

Risque d'Invasion Structuré Spatialement dans les Communautés de Poissons

Intégration de HMSC et de Réseaux de Neurones sur Graphe

Samudranil Basak¹ Gerard Masson² Hoai Minh Le¹ Baba Issa Camara¹

¹Université de Lorraine, LCOMS
3 Rue Augustin Fresnel, 57070 Metz, France

²Université de Lorraine, CNRS, LIEC
Avenue du Général Delestraint, 57070 Metz, France

2-3 février 2026 – CiSStats-RESSTE

Contexte : Écosystèmes d'eau douce sous pression

Les écosystèmes d'eau douce figurent parmi les écosystèmes les plus menacés au monde.

Principaux facteurs de stress :

- Modification des habitats
- Altérations hydrologiques
- Changement climatique
- Pollution chimique
- **Invasions biologiques**

Les invasions biologiques jouent un rôle particulièrement important dans la restructuration des communautés, l'altération des interactions entre espèces et la modification du fonctionnement des écosystèmes.

Problématique : Le Gobie à taches noires

Espèce modèle : *Neogobius melanostomus*

Caractéristiques :

- Originaire région Ponto-Caspienne
- Expansion rapide en Europe et Amérique du Nord
- Habitat benthique
- Comportement agressif
- Reproduction élevée
- Large tolérance environnementale



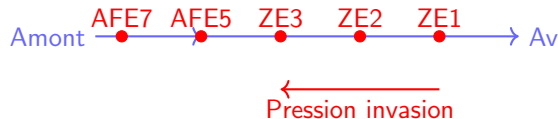
Impacts documentés : Compétition ressources, prédation œufs/juveniles natives, restructuration réseaux trophiques.

Spécificité des rivières : Structure directionnelle et connectivité élevée

Dispersion asymétrique :

- Transport passif aval dominant
- Dispersion par crues épisodique
- Corridors de navigation
- Secteurs aval : exposition précoce + longue
- Secteurs amont : refuges temporaires

Conséquence : Pression d'invasion cumulative plus élevée en aval



Hypothèse : Gradient aval → amont

Objectifs et Hypothèses

Système d'étude : Secteurs aval de la Moselle (ZE et AFE)

Objectif : Quantifier propagation spatiale du Gobie et conséquences écologiques

H1 – Dynamiques spatiales structurées

- Invasion anisotrope : secteurs aval exposés plus tôt et intensément
- Secteurs amont : refuges temporaires

H2 – Vulnérabilité médiée par traits et phylogénie

- Vulnérabilité différentielle selon position écologique (benthique/semi-benthique/non-benthique)
- Traits fonctionnels et parenté phylogénétique → patterns cohérents

Acquisition des Données

Zone : Secteurs aval Moselle (ZE et AFE) – corridor majeur d'invasion

Données collectées :

- **Communautés** : 19 espèces natives
- **GTN** : Présence/densité
- **Environnement (3 blocs)** :
 - **Bloc A** : Station/secteur (profondeur, pente, navigation, substrat)
 - **Bloc B** : Variables semi-fixes (débit, berge)
 - **Bloc C** : Variables dynamiques (végétation, substrat)
- **Traits** : 12 traits fonctionnels
- **Phylogénie** : Arbre 19 espèces

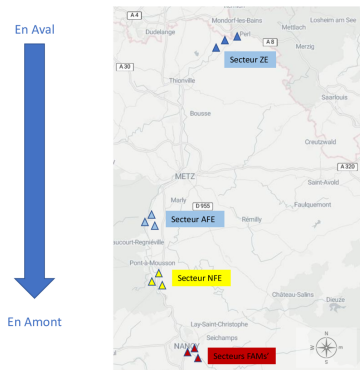


Figure – Enter Caption

Objectif : Quantifier similarité inter-stations pour graphe spatial

1. Distance de Gower pondérée par bloc

Pour chaque bloc (A, B, C) :

- **Variables continues** : PCA → poids basés sur variance expliquée
- **Variables catégorielles** : MCA → poids basés sur contribution inertie

Distance Gower pondérée :

$$D_{ij}^{(\text{bloc})} = \frac{\sum_{k \in \text{cont}} w_k \frac{|x_{ik} - x_{jk}|}{R_k} + \sum_{l \in \text{cat}} w_l \cdot \mathbb{I}[x_{il} \neq x_{jl}]}{\sum_k w_k + \sum_l w_l}$$

Similarité : $S_{ij}^{(\text{bloc})} = 1 - D_{ij}^{(\text{bloc})}$

2. Pondération blocs selon variabilité

$$\text{weight}_A = \frac{\text{Var}(D^{(A)})}{\text{Var}(D^{(A)}) + \text{Var}(D^{(B)})}$$

