse = ", "),
    "\n"
  )
}

d <- d %>%
  filter(lc %in% lc_order) %>%

```

```

mutate(
  lc_rank = unname(lc_rank_map[lc])
)

## 5) PRÉ-FILTRAGE MINIMAL ----
d <- d %>%
  filter(!is.na(id), id != "") %>%
  filter(!is.na(date_loc), is.finite(lon), is.finite(lat)) %>%
  filter(!is.na(`Sat.`), `Sat.` != "") %>%
  arrange(id, `Sat.`, date_loc)

## 6) DÉFINITION DES PASSAGES ----
d_pass <- d %>%
  group_by(id, `Sat.`) %>%
  arrange(date_loc, .by_group = TRUE) %>%
  mutate(
    dt_min = as.numeric(difftime(date_loc, lag(date_loc), units = "mins")),
    new_ep = ifelse(is.na(dt_min) | dt_min > gap_min, 1L, 0L),
    ep_id = cumsum(new_ep)
  ) %>%
  ungroup()

## 7) UNE UNIQUE LOCALISATION CONSERVÉE PAR PASSAGE ----
d_final <- d_pass %>%
  group_by(id, `Sat.`, ep_id) %>%
  arrange(
    lc_rank,
    is.na(smaj_km), smaj_km,
    dplyr::desc(coalesce(date_msg, date_loc)),
    .by_group = TRUE
  ) %>%
  slice(1) %>%
  ungroup() %>%
  select(id, `Sat.`, date_loc, lon, lat, lc, Origine, smaj_km, date_msg)

## 8) RÉSUMÉ CONSOLE ----
cat("\n=== ARGOS CLEANING – RÉSUMÉ ===\n")
cat("Lignes brutes      :", nrow(raw), "\n")
cat("Après filtrage + pré-filtrage :", nrow(d), "\n")
cat("Final (1 loc / passage)      :", nrow(d_final), "\n\n")

cat("Effectifs par LC (final) :\n")
print(table(d_final$lc, useNA = "ifany"))

cat("\nRépartition Origine (final) :\n")
print(table(d_final$Origine, useNA = "ifany"))

if (has_smaj) {
  cat("\nRésumé smaj_km (final) :\n")
  print(summary(d_final$smaj_km))
}

cat("\nAperçu d_final :\n")
print(utils::head(d_final, 10))

## 9) SAUVEGARDE AUTOMATIQUE DANS LE DOSSIER SOURCE ----
outfile <- file.path(dirname(csv_path), "argos_clean.csv")

ok_write <- tryCatch(
  {
    write.csv(
      d_final,
      outfile,

```

```

    row.names = FALSE,
    fileEncoding = "UTF-8"
  )
  TRUE
},
error = function(e) {
  message("Erreur d'écriture : ", conditionMessage(e))
  FALSE
}
)

if (ok_write) {
  cat("\nFichier enregistré dans :", outfile, "\n")
} else {
  cat("\nLe fichier n'a pas pu être enregistré.\n")
  cat("Vérifie qu'il n'est pas déjà ouvert dans Excel ou protégé en écriture.\n")
}

## FIN DU SCRIPT

```

ANNEXE 2 : Script R de clustering spatio-temporel des haltes migratoires

```
# =====
#
# SCRIPT : ARGOS_SPATIOTEMPORAL_CLUSTERING.R
#
# OBJECTIF
# Détecter des spots spatio-temporels de migration / hivernage
# uniquement chez les individus ayant franchi le Sahara.
#
# CRITÈRE INDIVIDU RETENU
# Un individu est retenu s'il franchit successivement :
#   1) 30°N vers le sud
#   2) 20°N vers le sud
#
# FILTRAGE QUALITÉ
# - classes ARGOS conservées : 1, 2, 3
# - coordonnées valides uniquement
# - champ Origine conservé dans les sorties
#
# SORTIES
# - argos_spots_points.csv
# - argos_spots_summary.csv
# - argos_spots_points.kml
# - argos_spots_centroids.kml
# - argos_spots_polygons.kml
#
# =====

suppressPackageStartupMessages({
  library(dplyr)
  library(lubridate)
  library(geosphere)
  library(sf)
  library(tcltk)
})

## 1) PARAMÈTRES ----

input_file <- "C:/Users/Utilisateur/Desktop/ARGOS FRANCE MAROC/argos_clean.csv"

# Seuils Sahara / Sahel
lat_sahara_north <- 30
lat_sahara_south <- 20

# Clustering
T_gap_hours <- 72
R_max_km <- 40
R_fusion_km <- 20
N_min_points <- 2
duration_min_days <- 1

## 2) FONCTIONS ----

to_time <- function(x) {
  suppressWarnings(as.POSIXct(
    as.character(x),
    tz = "UTC",
    tryFormats = c(
      "%Y-%m-%d %H:%M:%S",
      "%Y-%m-%d %H:%M",
      "%d/%m/%Y %H:%M:%S",
      "%d/%m/%Y %H:%M"
    )
  ))
}
```

```

    )
  })
}

safe_write_csv <- function(x, file) {
  tryCatch({
    write.csv(x, file, row.names = FALSE, fileEncoding = "UTF-8")
    TRUE
  }, error = function(e) {
    message("Erreur d'écriture de : ", file)
    message("→ Vérifie que le fichier n'est pas ouvert dans Excel.")
    message("Détail : ", conditionMessage(e))
    FALSE
  })
}

collapse_unique <- function(x) {
  x <- unique(trimws(as.character(x)))
  x <- x[!is.na(x) & x != ""]
  if (length(x) == 0) return(NA_character_)
  paste(x, collapse = " | ")
}

cluster_segment <- function(df_seg, R_max_km) {
  n <- nrow(df_seg)

  if (n == 0) return(df_seg %>% mutate(cluster_seq = NA_integer_))

  if (n == 1) {
    df_seg$cluster_seq <- 1L
    return(df_seg)
  }

  cluster_seq <- integer(n)
  current_k <- 1L
  cluster_seq[1] <- current_k

  cx <- df_seg$lon[1]
  cy <- df_seg$lat[1]
  c_n <- 1L

  for (i in 2:n) {
    dist_km <- geosphere::distHaversine(
      c(cx, cy),
      c(df_seg$lon[i], df_seg$lat[i])
    ) / 1000

    if (!is.finite(dist_km) || dist_km > R_max_km) {
      current_k <- current_k + 1L
      cx <- df_seg$lon[i]
      cy <- df_seg$lat[i]
      c_n <- 1L
    } else {
      c_n <- c_n + 1L
      cx <- (cx * (c_n - 1L) + df_seg$lon[i]) / c_n
      cy <- (cy * (c_n - 1L) + df_seg$lat[i]) / c_n
    }

    cluster_seq[i] <- current_k
  }

  df_seg$cluster_seq <- cluster_seq
  df_seg
}

```

```

}

## 3) LECTURE ----

if (!file.exists(input_file)) stop("Fichier d'entrée introuvable : ", input_file)

d_raw <- read.csv(
  input_file,
  stringsAsFactors = FALSE,
  check.names = FALSE
)

required_cols <- c("id", "date_loc", "lon", "lat", "lc", "Origine")
missing_cols <- setdiff(required_cols, names(d_raw))

if (length(missing_cols) > 0) {
  stop("Colonnes manquantes : ", paste(missing_cols, collapse = ", "))
}

d_raw$date_loc <- to_time(d_raw$date_loc)

d <- d_raw %>%
  mutate(
    id = trimws(as.character(id)),
    lon = suppressWarnings(as.numeric(lon)),
    lat = suppressWarnings(as.numeric(lat)),
    lc = toupper(trimws(as.character(lc))),
    Origine = trimws(as.character(Origine)),
    Origine = case_when(
      tolower(Origine) == "france" ~ "France",
      tolower(Origine) == "maroc" ~ "Maroc",
      TRUE ~ Origine
    )
  )

## 4) FILTRAGE QUALITÉ ----

bad_coords <- d %>%
  filter(
    is.na(date_loc) |
    !is.finite(lon) | !is.finite(lat) |
    lon < -180 | lon > 180 |
    lat < -90 | lat > 90
  )

cat("\n=== COORDONNÉES INVALIDES EXCLUES ===\n")
cat("Nombre :", nrow(bad_coords), "\n")

if (nrow(bad_coords) > 0) {
  print(head(bad_coords[, c("id", "date_loc", "lon", "lat", "lc", "Origine")], 20))
}

d <- d %>%
  filter(!is.na(id), id != "") %>%
  filter(!is.na(date_loc)) %>%
  filter(is.finite(lon), is.finite(lat)) %>%
  filter(lon >= -180, lon <= 180, lat >= -90, lat <= 90) %>%
  filter(lc %in% c("1", "2", "3")) %>%
  arrange(id, date_loc)

cat("\n=== APRÈS FILTRAGE QUALITÉ ===\n")
cat("Lignes brutes :", nrow(d_raw), "\n")
cat("Lignes conservées LC 1/2/3 :", nrow(d), "\n")

```

```

cat("\nEffectifs par LC :\n")
print(table(d$lc, useNA = "ifany"))

cat("\nEffectifs par Origine :\n")
print(table(d$Origine, useNA = "ifany"))

if (nrow(d) == 0) stop("Aucune donnée disponible après filtrage.")

## 5) FILTRE INDIVIDUS AYANT TRAVERSÉ LE SAHARA ----

crossings <- d %>%
  group_by(id) %>%
  arrange(date_loc, .by_group = TRUE) %>%
  mutate(
    lat_prev = lag(lat),
    cross_30S = !is.na(lat_prev) & lat_prev > lat_sahara_north & lat <= lat_sahara_north,
    cross_20S = !is.na(lat_prev) & lat_prev > lat_sahara_south & lat <= lat_sahara_south
  ) %>%
  summarise(
    Origine = collapse_unique(Origine),
    date_30N = {
      x <- date_loc[cross_30S]
      if (length(x) == 0) as.POSIXct(NA, tz = "UTC") else min(x, na.rm = TRUE)
    },
    date_20N = {
      x <- date_loc[cross_20S]
      if (length(x) == 0) as.POSIXct(NA, tz = "UTC") else min(x, na.rm = TRUE)
    },
    .groups = "drop"
  ) %>%
  mutate(
    crossed_sahara = !is.na(date_30N) & !is.na(date_20N) & date_20N > date_30N
  )

cat("\n=== INDIVIDUS ET FRANCHISSEMENT SAHARA ===\n")
print(crossings)

sahara_migrants <- crossings %>%
  filter(crossed_sahara)

cat("\n=== INDIVIDUS CONSERVÉS POUR ANALYSE CLUSTERS ===\n")
cat("Nombre :", nrow(sahara_migrants), "\n")
print(table(sahara_migrants$Origine, useNA = "ifany"))

if (nrow(sahara_migrants) == 0) {
  stop("Aucun individu n'a franchi successivement 30°N puis 20°N vers le sud.")
}

d <- d %>%
  filter(id %in% sahara_migrants$id) %>%
  mutate(row_id = dplyr::row_number())

cat("\nPoints conservés après filtre Sahara :", nrow(d), "\n")

## 6) SEGMENTATION TEMPORELLE ----

d_seg <- d %>%
  group_by(id) %>%
  arrange(date_loc, .by_group = TRUE) %>%
  mutate(
    dt_h = as.numeric(difftime(date_loc, lag(date_loc), units = "hours")),
    segment = cumsum(ifelse(is.na(dt_h) | dt_h > T_gap_hours, 1L, 0L))
  )

```

```

) %>%
ungroup()

## 7) CLUSTERING SÉQUENTIEL ----

d_clustered <- d_seg %>%
  arrange(id, segment, date_loc) %>%
  group_by(id, segment) %>%
  group_modify(~ cluster_segment(.x, R_max_km = R_max_km)) %>%
  ungroup()

clusters_seq <- d_clustered %>%
  filter(!is.na(cluster_seq)) %>%
  group_by(id, segment, cluster_seq) %>%
  summarise(
    n_points    = n(),
    start_time  = min(date_loc),
    end_time    = max(date_loc),
    duration_days = as.numeric(difftime(end_time, start_time, units = "days")),
    lon_centroid = mean(lon, na.rm = TRUE),
    lat_centroid = mean(lat, na.rm = TRUE),
    Origine     = collapse_unique(Origine),
    .groups     = "drop"
  )

cat("\n=== CLUSTERS SÉQUENTIELS AVANT FUSION ===\n")
print(head(clusters_seq, 10))

## 8) FUSION DES CLUSTERS PROCHES ----

if (nrow(clusters_seq) > 0) {

  clusters_merged <- clusters_seq %>%
    arrange(id, start_time) %>%
    group_by(id) %>%
    mutate(
      prev_lon = lag(lon_centroid),
      prev_lat = lag(lat_centroid),
      dist_prev_km = ifelse(
        is.na(prev_lon),
        NA_real_,
        geosphere::distHaversine(
          cbind(prev_lon, prev_lat),
          cbind(lon_centroid, lat_centroid)
        ) / 1000
      ),
      new_group = ifelse(is.na(dist_prev_km) | dist_prev_km > R_fusion_km, 1L, 0L),
      final_group_index = cumsum(new_group)
    ) %>%
    select(-prev_lon, -prev_lat, -new_group) %>%
    ungroup()

  clusters_final <- clusters_merged %>%
    group_by(id, final_group_index) %>%
    summarise(
      n_points    = sum(n_points),
      start_time  = min(start_time),
      end_time    = max(end_time),
      duration_days = as.numeric(difftime(end_time, start_time, units = "days")),
      lon_centroid = sum(lon_centroid * n_points) / sum(n_points),
      lat_centroid = sum(lat_centroid * n_points) / sum(n_points),
      Origine     = collapse_unique(Origine),
      .groups     = "drop"
    )
}

```

```

) %>%
mutate(
  spot_id = paste0("SPOT_", id, "_", final_group_index)
)

d_clustered <- d_clustered %>%
left_join(
  clusters_merged %>%
  select(id, segment, cluster_seq, final_group_index),
  by = c("id", "segment", "cluster_seq")
)

clusters_final <- clusters_final %>%
rowwise() %>%
mutate(
  radius_km = {
    pts <- d_clustered %>%
    filter(
      id == .data$id,
      final_group_index == .data$final_group_index
    )

    if (nrow(pts) == 0) {
      NA_real_
    } else {
      max(
        geosphere::distHaversine(
          cbind(lon_centroid, lat_centroid),
          cbind(pts$lon, pts$lat)
        ) / 1000,
        na.rm = TRUE
      )
    }
  }
) %>%
ungroup()

} else {
clusters_final <- tibble()
d_clustered$final_group_index <- NA_integer_
}

## 9) FILTRE DES SPOTS VALIDES ----

if (nrow(clusters_final) > 0) {

valid_spots <- clusters_final %>%
filter(
  n_points >= N_min_points,
  duration_days >= duration_min_days
)

d_clustered <- d_clustered %>%
left_join(
  valid_spots %>%
  select(id, final_group_index, spot_id),
  by = c("id", "final_group_index")
) %>%
mutate(
  is_spot_point = !is.na(spot_id)
)

} else {

```

```

valid_spots <- tibble()
d_clustered$spot_id <- NA_character_
d_clustered$is_spot_point <- FALSE
}

cat("\n=== SPOTS VALIDES ===\n")
cat("Nombre :", nrow(valid_spots), "\n")

if (nrow(valid_spots) > 0) {
  print(table(valid_spots$Origine, useNA = "ifany"))
}

cat("\nNombre de points assignés à un spot : ",
    sum(d_clustered$is_spot_point, na.rm = TRUE), "\n")

## 10) OBJETS SF ----

points_sf <- st_as_sf(
  d_clustered,
  coords = c("lon", "lat"),
  crs = 4326,
  remove = FALSE
)

if (nrow(valid_spots) > 0) {

  spots_sf <- st_as_sf(
    valid_spots,
    coords = c("lon_centroid", "lat_centroid"),
    crs = 4326,
    remove = FALSE
  )

  spots_poly_sf <- points_sf %>%
    filter(is_spot_point) %>%
    group_by(id, spot_id, final_group_index, Origine) %>%
    filter(n() >= 3) %>%
    summarise(
      geometry = sf::st_convex_hull(sf::st_union(geometry)),
      .groups = "drop"
    ) %>%
    left_join(
      valid_spots,
      by = c("id", "final_group_index", "spot_id", "Origine")
    )

} else {
  spots_sf <- NULL
  spots_poly_sf <- NULL
}

## 11) DOSSIER DE SORTIE ----

cat("\nUne fenêtre va s'ouvrir pour choisir le dossier de sortie...\n")

dest_dir <- tcltk::tk_choose.dir(
  caption = "Choisissez un dossier pour enregistrer les fichiers"
)

if (is.na(dest_dir) || !nzchar(dest_dir)) {
  stop("Aucun dossier choisi. Arrêt du script.")
}

```

```

dest_dir <- normalizePath(dest_dir, winslash = "/", mustWork = TRUE)

outfile_points <- file.path(dest_dir, "argos_spots_points.csv")
outfile_spots <- file.path(dest_dir, "argos_spots_summary.csv")
outfile_kml_pts <- file.path(dest_dir, "argos_spots_points.kml")
outfile_kml_cent <- file.path(dest_dir, "argos_spots_centroids.kml")
outfile_kml_poly <- file.path(dest_dir, "argos_spots_polygons.kml")

## 12) EXPORTS ----

ok_points <- safe_write_csv(d_clustered, outfile_points)
ok_spots <- safe_write_csv(valid_spots, outfile_spots)

st_write(
  points_sf,
  outfile_kml_pts,
  driver = "KML",
  layer = "TrackPoints",
  append = FALSE,
  quiet = TRUE
)

if (!is.null(spots_sf) && nrow(spots_sf) > 0) {
  st_write(
    spots_sf,
    outfile_kml_cent,
    driver = "KML",
    layer = "SpotsCentroids",
    append = FALSE,
    quiet = TRUE
  )
}

if (!is.null(spots_poly_sf) && nrow(spots_poly_sf) > 0) {
  st_write(
    spots_poly_sf,
    outfile_kml_poly,
    driver = "KML",
    layer = "SpotsPolygons",
    append = FALSE,
    quiet = TRUE
  )
}

cat("\n=== EXPORT TERMINÉ ===\n")
cat("Points + annotations :", outfile_points, if (!ok_points) " (échec écriture)" else "", "\n")
cat("Résumé des spots :", outfile_spots, if (!ok_spots) " (échec écriture)" else "", "\n")
cat("KML points :", outfile_kml_pts, "\n")

if (!is.null(spots_sf) && nrow(spots_sf) > 0) {
  cat("KML centroïdes spots :", outfile_kml_cent, "\n")
}

if (!is.null(spots_poly_sf) && nrow(spots_poly_sf) > 0) {
  cat("KML polygones spots :", outfile_kml_poly, "\n")
} else {
  cat("Pas de polygones exportés : aucun spot valide avec >= 3 points.\n")
}