3. Similarité totale : $S_{ij}^{\text{total}} = \text{weight}_A \cdot S_{ij}^{(A)} + \text{weight}_B \cdot S_{ij}^{(B)}$

4. Indices finaux : within_sector_sim, between_sector_sim

Modèle 1 : HMSC

Objectif : Quantifier réponses espèce-environnement, traits, phylogénie

Formulation :

$$Y_{ij} \sim \text{Distribution}(\eta_{ij})$$

$$\eta_{ij} = X_i \beta_j + X_i (T_j \Gamma) + u_i + \epsilon_{ij}$$

$$\beta_j \sim \mathcal{N}(0, \sigma_\beta^2 C) \quad (\text{phylogénie})$$

$$\Gamma \sim \mathcal{N}(0, \sigma_\Gamma^2 I)$$

Composantes :

- β_j : réponses environnementales
- Γ : interactions traits-env.
- u_i : effet station, Ω : corrélations résiduelles

Estimation Bayésienne :

- MCMC (Gibbs/Metropolis-Hastings)
- 15,000 échantillons, thin=10, transient=2,000
- 4 chaînes indépendantes
- Convergence : trace plots, ESS

Covariates : héliophytes, hydrophytes, substrat, arbres, **GTN**, temps_depuis_arrivée, between_sector_sim

Objectif : Estimer probabilité propagation GTN station \rightarrow station

Architecture GCN (2 couches) :

$$h^{(1)} = \text{ReLU}(\text{GCN}(X, E))$$

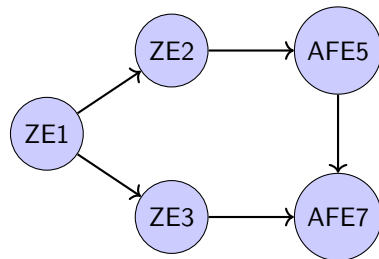
$$h^{(2)} = \text{ReLU}(\text{GCN}(h^{(1)}, E))$$

$$\hat{y} = \text{Linear}(h^{(2)})$$

- X : env. + traits GTN (BT, LMAX, RESIDENT, EUR)
- E : adjacence basée sur S_{ij}^{total} (seuil)
- \hat{y} : probabilité propagation

Paramètres : $d_{\text{in}} \times 32 + 32 \times 16 + 16 \times 1$

Entraînement : Adam, lr=0.01, 10K époques, BCE, perte ≈ 0.6



Graphe dirigé (flux rivière)

Formulation station $i \rightarrow j$:

$$\hat{p}_{i \rightarrow j} = f_{\text{distance}} \cdot f_{\text{traits}}$$

- f_{distance} : connectivité spatiale/hydrologique

$$f_{\text{distance}} = \frac{1}{1 + d_{ij}} \cdot \text{poids_connectivité}_{ij}$$

- f_{traits} : traits GTN (LMAX=2, RESIDENT, benthique, EUR)

Probabilité cumulative station j :

$$\hat{p}_{s_j} = 1 - \prod_{i \in \text{aval}(j)} (1 - \hat{p}_{i \rightarrow j})$$

⇒ Contributions cumulatives toutes stations aval (constant années/campagnes)

Intégration HMSC + GNN : Indices Risque

1. Score Vulnérabilité espèce i (HMSC) :

$$V_i = \sum_e \beta_{i,e} + \sum_{t \in T_i} \sum_e \gamma_{t,e}$$

$\beta_{i,e}$: réponse espèce à env. e , $\gamma_{t,e}$: effet trait t sur env. e , $V_i < 0$: vulnérable

2. Temps depuis Arrivée GTN :

$$T(s, y) = \max(0, y - A_s)$$

où A_s = année première observation GTN station s

3. Indice Risque (station-espèce-année) :

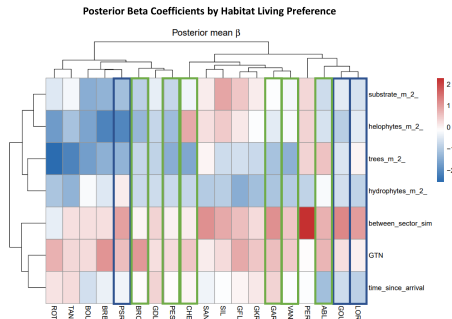
$$R_{i,s,y} = \hat{p}_s \times V_i \times (1 + T(s, y))$$

- \hat{p}_s : probabilité propagation GNN (constant)
- V_i : vulnérabilité HMSC
- $T(s, y)$: exposition temporelle croissante

$R_{i,s,y}$ plus négatif \Rightarrow risque écologique élevé

Résultat 1 : Réponses Espèce-Environnement (HMSC)

Hétérogénéité prononcée des réponses β :



Observations :

- Profils réponse distincts entre espèces
- Covariates env. : effets positifs/négatifs cohérents par groupes
- Filtrage environnemental + idiosyncrasies spécifiques
- Niches écologiques différenciées