## FIN DU SCRIPT

```

ANNEXE 3 : Script R d'analyse de la traversée saharienne

```
# =====
# SCRIPT : TRAVERSÉE DU SAHARA 30°N -> 20°N
#
# OBJECTIF
#   Tester les différences entre origines pour :
#   - date de franchissement de 30°N vers le sud
#   - date de franchissement de 20°N vers le sud
#   - durée de traversée saharienne entre 30°N et 20°N
#
# BASE
#   argos_clean.csv
#   supposé déjà filtré sur les localisations ARGOS 1, 2 et 3.
#
# SORTIES
#   - sahara_crossing_30N_20N.csv
#   - résumé par origine
#   - tests Wilcoxon exploratoires
#
# =====

suppressPackageStartupMessages({
  library(dplyr)
  library(lubridate)
})

## 1) PARAMÈTRES ----

input_file <- "C:/Users/Utilisateur/Desktop/ARGOS FRANCE MAROC/argos_clean.csv"

lat_sahara_north <- 30
lat_sahara_south <- 20

if (!file.exists(input_file)) {
  stop("Fichier introuvable : ", input_file)
}

## 2) LECTURE ----

d <- read.csv(
  input_file,
  stringsAsFactors = FALSE,
  check.names = FALSE
)

required_cols <- c("id", "date_loc", "lon", "lat", "Origine")
missing_cols <- setdiff(required_cols, names(d))

if (length(missing_cols) > 0) {
  stop("Colonnes manquantes : ", paste(missing_cols, collapse = ", "))
}

d$date_loc <- as.POSIXct(
  as.character(d$date_loc),
  tz = "UTC",
  tryFormats = c(
    "%Y-%m-%d %H:%M:%S",
    "%Y-%m-%d %H:%M",
    "%d/%m/%Y %H:%M:%S",
    "%d/%m/%Y %H:%M"
  )
)
```

3) NETTOYAGE MINIMAL ----

```
d <- d %>%
  mutate(
    id = trimws(as.character(id)),
    lon = suppressWarnings(as.numeric(lon)),
    lat = suppressWarnings(as.numeric(lat)),
    Origine = trimws(as.character(Origine)),
    Origine = case_when(
      tolower(Origine) == "france" ~ "France",
      tolower(Origine) == "maroc" ~ "Maroc",
      TRUE ~ Origine
    )
  ) %>%
  filter(!is.na(id), id != "") %>%
  filter(!is.na(date_loc)) %>%
  filter(is.finite(lon), is.finite(lat)) %>%
  filter(lon >= -180, lon <= 180, lat >= -90, lat <= 90) %>%
  arrange(id, date_loc)

cat("\n=== DONNÉES UTILISÉES ===\n")
cat("Nombre de points :", nrow(d), "\n")
cat("Nombre d'individus :", length(unique(d$id)), "\n")

cat("\nEffectifs par origine :\n")
print(table(d$Origine, useNA = "ifany"))

if ("lc" %in% names(d)) {
  cat("\nClasses ARGOS présentes :\n")
  print(table(d$lc, useNA = "ifany"))
}
```

4) DÉTECTION DES FRANCHISSEMENTS VERS LE SUD ----

```
d_cross <- d %>%
  group_by(id) %>%
  arrange(date_loc, .by_group = TRUE) %>%
  mutate(
    lat_prev = lag(lat),

    cross_30S = !is.na(lat_prev) &
      lat_prev > lat_sahara_north &
      lat <= lat_sahara_north,

    cross_20S = !is.na(lat_prev) &
      lat_prev > lat_sahara_south &
      lat <= lat_sahara_south
  ) %>%
  ungroup()

cross_30 <- d_cross %>%
  filter(cross_30S) %>%
  group_by(id) %>%
  slice(1) %>%
  ungroup() %>%
  transmute(
    id,
    Origine,
    date_30N = date_loc,
    lon_30N = lon,
    lat_30N = lat
  )
```

```

cross_20 <- d_cross %>%
  filter(cross_20S) %>%
  group_by(id) %>%
  slice(1) %>%
  ungroup() %>%
  transmute(
    id,
    date_20N = date_loc,
    lon_20N = lon,
    lat_20N = lat
  )

```

5) TABLE INDIVIDUELLE DE TRAVERSÉE ----

```

sahara_crossing <- cross_30 %>%
  left_join(cross_20, by = "id") %>%
  mutate(
    duration_days = as.numeric(difftime(date_20N, date_30N, units = "days")),
    doy_30N = yday(date_30N),
    doy_20N = yday(date_20N),
    crossed_sahara = !is.na(date_30N) &
      !is.na(date_20N) &
      date_20N > date_30N
  ) %>%
  arrange(Origine, id)

```

```

sahara_complete <- sahara_crossing %>%
  filter(crossed_sahara)

```

6) AFFICHAGE DES EFFECTIFS ----

```

cat("\n=== EFFECTIFS ===\n")

cat("\nIndividus avec passage 30°N vers le sud :\n")
print(table(cross_30$Origine, useNA = "ifany"))

cat("\nIndividus avec traversée complète 30°N -> 20°N :\n")
print(table(sahara_complete$Origine, useNA = "ifany"))

cat("\n=== TABLE INDIVIDUELLE TRAVERSÉE SAHARA ===\n")
print(sahara_complete)

```

7) RÉSUMÉ PAR ORIGINE ----

```

summary_crossing <- sahara_complete %>%
  group_by(Origine) %>%
  summarise(
    n = n(),

    mean_doy_30N = mean(doy_30N, na.rm = TRUE),
    sd_doy_30N = sd(doy_30N, na.rm = TRUE),
    min_doy_30N = min(doy_30N, na.rm = TRUE),
    max_doy_30N = max(doy_30N, na.rm = TRUE),

    mean_doy_20N = mean(doy_20N, na.rm = TRUE),
    sd_doy_20N = sd(doy_20N, na.rm = TRUE),
    min_doy_20N = min(doy_20N, na.rm = TRUE),
    max_doy_20N = max(doy_20N, na.rm = TRUE),

    mean_duration_days = mean(duration_days, na.rm = TRUE),
    sd_duration_days = sd(duration_days, na.rm = TRUE),
    min_duration_days = min(duration_days, na.rm = TRUE),

```

```

max_duration_days = max(duration_days, na.rm = TRUE),

.groups = "drop"
)

cat("\n=== RÉSUMÉ PAR ORIGINE ===\n")
print(summary_crossing)

## 8) TESTS EXPLORATOIRES ----

if (length(unique(na.omit(sahara_complete$Origine))) == 2) {

  cat("\n=== TEST WILCOXON : DATE DE PASSAGE 30°N ===\n")
  print(wilcox.test(doy_30N ~ Origine, data = sahara_complete))

  cat("\n=== TEST WILCOXON : DATE DE PASSAGE 20°N ===\n")
  print(wilcox.test(doy_20N ~ Origine, data = sahara_complete))

  cat("\n=== TEST WILCOXON : DURÉE DE TRAVERSÉE SAHARIENNE ===\n")
  print(wilcox.test(duration_days ~ Origine, data = sahara_complete))
}

## 9) EXPORTS ----

out_dir <- dirname(input_file)

outfile_crossing <- file.path(out_dir, "sahara_crossing_30N_20N.csv")
outfile_summary <- file.path(out_dir, "sahara_crossing_summary_by_origin.csv")

write.csv(
  sahara_complete,
  outfile_crossing,
  row.names = FALSE,
  fileEncoding = "UTF-8"
)

write.csv(
  summary_crossing,
  outfile_summary,
  row.names = FALSE,
  fileEncoding = "UTF-8"
)

cat("\n=== EXPORT TERMINÉ ===\n")
cat("Table individuelle :", outfile_crossing, "\n")
cat("Résumé par origine :", outfile_summary, "\n")

## FIN DU SCRIPT

```

ANNEXE 4 : Script R de délimitation des zones d'hivernage

```
# =====
# SCRIPT : CLUSTERS HIVERNAUX SAHEL - SECTEURS REGIONAUX
#
# Objectif :
# Identifier des secteurs d'utilisation hivernale au Sahel.
#
# Critères de sélection :
# - hivernage avec au moins 90 jours de présence au Sahel
# - au moins 50 localisations
#
# Clustering :
# - rupture temporelle si gap > 14 jours
# - noyaux locaux : 50 km
# - fusion régionale : 100 km
# =====

suppressPackageStartupMessages({
  library(dplyr)
  library(lubridate)
  library(geosphere)
  library(sf)
})

input_file <- "C:/Users/Utilisateur/Desktop/ARGOS FRANCE MAROC/argos_clean.csv"

min_winter_duration_days <- 90
min_winter_points <- 50

T_gap_hours <- 336
R_max_km <- 50
R_fusion_km <- 100
N_min_points <- 5
duration_min_days <- 7

to_time <- function(x) {
  suppressWarnings(as.POSIXct(
    as.character(x),
    tz = "UTC",
    tryFormats = c(
      "%Y-%m-%d %H:%M:%S",
      "%Y-%m-%d %H:%M",
      "%d/%m/%Y %H:%M:%S",
      "%d/%m/%Y %H:%M"
    )
  ))
}

cluster_segment <- function(df_seg, R_max_km) {
  n <- nrow(df_seg)

  if (n == 0) return(df_seg %>% mutate(cluster_seq = NA_integer_))

  if (n == 1) {
    df_seg$cluster_seq <- 1L
    return(df_seg)
  }

  cluster_seq <- integer(n)
  current_k <- 1L
  cluster_seq[1] <- current_k
```

```

cx <- df_seg$lon[1]
cy <- df_seg$lat[1]
c_n <- 1L

for (i in 2:n) {
  dist_km <- geosphere::distHaversine(
    c(cx, cy),
    c(df_seg$lon[i], df_seg$lat[i])
  ) / 1000

  if (!is.finite(dist_km) || dist_km > R_max_km) {
    current_k <- current_k + 1L
    cx <- df_seg$lon[i]
    cy <- df_seg$lat[i]
    c_n <- 1L
  } else {
    c_n <- c_n + 1L
    cx <- (cx * (c_n - 1L) + df_seg$lon[i]) / c_n
    cy <- (cy * (c_n - 1L) + df_seg$lat[i]) / c_n
  }

  cluster_seq[i] <- current_k
}

df_seg$cluster_seq <- cluster_seq
df_seg
}

safe_write_csv <- function(x, file) {
  tryCatch({
    write.csv(x, file, row.names = FALSE, fileEncoding = "UTF-8")
    TRUE
  }, error = function(e) {
    message("Erreur d'écriture : ", conditionMessage(e))
    FALSE
  })
}

if (!file.exists(input_file)) stop("Fichier introuvable : ", input_file)

d <- read.csv(input_file, stringsAsFactors = FALSE, check.names = FALSE)

required_cols <- c("id", "date_loc", "lon", "lat", "Origine")
if (!all(required_cols %in% names(d))) {
  stop("Colonnes manquantes : ", paste(setdiff(required_cols, names(d)), collapse = ", "))
}

d$date_loc <- to_time(d$date_loc)

d <- d %>%
  mutate(
    id = trimws(as.character(id)),
    lon = suppressWarnings(as.numeric(lon)),
    lat = suppressWarnings(as.numeric(lat)),
    Origine = trimws(as.character(Origine)),
    Origine = case_when(
      tolower(Origine) == "france" ~ "France",
      tolower(Origine) == "maroc" ~ "Maroc",
      TRUE ~ Origine
    )
  ) %>%
  filter(!is.na(id), id != "") %>%
  filter(!is.na(date_loc), is.finite(lon), is.finite(lat)) %>%

```

```

filter(lon >= -180, lon <= 180, lat >= -90, lat <= 90) %>%
arrange(id, date_loc)

cat("\n=== DONNÉES D'ENTRÉE ===\n")
cat("Points :", nrow(d), "\n")
print(table(d$Origine, useNA = "ifany"))

## 1) Détection entrée/sortie Sahel ----

d_cross <- d %>%
  group_by(id) %>%
  arrange(date_loc, .by_group = TRUE) %>%
  mutate(
    lat_prev = lag(lat),
    crossing_south_20 = !is.na(lat_prev) & lat_prev > 20 & lat <= 20,
    crossing_north_20 = !is.na(lat_prev) & lat_prev <= 20 & lat > 20
  ) %>%
  ungroup()

entries <- d_cross %>%
  filter(crossing_south_20) %>%
  transmute(
    id,
    Origine,
    winter_year = year(date_loc),
    date_start = date_loc
  )

exits <- d_cross %>%
  filter(crossing_north_20) %>%
  transmute(
    id,
    date_end = date_loc
  )

winter_periods <- entries %>%
  rowwise() %>%
  mutate(
    date_end = {
      this_id <- id
      this_start <- date_start

      e <- exits %>%
        filter(id == this_id, date_end > this_start) %>%
        arrange(date_end)

      if (nrow(e) == 0) {
        as.POSIXct(NA, tz = "UTC")
      } else {
        e$date_end[1]
      }
    }
  ) %>%
  ungroup() %>%
  mutate(
    winter_id = paste0(id, "_", winter_year, "_", winter_year + 1)
  )

cat("\n=== PÉRIODES D'HIVERNAGE DÉTECTÉES ===\n")
print(winter_periods)

## 2) Extraction des points hivernaux ----

```

```

winter_points_list <- list()

for (i in seq_len(nrow(winter_periods))) {

  this_id <- winter_periods$id[i]
  this_start <- winter_periods$date_start[i]
  this_end <- winter_periods$date_end[i]
  this_wid <- winter_periods$winter_id[i]
  this_year <- winter_periods$winter_year[i]
  this_orig <- winter_periods$Origine[i]

  tmp <- d %>%
    filter(id == this_id, date_loc >= this_start) %>%
    { if (!is.na(this_end)) filter(., date_loc <= this_end) else . } %>%
    filter(lat <= 20) %>%
    mutate(
      winter_id = this_wid,
      winter_year = this_year,
      Origine = this_orig
    )

  winter_points_list[[i]] <- tmp
}

winter_points <- bind_rows(winter_points_list)

if (nrow(winter_points) == 0) stop("Aucun point hivernal détecté.")

## 3) Filtre des hivernages exploitables ----

winter_summary_all <- winter_points %>%
  group_by(id, winter_id, Origine, winter_year) %>%
  summarise(
    date_start_obs = min(date_loc, na.rm = TRUE),
    date_end_obs = max(date_loc, na.rm = TRUE),
    duration_days = as.numeric(difftime(date_end_obs, date_start_obs, units = "days")),
    n_points = n(),
    .groups = "drop"
  )

winter_summary_valid <- winter_summary_all %>%
  filter(
    duration_days >= min_winter_duration_days,
    n_points >= min_winter_points
  )

valid_winter_ids <- winter_summary_valid$winter_id

winter_points_valid <- winter_points %>%
  filter(winter_id %in% valid_winter_ids)

cat("\n=== HIVERNAGES VALIDES — >= 90 jours au Sahel et >= 50 points ===\n")
print(winter_summary_valid)

cat("\n=== POINTS HIVERNAUX VALIDES PAR ORIGINE ===\n")
print(table(winter_points_valid$Origine, useNA = "ifany"))

if (nrow(winter_points_valid) == 0) {
  stop("Aucun hivernage valide après filtre durée/points.")
}

## 4) Segmentation temporelle ----

```

```
winter_seg <- winter_points_valid %>%
  arrange(id, winter_id, date_loc) %>%
  group_by(id, winter_id) %>%
  mutate(
    dt_h = as.numeric(difftime(date_loc, lag(date_loc), units = "hours")),
    segment = cumsum(ifelse(is.na(dt_h) | dt_h > T_gap_hours, 1L, 0L))
  ) %>%
  ungroup()
```

5) Clustering séquentiel ----

```
winter_clustered <- winter_seg %>%
  arrange(id, winter_id, segment, date_loc) %>%
  group_by(id, winter_id, segment) %>%
  group_modify(~ cluster_segment(.x, R_max_km = R_max_km)) %>%
  ungroup()

clusters_seq <- winter_clustered %>%
  filter(!is.na(cluster_seq)) %>%
  group_by(id, winter_id, Origine, winter_year, segment, cluster_seq) %>%
  summarise(
    n_points_seq = n(),
    start_time = min(date_loc),
    end_time = max(date_loc),
    duration_days = as.numeric(difftime(end_time, start_time, units = "days")),
    lon_centroid = mean(lon, na.rm = TRUE),
    lat_centroid = mean(lat, na.rm = TRUE),
    .groups = "drop"
  )
```

6) Fusion régionale des clusters ----

```
clusters_merged <- clusters_seq %>%
  arrange(id, winter_id, start_time) %>%
  group_by(id, winter_id) %>%
  mutate(
    prev_lon = lag(lon_centroid),
    prev_lat = lag(lat_centroid),
    dist_prev_km = ifelse(
      is.na(prev_lon),
      NA_real_,
      geosphere::distHaversine(
        cbind(prev_lon, prev_lat),
        cbind(lon_centroid, lat_centroid)
      ) / 1000
    ),
    new_group = ifelse(is.na(dist_prev_km) | dist_prev_km > R_fusion_km, 1L, 0L),
    winter_cluster_index = cumsum(new_group)
  ) %>%
  select(-prev_lon, -prev_lat, -new_group) %>%
  ungroup()
```

```
clusters_final <- clusters_merged %>%
  group_by(id, winter_id, Origine, winter_year, winter_cluster_index) %>%
  summarise(
    n_points = sum(n_points_seq),
    start_time = min(start_time),
    end_time = max(end_time),
    duration_days = as.numeric(difftime(end_time, start_time, units = "days")),
    lon_centroid = sum(lon_centroid * n_points_seq) / sum(n_points_seq),
    lat_centroid = sum(lat_centroid * n_points_seq) / sum(n_points_seq),
    .groups = "drop"
  ) %>%
```

```

mutate(
  winter_cluster_id = paste0(winter_id, "_CL", winter_cluster_index)
) %>%
filter(
  n_points >= N_min_points,
  duration_days >= duration_min_days
)

cat("\n=== CLUSTERS HIVERNAUX VALIDES ===\n")
cat("Nombre de clusters :", nrow(clusters_final), "\n")
print(table(clusters_final$Origine, useNA = "ifany"))

```

7) Attribution des clusters aux points ----

```

winter_clustered <- winter_clustered %>%
left_join(
  clusters_merged %>%
  select(id, winter_id, segment, cluster_seq, winter_cluster_index),
  by = c("id", "winter_id", "segment", "cluster_seq")
) %>%
left_join(
  clusters_final %>%
  select(id, winter_id, winter_cluster_index, winter_cluster_id),
  by = c("id", "winter_id", "winter_cluster_index")
) %>%
mutate(
  is_winter_cluster_point = !is.na(winter_cluster_id)
)

```

8) Rayon des clusters ----