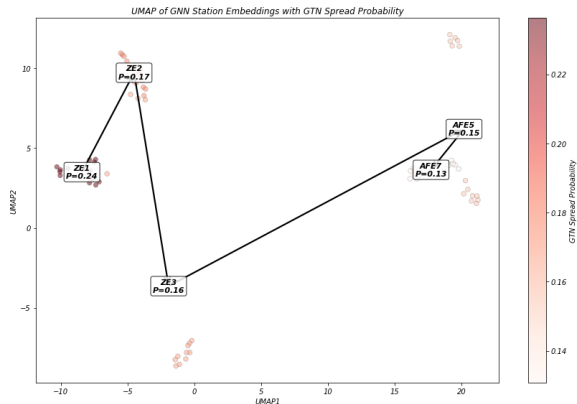
Interactions Traits-Env. (Γ) :

- Modulation cohérente selon traits
- Groupes traits amplifient/atténuent sensibilités
- Traits fonctionnels \rightarrow rôle clé au-delà taxonomie

Fig. 1 : Coefficients β par Habitat Living

Résultat 2 : Structure Latente GNN

Espace latent UMAP : structure organisée



Observations :

- Assemblages : régions distinctes (non homogène)
- GNN capture dépendances non-linéaires
- Gradient probabilité propagation aval → amont
- Zones haute probabilité : clusters localisés (combinaisons écologiques spécifiques)
- Zones basse probabilité : assemblages stables

Confirme H1 : Dispersion directionnelle, discrimination efficace assemblages à risque

Fig. 2 : UMAP embeddings + probabilité propagation GTN

Résultat 3 : Indices Risque et Probabilité Déclin

Transformation logistique risque \rightarrow probabilité :

$$P_{\text{déclin}}(i, s, y) = \text{logit}^{-1}(-\kappa(R_{i,s,y} - \tau))$$

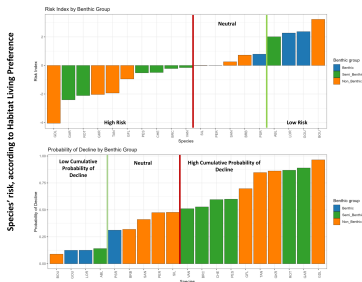


Fig. 3 : Indices risque + probabilité déclin

Patterns observés :

- Variation substantielle espèces/temps/secteurs
- Certaines espèces : risque élevé constant
- D'autres : fluctuations temporelles marquées
- Hétérogénéité spatiale : hotspots récurrents
- Catégories extrêmes : robustes entre métriques
- Espèces neutres : zone transitionnelle

Paramètres : $\kappa = 2$, $\tau = \text{médiane}(R)$

Propriétés : R très négatif $\Rightarrow P > 0.6$

Résultat 4 : Groupes à Risque

Agrégation par position écologique et traits trophiques :

Habitat Living (position écologique) :

Risque	Groupe	Espèces
Élevé	Semi-benthique	GAR, ROT, PES, CHE, BRC
	Non-benthique	GFL, TAN, GKR, GDL
Neutre	Non-benthique	SIL, PER, SAN, BRB
	Benthique	PSR
Faible	Semi-benthique	ABL
	Benthique	GOU, LOR
	Non-benthique	BOU

Habitat Trophic (traits trophiques) :

Risque	Groupe	Espèces
Élevé	Benthique	GDL, GKR, TAN, GFL, BRC
	Non-benthique	VAN, CHE, ROT, PES, GAR
Neutre	Benthique	BRB, SIL
	Non-benthique	PSR, SAN, PER
Faible	Benthique	GOU, LOR
	Non-benthique	BOU, ABL

Observations : Contrastes nets benthique/semi-benthique/non-benthique. Profils risque : variables selon traits, espèces benthiques trophiques plus vulnérables.

Discussion : Synthèse

Validation hypothèses :

H1 – Dynamiques spatiales structurées

Validée :

- GNN : gradient aval → amont clair
- Aval : exposition précoce/intense (transport passif)
- Amont : refuges temporaires (dispersion ralentie)
- Structure anisotrope : impacts non uniformes

H2 – Vulnérabilité traits/phylogénie

Validée :

- HMSC : effets phylogénétiques significatifs
- Traits (benthique, guildes, rhéophilie) prédictifs
- Espèces phylogénétiquement proches : réponses similaires
- Organisation multi-couches : gradients env. + traits

Synthèse :

- ① **Cadre intégré GNN + HMSC** : Propagation spatiale + vulnérabilité espèces
- ② **Quantification explicite** : Probabilité propagation, indices risque, probabilité déclin
- ③ **Validation hypothèses** : Structuration spatiale + médiation traits/phylogénie
- ④ **Applicabilité gestion** : Priorisation, surveillance, anticipation

Merci de votre attention !