```

if (nrow(clusters_final) > 0) {
  clusters_final <- clusters_final %>%
  rowwise() %>%
  mutate(
    radius_km = {
      pts <- winter_clustered %>%
      filter(
        id == .data$id,
        winter_id == .data$winter_id,
        winter_cluster_id == .data$winter_cluster_id
      )

      if (nrow(pts) == 0) {
        NA_real_
      } else {
        max(
          geosphere::distHaversine(
            cbind(lon_centroid, lat_centroid),
            cbind(pts$lon, pts$lat)
          ) / 1000,
          na.rm = TRUE
        )
      }
    }
  ) %>%
  ungroup()
}

```

9) Objets spatiaux ----

```

points_sf <- st_as_sf(
  winter_clustered,

```

```

coords = c("lon", "lat"),
crs = 4326,
remove = FALSE
)

centroids_sf <- st_as_sf(
  clusters_final,
  coords = c("lon_centroid", "lat_centroid"),
  crs = 4326,
  remove = FALSE
)

cluster_poly_sf <- points_sf %>%
  filter(is_winter_cluster_point) %>%
  group_by(id, winter_id, Origine, winter_year, winter_cluster_id) %>%
  filter(n() >= 3) %>%
  summarise(
    geometry = st_convex_hull(st_union(geometry)),
    .groups = "drop"
  ) %>%
  left_join(
    clusters_final,
    by = c("id", "winter_id", "Origine", "winter_year", "winter_cluster_id")
  )

if (nrow(cluster_poly_sf) > 0) {
  cluster_poly_eq <- st_transform(cluster_poly_sf, 8857)
  cluster_poly_sf$cluster_area_km2 <- as.numeric(st_area(cluster_poly_eq)) / 1e6
}

## 10) Exports ----

out_dir <- dirname(input_file)

outfile_points <- file.path(out_dir, "winter_clusters_points.csv")
outfile_summary <- file.path(out_dir, "winter_clusters_summary.csv")
outfile_centroids <- file.path(out_dir, "winter_clusters_centroids.gpkg")
outfile_polygons <- file.path(out_dir, "winter_clusters_polygons.gpkg")

ok_points <- safe_write_csv(winter_clustered, outfile_points)
ok_summary <- safe_write_csv(clusters_final, outfile_summary)

if (nrow(centroids_sf) > 0) {
  st_write(
    centroids_sf,
    outfile_centroids,
    delete_dsn = TRUE,
    quiet = TRUE
  )
}

if (nrow(cluster_poly_sf) > 0) {
  st_write(
    cluster_poly_sf,
    outfile_polygons,
    delete_dsn = TRUE,
    quiet = TRUE
  )
}

cat("\n=== EXPORT TERMINÉ ===\n")
cat("Points clusters hivernaux :", outfile_points, if (!ok_points) " (échec)" else "", "\n")
cat("Résumé clusters hivernaux :", outfile_summary, if (!ok_summary) " (échec)" else "", "\n")

```

```
cat("Centroides clusters      :", outfile_centroids, "\n")
cat("Polygones clusters      :", outfile_polygons, "\n")

## FIN DU SCRIPT
```

ANNEXE 5 : Script R d'extraction NDVI MODIS

```
# =====
# SCRIPT : NDVI ET TRANSITIONS ENTRE CLUSTERS HIVERNAUX
# =====
# Objectifs :
# Tester si les transitions intra-hivernales entre clusters
# se font vers des secteurs présentant un NDVI plus élevé.
# Décrire l'évolution du NDVI dans les zones de départ
# pendant leur occupation et dans les 48 jours avant le départ.
#
# Transition analysée :
# cluster A -> cluster B
#
# Métrique principale :
# delta_ndvi_transition = NDVI arrivée B - NDVI départ A
#
# Dynamique des zones de départ :
# delta_end_start = NDVI final - NDVI initial
# Calcul sur les mêmes pixels valides aux deux dates.
#
# Données :
# winter_clusters_summary.csv (échantillon sélectionné en amont).
# MODIS MOD13Q1 : NDVI, qualité des pixels et jour d'observation.
# Fenêtres autour des centroïdes, d'environ 50 x 50 km.
# Ces fenêtres ne correspondent pas aux contours des clusters.
#
# Sorties :
# Tableaux, contrôles de qualité, statistiques et figures
# dans le dossier NDVI_analyse.
# =====

suppressPackageStartupMessages({
  library(dplyr)
  library(ggplot2)
  library(MODISTools)
  library(geosphere)
})
# =====
# 1) PARAMÈTRES
# =====

base_dir <- "C:/Users/Utilisateur/Desktop/ARGOS FRANCE MAROC"
clusters_file <- file.path(base_dir, "winter_clusters_summary.csv")
out_dir <- file.path(base_dir, "NDVI_analyse")
cache_dir <- file.path(out_dir, "cache_MOD13Q1_v1")
product <- "MOD13Q1"
ndvi_km_lr <- 25L # Extension de part et d'autre : zone ~50 x 50 km
ndvi_km_ab <- 25L
recent_days <- 48L
min_coverage <- 0.50 # Contrôle de couverture, configurable; pas un seuil biologique
min_common_fraction <- 0.50
min_pixel_count <- 30L
min_pair_separation_days <- 8L
min_slope_dates <- 3L
min_slope_span_days <- 16L
max_observation_lag_days <- 8L # Différence des dates médianes A/B, même composite
max_composite_offset_days <- 16L
run_downloads <- TRUE
primary_comparison_mode <- "nearest_composite" # Choix fixe, avant examen des résultats
qa_scenarios <- list(good_only = 0L, good_and_marginal = c(0L, 1L))
# Analyse principale QA=0, sensibilité QA=0 ou 1. QA -1/2/3 exclus.
# Comparaison principale : nearest_composite (composites les plus proches
```

```

# du depart et de l'arrivee, avec controle des valeurs NDVI et de la qualite).
# Sensibilites : common_composite (meme periode MODIS A et B),
# strict_before_after (observations pendant occupation,
# A avant derniere localisation A; B apres premiere localisation B).
# Les bornes ARGOS encadrent la presence detectee, pas le depart/arrivee exacts.
bands <- c("250m_16_days_NDVI", "250m_16_days_pixel_reliability",
          "250m_16_days_composite_day_of_the_year")

# =====
# 2) FONCTIONS DE TRAITEMENT ET DE CONTRÔLE
# =====

validate_raw <- function(z) {
  need <- c("band", "calendar_date", "pixel", "value", "scale")
  if (!is.data.frame(z) || !nrow(z) || length(setdiff(need, names(z))))
    stop("Reponse MODIS vide ou schema inattendu.")
  z <- z[z$band %in% bands, , drop = FALSE]
  z$calendar_date <- as.Date(z$calendar_date)
  if (length(setdiff(bands, unique(z$band)))) stop("Une bande requise manque.")
  if ("complete" %in% names(z) && any(!is.na(z$complete) &
    tolower(as.character(z$complete)) %in% c("false", "0")))
    stop("Le service signale une extraction incomplete.")
  if (anyNA(z$calendar_date) || anyNA(z$pixel)) stop("Dates/pixels manquants.")
  k <- paste(z$band, z$calendar_date, z$pixel, sep = "|")
  if (anyDuplicated(k)) stop("Pixels dupliques : appariement ambigu.")
  z
}

fetch_chunk <- function(row, first_day, last_day) {
  key <- paste(product, gsub("[^A-Za-z0-9]", "_", row$cluster_id),
    sprintf("%.8f", row$lat), sprintf("%.8f", row$lon),
    row$km_lr, row$km_ab,
    first_day, last_day, sep = "_")
  path <- file.path(cache_dir, paste0(key, ".rds"))
  if (file.exists(path)) return(validate_raw(readRDS(path)))
  if (!run_downloads) stop("Cache absent : ", basename(path))
  last_error <- "Erreur inconnue"
  for (attempt in seq_len(3L)) {
    z <- tryCatch({
      ans <- MODISTools::mt_subset(product = product, band = bands,
        lat = row$lat, lon = row$lon, start = as.character(first_day),
        end = as.character(last_day), km_lr = row$km_lr,
        km_ab = row$km_ab, site_name = paste0("DEP_", row$cluster_id),
        internal = TRUE, progress = FALSE)
      validate_raw(ans)
    }, error = function(e) {
      if (grepl("No data points exist for the selected date range", conditionMessage(e), fixed = TRUE))
        return(data.frame())
      last_error <- conditionMessage(e); NULL
    })
    if (!is.null(z) && !nrow(z)) return(z) # Court fragment sans date MODIS
    if (!is.null(z)) { saveRDS(z, path); return(z) }
    if (attempt < 3L) Sys.sleep(2)
  }
  stop(last_error)
}

fetch_series <- function(row) {
  first_day <- row$query_start
  last_day <- row$query_end
  # Requetes courtes pour limiter la charge du service ORNL.
  starts <- seq(first_day, last_day, by = "48 days")
  parts <- lapply(seq_along(starts), function(i)

```

```

    fetch_chunk(row, starts[i], min(starts[i] + 47L, last_day)))
  raw <- bind_rows(parts)
  if (!nrow(raw)) stop("Aucune date MODIS dans la plage demandee.")
  for (nm in intersect(c("xllcorner", "yllcorner", "cellsize", "nrows", "ncols"),
    names(raw))) {
    if (n_distinct(raw[[nm]], na.rm = TRUE) > 1L)
      stop("Grille changeante entre requetes : ", nm)
  }
  # Un service peut inclure un composite frontalier dans deux requetes.
  if (anyDuplicated(paste(raw$band, raw$calendar_date, raw$pixel))) {
    conflicts <- raw %>% group_by(band, calendar_date, pixel) %>%
      summarise(n_values = n_distinct(value), .groups = "drop")
    if (any(conflicts$n_values > 1L)) stop("Valeurs contradictoires aux frontieres.")
    raw <- distinct(raw, band, calendar_date, pixel, .keep_all = TRUE)
  }
  n <- raw %>% filter(band == bands[1]) %>%
    transmute(calendar_date, pixel, ndvi = as.numeric(value) * as.numeric(scale))
  q <- raw %>% filter(band == bands[2]) %>%
    transmute(calendar_date, pixel, qa = as.integer(value))
  d <- raw %>% filter(band == bands[3]) %>%
    transmute(calendar_date, pixel, doy = as.integer(value))
  z <- left_join(left_join(n, q, by = c("calendar_date", "pixel")), d,
    by = c("calendar_date", "pixel"))
  year <- as.integer(format(z$calendar_date, "%Y"))
  jan1 <- as.Date(paste0(year, "-01-01"))
  next_jan1 <- as.Date(paste0(year + 1L, "-01-01"))
  valid_doy <- !is.na(z$doy) & z$doy >= 1L & z$doy <= as.integer(next_jan1 - jan1)
  z$observation_date <- jan1 + z$doy - 1L
  z$observation_date[!valid_doy] <- as.Date(NA)
  # Date DOY = date de l'observation retenue pour CHAQUE pixel.
  # La date du composite n'est pas assimilee a la date du depart.
  z$valid_ndvi <- is.finite(z$ndvi) & z$ndvi >= -0.2 & z$ndvi <= 1
  z$valid_observation <- !is.na(z$observation_date) &
    z$observation_date >= z$calendar_date &
    z$observation_date <= z$calendar_date + 15L
  z
}

analyse_window <- function(z, row, first_day, window_name, qa_keep) {
  total_pixels <- n_distinct(z$pixel)
  # Exclusion au pixel des observations hors fenetre ou apres le depart.
  v <- z %>% filter(valid_ndvi, valid_observation, qa %in% qa_keep,
    observation_date >= first_day,
    observation_date <= row$end_date_A)
  series <- v %>% group_by(calendar_date) %>% summarise(
    ndvi_mean = mean(ndvi), ndvi_median = median(ndvi),
    n_valid_pixels = n_distinct(pixel),
    observation_day_median = median(as.numeric(observation_date)),
    observation_date_min = min(observation_date),
    observation_date_max = max(observation_date), .groups = "drop") %>%
    mutate(coverage = n_valid_pixels / total_pixels,
      observation_date_median = as.Date(observation_day_median, origin = "1970-01-01"),
      day_from_departure = observation_day_median - as.numeric(row$end_date_A),
      usable_date = coverage >= min_coverage & n_valid_pixels >= min_pixel_count) %>%
    arrange(calendar_date)
  eligible <- series %>% filter(usable_date)
  summary <- tibble(window = window_name, window_start = first_day,
    n_dates_with_pixels = nrow(series), n_usable_dates = nrow(eligible),
    n_total_pixels = total_pixels, ndvi_first = NA_real_, ndvi_last = NA_real_,
    delta_end_start = NA_real_, separation_days = NA_real_,
    n_common_endpoint_pixels = 0L, common_endpoint_fraction = NA_real_,
    slope_ndvi_per_day = NA_real_, n_slope_dates = 0L,
    slope_span_days = NA_real_, n_common_slope_pixels = 0L,

```

```

pair_status = "fewer_than_two_usable_dates", slope_status = "insufficient_dates")
if (nrow(eligible) >= 2L) {
  first_date <- eligible$calendar_date[1]
  last_date <- tail(eligible$calendar_date, 1)
  pair <- inner_join(
    v %>% filter(calendar_date == first_date) %>%
      select(pixel, first_ndvi = ndvi, first_obs = observation_date),
    v %>% filter(calendar_date == last_date) %>%
      select(pixel, last_ndvi = ndvi, last_obs = observation_date), by = "pixel")
  summary$n_common_endpoint_pixels <- nrow(pair)
  summary$common_endpoint_fraction <- nrow(pair) / total_pixels
  summary$pair_status <- "insufficient_common_pixels"
  if (nrow(pair) >= min_pixel_count &&
      summary$common_endpoint_fraction >= min_common_fraction) {
    summary$separation_days <- median(as.numeric(pair$last_obs - pair$first_obs))
    summary$pair_status <- "dates_too_close"
    if (summary$separation_days >= min_pair_separation_days) {
      summary$ndvi_first <- mean(pair$first_ndvi)
      summary$ndvi_last <- mean(pair$last_ndvi)
      summary$delta_end_start <- mean(pair$last_ndvi - pair$first_ndvi)
      summary$pair_status <- "usable_descriptive_change"
    }
  }
}
# Pente sur un ENSEMBLE FIXE de pixels valides a toutes les dates retenues.
# Ne pas utiliser les milliers de pixels comme repetitions statistiques.
if (nrow(eligible) >= min_slope_dates) {
  w <- v %>% filter(calendar_date %in% eligible$calendar_date)
  common <- w %>% count(pixel, name = "n_dates") %>%
    filter(n_dates == nrow(eligible)) %>% pull(pixel)
  summary$n_common_slope_pixels <- length(common)
  summary$slope_status <- "insufficient_common_pixels"
  if (length(common) >= min_pixel_count &&
      length(common) / total_pixels >= min_common_fraction) {
    fixed <- w %>% filter(pixel %in% common) %>% group_by(calendar_date) %>%
      summarise(y = mean(ndvi), x = median(as.numeric(observation_date)),
        .groups = "drop") %>% arrange(x)
    summary$slope_span_days <- diff(range(fixed$x))
    summary$n_slope_dates <- nrow(fixed)
    summary$slope_status <- "span_too_short"
    if (summary$slope_span_days >= min_slope_span_days && n_distinct(fixed$x) >= 3L) {
      fixed$x <- fixed$x - min(fixed$x)
      summary$slope_ndvi_per_day <- unname(coef(lm(y ~ x, data = fixed)))[2]
      summary$slope_status <- "usable_descriptive_slope"
    }
  }
}
list(series = series, summary = summary)
}

period_stats <- function(z, qa_keep, first_day = as.Date("1900-01-01"),
  last_day = as.Date("2100-01-01")) {
  total <- n_distinct(z$pixel)
  v <- z %>% filter(valid_ndvi, valid_observation, qa %in% qa_keep,
    observation_date >= first_day, observation_date <= last_day)
  if (!nrow(v)) return(tibble(calendar_date = as.Date(character()),
    ndvi = double(), n_pixels = integer(), coverage = double(),
    observation_day_median = double(), observation_date_min = as.Date(character()),
    observation_date_max = as.Date(character()), usable = logical()))
  v %>% group_by(calendar_date) %>% summarise(ndvi = mean(ndvi),
    n_pixels = n_distinct(pixel), observation_day_median = median(as.numeric(observation_date)),
    observation_date_min = min(observation_date), observation_date_max = max(observation_date),

```

```

.groups = "drop") %>% mutate(coverage = n_pixels / total,
  usable = n_pixels >= min_pixel_count & coverage >= min_coverage) %>%
  arrange(calendar_date)
}

compare_areas <- function(a, b, row, qa_keep, mode) {
  res <- tibble(comparison_mode = mode, comparison_status = "no_usable_pair",
    ndvi_A_depart = NA_real_, ndvi_B_arrivee = NA_real_, delta_ndvi_transition = NA_real_,
    composite_date_A = as.Date(NA), composite_date_B = as.Date(NA),
    observation_date_A_median = as.Date(NA), observation_date_B_median = as.Date(NA),
    observation_date_A_min = as.Date(NA), observation_date_A_max = as.Date(NA),
    observation_date_B_min = as.Date(NA), observation_date_B_max = as.Date(NA),
    observation_lag_days = NA_real_, n_valid_pixels_A = NA_integer_,
    n_valid_pixels_B = NA_integer_, coverage_A = NA_real_, coverage_B = NA_real_)
  if (is.null(a) || is.null(b)) {
    res$comparison_status <- "download_failed"; return(res)
  }
  if (mode == "strict_before_after") {
    sa <- period_stats(a, qa_keep, max(row$start_date_A, row$end_date_A - 16L), row$end_date_A)
    sb <- period_stats(b, qa_keep, row$start_date_B, min(row$end_date_B, row$start_date_B + 16L))
  } else {
    sa <- period_stats(a, qa_keep); sb <- period_stats(b, qa_keep)
  }
  sa <- sa %>% filter(usable); sb <- sb %>% filter(usable)
  if (!nrow(sa) || !nrow(sb)) return(res)
  if (mode == "common_composite") {
    reference <- (as.numeric(row$end_date_A) + as.numeric(row$start_date_B)) / 2
    common <- intersect(as.character(sa$calendar_date), as.character(sb$calendar_date))
    common <- as.Date(common)
    # Date de reference = milieu de l'intervalle entre les fixes ARGOS A et B.
    # Choix du composite par son centre nominal; les dates reelles sont exportees.
    offsets <- abs(as.numeric(common) + 7.5 - reference)
    common <- common[offsets <= max_composite_offset_days]
    if (!length(common)) { res$comparison_status <- "no_common_period"; return(res) }
    date <- common[which.min(abs(as.numeric(common) + 7.5 - reference))]
    aa <- sa[match(date, sa$calendar_date), , drop = FALSE]
    bb <- sb[match(date, sb$calendar_date), , drop = FALSE]
  } else if (mode == "nearest_composite") {
    sa <- sa %>% filter(abs(as.numeric(calendar_date - row$end_date_A)) <= max_composite_offset_days)
    sb <- sb %>% filter(abs(as.numeric(calendar_date - row$start_date_B)) <= max_composite_offset_days)
    if (!nrow(sa) || !nrow(sb)) { res$comparison_status <- "no_nearby_period"; return(res) }
    aa <- sa[which.min(abs(as.numeric(sa$calendar_date - row$end_date_A))), , drop = FALSE]
    bb <- sb[which.min(abs(as.numeric(sb$calendar_date - row$start_date_B))), , drop = FALSE]
  } else if (mode == "strict_before_after") {
    aa <- sa[which.min(abs(sa$observation_day_median - as.numeric(row$end_date_A))), , drop = FALSE]
    bb <- sb[which.min(abs(sb$observation_day_median - as.numeric(row$start_date_B))), , drop = FALSE]
  } else stop("Mode inconnu : ", mode)
  res$ndvi_A_depart <- aa$ndvi; res$ndvi_B_arrivee <- bb$ndvi
  res$composite_date_A <- aa$calendar_date; res$composite_date_B <- bb$calendar_date
  res$observation_date_A_median <- as.Date(aa$observation_day_median, origin = "1970-01-01")
  res$observation_date_B_median <- as.Date(bb$observation_day_median, origin = "1970-01-01")
  res$observation_date_A_min <- aa$observation_date_min; res$observation_date_A_max <-
aa$observation_date_max
  res$observation_date_B_min <- bb$observation_date_min; res$observation_date_B_max <-
bb$observation_date_max
  res$observation_lag_days <- bb$observation_day_median - aa$observation_day_median
  res$n_valid_pixels_A <- aa$n_pixels; res$n_valid_pixels_B <- bb$n_pixels
  res$coverage_A <- aa$coverage; res$coverage_B <- bb$coverage
  if (mode == "common_composite" && abs(res$observation_lag_days) > max_observation_lag_days) {
    res$comparison_status <- "actual_dates_too_far_apart"; return(res)
  }
  res$delta_ndvi_transition <- res$ndvi_B_arrivee - res$ndvi_A_depart
  res$comparison_status <- "usable"
}

```

```

    res
  }

read_clusters <- function(path) {
  x <- read.csv(path, stringsAsFactors = FALSE, check.names = FALSE)
  req <- c("id", "winter_id", "Origine", "winter_cluster_id", "start_time", "end_time",
    "lon_centroid", "lat_centroid")
  if (length(setdiff(req, names(x)))) stop("Colonnes manquantes : ", paste(setdiff(req, names(x)), collapse = ",
"))
  for (nm in c("id", "winter_id", "Origine", "winter_cluster_id")) x[[nm]] <- trimws(as.character(x[[nm]]))
  x$Origine[tolower(x$Origine) == "france"] <- "France"
  x$Origine[tolower(x$Origine) == "maroc"] <- "Maroc"
  for (nm in c("start_time", "end_time")) x[[nm]] <- as.POSIXct(x[[nm]], tz = "UTC",
    tryFormats = c("%Y-%m-%d %H:%M:%S", "%Y-%m-%d %H:%M", "%d/%m/%Y %H:%M:%S",
"%d/%m/%Y %H:%M"))
  x$lon_centroid <- suppressWarnings(as.numeric(x$lon_centroid))
  x$lat_centroid <- suppressWarnings(as.numeric(x$lat_centroid))
  bad <- !complete.cases(x[, req]) | !is.finite(x$lon_centroid) | !is.finite(x$lat_centroid) |
    abs(x$lon_centroid) > 180 | abs(x$lat_centroid) > 90 | x$end_time < x$start_time
  bad <- bad | x$id == "" | x$winter_id == "" | x$winter_cluster_id == ""
  if (any(bad)) {
    write.csv(x[bad, ], file.path(out_dir, "invalid_clusters.csv"), row.names = FALSE)
    stop("Clusters invalides : voir invalid_clusters.csv. Corriger les entrees avant de continuer.")
  }
  if (anyDuplicated(x$winter_cluster_id)) stop("Identifiants de clusters dupliques.")
  x %>% mutate(start_date = as.Date(start_time, tz = "UTC"), end_date = as.Date(end_time, tz = "UTC"))
  %>%
    arrange(id, winter_id, start_time)
}

build_transitions <- function(clusters) {
  z <- clusters %>% group_by(id, winter_id) %>% arrange(start_time, .by_group = TRUE) %>% mutate(
    cluster_A = winter_cluster_id, cluster_B = lead(winter_cluster_id),
    start_time_A = start_time, end_time_A = end_time,
    start_time_B = lead(start_time), end_time_B = lead(end_time),
    start_date_A = start_date, end_date_A = end_date,
    start_date_B = lead(start_date), end_date_B = lead(end_date),
    lon_A = lon_centroid, lat_A = lat_centroid,
    lon_B = lead(lon_centroid), lat_B = lead(lat_centroid)) %>%
    ungroup() %>% filter(!is.na(cluster_B)) %>% mutate(
    transition_id = paste(id, winter_id, cluster_A, cluster_B, sep = "_"),
    transition_delay_days = as.numeric(difftime(start_time_B, end_time_A, units = "days"))) %>%
    select(transition_id, id, winter_id, Origine, cluster_A, cluster_B,
    start_time_A, end_time_A, start_time_B, end_time_B, start_date_A, end_date_A,
    start_date_B, end_date_B, lon_A, lat_A, lon_B, lat_B, transition_delay_days)
  if (any(z$transition_delay_days < 0)) stop("Clusters successifs se chevauchant dans le temps.")
  if (anyDuplicated(z$transition_id)) stop("Transitions dupliques.")
  if (!nrow(z)) stop("Aucune transition dans le fichier de clusters.")
  z$distance_km <- geosphere::distHaversine(cbind(z$lon_A, z$lat_A), cbind(z$lon_B, z$lat_B)) / 1000
  z
}

# =====
# 3) LECTURE ET PRÉPARATION DES CLUSTERS
# =====

dir.create(out_dir, recursive = TRUE, showWarnings = FALSE)
dir.create(cache_dir, recursive = TRUE, showWarnings = FALSE)
if (!file.exists(clusters_file)) stop("Fichier introuvable : ", clusters_file)
clusters <- read_clusters(clusters_file)
tr <- build_transitions(clusters)
write.csv(tr, file.path(out_dir, "transitions_input_audit.csv"), row.names = FALSE)
cat("Clusters :", nrow(clusters), "| transitions :", nrow(tr), "| oiseaux :", n_distinct(tr$id), "\n")

```

```

# =====
# 4) TÉLÉCHARGEMENT DES DONNÉES MODIS
# =====

requests <- bind_rows(
  tr %>% transmute(cluster_id = cluster_A, lat = lat_A, lon = lon_A,
    query_start = pmin(start_date_A, end_date_A - recent_days) - 15L,
    query_end = end_date_A + 16L),
  tr %>% transmute(cluster_id = cluster_B, lat = lat_B, lon = lon_B,
    query_start = start_date_B - 16L, query_end = start_date_B + 16L)) %>%
group_by(cluster_id, lat, lon) %>% summarise(query_start = min(query_start),
  query_end = max(query_end), .groups = "drop") %>%
mutate(km_lr = ndvi_km_lr, km_ab = ndvi_km_ab)
if (anyDuplicated(requests$cluster_id)) stop("Coordonnées incohérentes pour un même cluster.")
write.csv(requests, file.path(out_dir, "download_requests.csv"), row.names = FALSE)
zone_data <- list(); failures <- list()
for (i in seq_len(nrow(requests))) {
  row <- requests[i, , drop = FALSE]
  message("MODIS : zone ", i, "/", nrow(requests), " ", row$cluster_id)
  z <- tryCatch(fetch_series(row), error = function(e) {
    failures[[length(failures) + 1L]] <- tibble(cluster_id = row$cluster_id,
      error = conditionMessage(e)); NULL
  })
  if (!is.null(z)) zone_data[[row$cluster_id]] <- z
}
error_table <- if (length(failures)) bind_rows(failures) else tibble(cluster_id = character(), error = character())
write.csv(error_table, file.path(out_dir, "download_errors.csv"), row.names = FALSE)

# =====
# 5) COMPARAISON ENTRE CLUSTERS ET DYNAMIQUE DES ZONES DE DÉPART
# =====

comparisons <- list(); dynamics <- list(); trajectories <- list(); pixel_audits <- list()
for (id_zone in names(zone_data)) {
  z <- zone_data[[id_zone]]
  pixel_audits[[length(pixel_audits) + 1L]] <- z %>% group_by(calendar_date) %>% summarise(
    n_total = n(), n_valid_range = sum(valid_ndvi),
    n_good = sum(valid_ndvi & valid_observation & qa == 0L, na.rm = TRUE),
    n_good_or_marginal = sum(valid_ndvi & valid_observation & qa %in% c(0L, 1L)),
    .groups = "drop") %>% mutate(cluster_id = id_zone)
}
for (i in seq_len(nrow(tr))) {
  row <- tr[i, , drop = FALSE]
  a <- zone_data[[row$cluster_A]]; b <- zone_data[[row$cluster_B]]
  for (scenario in names(qa_scenarios)) {
    for (mode in c("common_composite", "nearest_composite", "strict_before_after")) {
      cc <- compare_areas(a, b, row, qa_scenarios[[scenario]], mode)
      comparisons[[length(comparisons) + 1L]] <- bind_cols(row, cc) %>%
        mutate(qa_scenario = scenario, km_lr = ndvi_km_lr, km_ab = ndvi_km_ab)
    }
    if (is.null(a)) next
    windows <- list(occupation = row$start_date_A)
    windows[[paste0("last_", recent_days, "_days")]] <- row$end_date_A - recent_days
    for (window in names(windows)) {
      aw <- analyse_window(a, row, windows[[window]], window, qa_scenarios[[scenario]])
      meta <- row %>% select(transition_id, id, winter_id, Origine, cluster_A, start_date_A, end_date_A)
      meta$qa_scenario <- scenario
      dynamics[[length(dynamics) + 1L]] <- bind_cols(meta, aw$summary)
      if (nrow(aw$series)) trajectories[[length(trajectories) + 1L]] <-
        bind_cols(meta[rep(1L, nrow(aw$series)), , drop = FALSE], aw$series) %>% mutate(window = window)
    }
  }
}

```

```

}
comp <- bind_rows(comparisons)
comp$ndvi_transition_positive <- ifelse(is.na(comp$delta_ndvi_transition), NA, comp$delta_ndvi_transition >
0)
dyn <- bind_rows(dynamics); ts <- bind_rows(trjectories)
write.csv(comp, file.path(out_dir, "winter_cluster_transitions_ndvi_all_modes.csv"), row.names = FALSE)
main <- comp %>% filter(qa_scenario == "good_only", comparison_mode ==
.env$primary_comparison_mode)
write.csv(main, file.path(out_dir, "winter_cluster_transitions_ndvi_primary.csv"), row.names = FALSE)
write.csv(dyn, file.path(out_dir, "departure_dynamics_summary.csv"), row.names = FALSE)
write.csv(ts, file.path(out_dir, "departure_ndvi_timeseries.csv"), row.names = FALSE)
write.csv(bind_rows(pixel_audits), file.path(out_dir, "pixel_quality_audit.csv"), row.names = FALSE)

# =====
# 6) RÉSULTATS STATISTIQUES ET ANALYSES DE SENSIBILITÉ
# =====

valid <- comp %>% filter(comparison_status == "usable")
if (!nrow(valid)) stop("Aucune comparaison exploitable : examiner les audits et download_errors.csv.")
summary_table <- valid %>% group_by(qa_scenario, comparison_mode) %>% summarise(
  n_transitions = n(), n_birds = n_distinct(id),
  mean_ndvi_A = mean(ndvi_A_depart), mean_ndvi_B = mean(ndvi_B_arrivee),
  mean_delta = mean(delta_ndvi_transition), sd_delta = sd(delta_ndvi_transition),
  median_delta = median(delta_ndvi_transition), n_positive = sum(delta_ndvi_transition > 0),
  n_zero = sum(delta_ndvi_transition == 0), proportion_positive = mean(delta_ndvi_transition > 0),
  .groups = "drop")
write.csv(summary_table, file.path(out_dir, "transition_summary.csv"), row.names = FALSE)
by_origin <- valid %>% group_by(qa_scenario, comparison_mode, Origine) %>% summarise(
  n_transitions = n(), n_birds = n_distinct(id), mean_delta = mean(delta_ndvi_transition),
  .groups = "drop")
write.csv(by_origin, file.path(out_dir, "transition_summary_by_origin.csv"), row.names = FALSE)
# Origine et annee ne sont pas separees par ce plan d'echantillonnage : pas
# de test causal/populationnel France-Maroc dans ce script.
bird <- valid %>% group_by(qa_scenario, comparison_mode, id) %>% summarise(
  n_transitions = n(), mean_delta = mean(delta_ndvi_transition), .groups = "drop")
write.csv(bird, file.path(out_dir, "transition_delta_by_bird.csv"), row.names = FALSE)

sink(file.path(out_dir, "statistics_and_diagnostics.txt"))
cat("ANALYSE PRINCIPALE :", primary_comparison_mode, "| QA=0. Les autres modes sont des
sensibilites.\n")
print(comp %>% count(qa_scenario, comparison_mode, comparison_status))
print(summary_table)
for (scenario in names(qa_scenarios)) for (mode in unique(comp$comparison_mode)) {
  cat("\nQA :", scenario, "| mode :", mode, "\n")
  x <- valid %>% filter(qa_scenario == .env$scenario, comparison_mode == .env$mode) %>%
pull(delta_ndvi_transition)
  y <- bird %>% filter(qa_scenario == .env$scenario, comparison_mode == .env$mode) %>%
pull(mean_delta)
  cat("Tests par TRANSITION, non ajustes pour les repetitions : exploratoires seulement.\n")
  if (length(x) >= 2L && any(x != 0)) print(wilcox.test(x, mu = 0, alternative = "two.sided", exact = FALSE))
  xnz <- x[x != 0]
  if (length(xnz)) {
    cat("Binomial unilateral directionnel (greater), zeros exclus :\n")
    print(binom.test(sum(xnz > 0), length(xnz), .5, alternative = "greater"))
    cat("Binomial bilateral, zeros exclus :\n")
    print(binom.test(sum(xnz > 0), length(xnz), .5, alternative = "two.sided"))
  }
  cat("Sensibilite : une moyenne de delta par OISEAU :\n")
  if (length(y) >= 5L && any(y != 0)) print(wilcox.test(y, mu = 0, alternative = "two.sided", exact = FALSE))
}
if (nrow(dyn)) {
  cat("\nDynamique : disponibilite de l'information temporelle\n")
  print(dyn %>% count(qa_scenario, window, pair_status, slope_status))
}

```

```

}
cat("\nPARAMETRES\n")
print(list(product = product, primary_comparison_mode = primary_comparison_mode, ndvi_km_lr =
ndvi_km_lr, ndvi_km_ab = ndvi_km_ab,
min_coverage = min_coverage, min_common_fraction = min_common_fraction,
min_pixel_count = min_pixel_count, min_slope_dates = min_slope_dates,
min_slope_span_days = min_slope_span_days, min_pair_separation_days = min_pair_separation_days,
max_observation_lag_days = max_observation_lag_days, recent_days = recent_days))
print(sessionInfo()); sink()

# =====
# 7) FIGURES
# =====

for (scenario in names(qa_scenarios)) for (mode in unique(comp$comparison_mode)) {
  vv <- valid %>% filter(qa_scenario == .env$scenario, comparison_mode == .env$mode)
  if (!nrow(vv)) next
  long <- bind_rows(vv %>% transmute(transition_id, id, delta_ndvi_transition,
area = "Departure", NDVI = ndvi_A_depart),
vv %>% transmute(transition_id, id, delta_ndvi_transition,
area = "Arrival", NDVI = ndvi_B_arrivee))
  long$area <- factor(long$area, levels = c("Departure", "Arrival"))
  long$direction <- ifelse(long$delta_ndvi_transition > 0, "Higher at arrival", "Equal or lower at arrival")
  p <- ggplot(long, aes(area, NDVI, group = transition_id, colour = direction)) +
geom_line(alpha = .6, linewidth = .5) + geom_point(size = 2) +
scale_colour_manual(values = c("Higher at arrival" = "#28786B", "Equal or lower at arrival" = "#BA8950"))
+
labs(x = NULL, colour = NULL, title = "NDVI across successive wintering areas",
subtitle = paste(nrow(vv), "transitions;", n_distinct(vv$id), "birds;", mode, scenario)) +
theme_minimal(base_size = 12) + theme(legend.position = "bottom")
ggsave(file.path(out_dir, paste0("paired_ndvi_", mode, "_", scenario, ".png")), p, width = 7, height = 6, dpi =
300)
}
if (nrow(ts)) for (current_window in unique(ts$window)) {
  pp <- ts %>% filter(qa_scenario == "good_only", .data$window == .env$current_window)
  if (!nrow(pp)) next
  p <- ggplot(pp, aes(day_from_departure, ndvi_mean, group = cluster_A)) +
geom_vline(xintercept = 0, linetype = "dashed", colour = "#BA8950") +
geom_line(data = pp %>% filter(usable_date) %>% group_by(cluster_A) %>%
filter(n() >= 2L) %>% ungroup(), colour = "#28786B", linewidth = .5) +
geom_point(aes(shape = usable_date), colour = "#173A3C", size = 1.8) +
facet_wrap(~ cluster_A, scales = "free_x", ncol = 4) +
scale_shape_manual(values = c(`FALSE` = 1, `TRUE` = 16)) +
labs(x = "Days relative to last ARGOS fix in departure area", y = "Mean NDVI",
shape = "Coverage sufficient", title = paste("Departure-area NDVI:", current_window),
subtitle = "Lines connect dates meeting coverage criteria; open points show excluded dates") +
theme_minimal(base_size = 10) + theme(legend.position = "bottom")
ggsave(file.path(out_dir, paste0("departure_ndvi_", current_window, ".png")), p,
width = 13, height = max(5, ceiling(n_distinct(pp$cluster_A)/4) * 2.4), dpi = 300)
}
cat("\nTermine :", out_dir, "\nZones en echec de telechargement :", nrow(error_table), "\n")

## FIN DU SCRIPT

```