Qui mangera qui en mer de Barents ?

Marine Food Web Prediction using Graph Machine Learning

Sébastien Ploix, Laurene Pecuchet

Plan

Introduction

I. How is the model working ?
II. How well is the model working ?
    - Comparison with the mechanistic niche model
III. How to predict interactions of unknown species ?
    - Phylogenetic Transfer Learning

Conclusion

Introduction

Context

- **Climate change :**
  - Change of species distribution
- **Food webs are hard to measure empirically**
  - a lot of existing interactions are not observed

-> Need for Food Web Generation

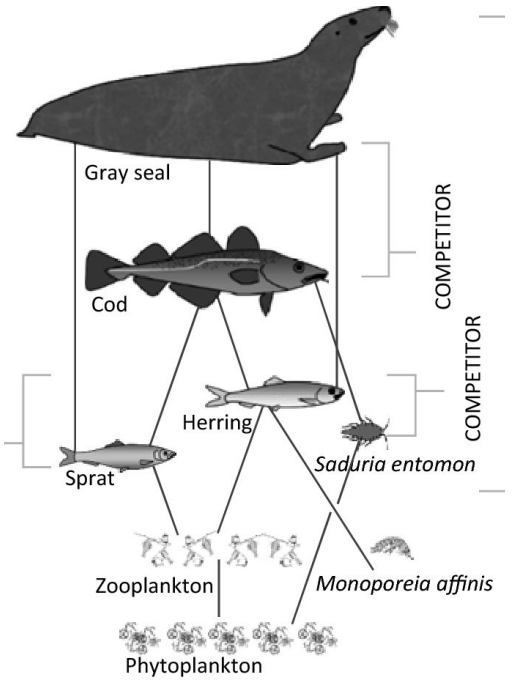
Understand a marine food web

Structured by :

- the Body Weight
- Position in the Water Column

Pecuchet et al., « Novel Feeding Interactions Amplify the Impact of Species Redistribution on an Arctic Food Web ».

Simplified food web of the baltic sea



Karlson, A.M.L., Gorokhova, E., Gårdmark, A. *et al.* Linking consumer physiological status to food-web structure and prey food value in the Baltic Sea. *Ambio* **49**, 391–406 (2020). https://doi.org/10.1007/s13280-019-01201-1

The data

Species traits :

- Body Weight *(numeric)*
- Position in the water column *(numeric)*
- Taxonomy
  - *Converted to a vector using a sentence transformer*

Food webs

Barents Sea

Kongsfjorden

Coral Reefs

Weddell Sea

1168 Species, 23171 documented interactions

Current approaches

Current approaches :

- mechanistic
- trait-based
- network-based

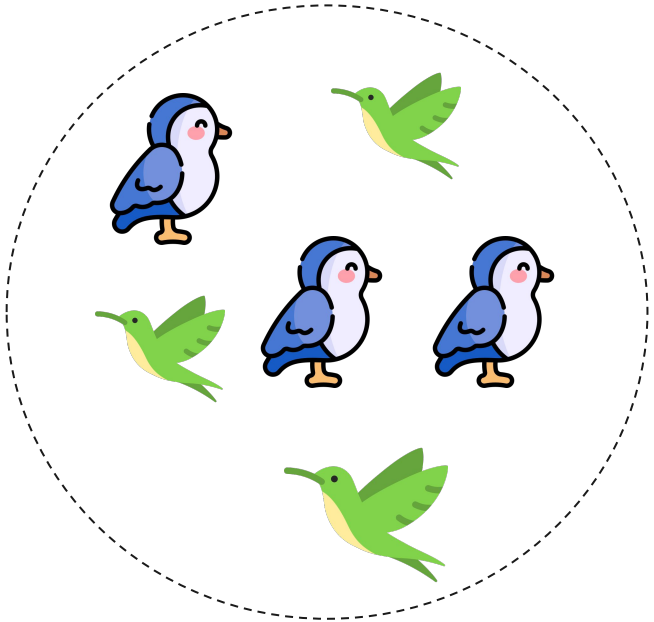
-> Merge traits and network based approaches

- increase precision

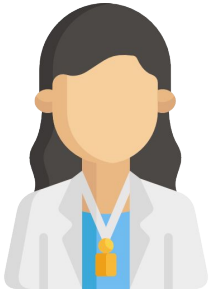
- reduce amount of traits needed

I. How is the model working ?

Analogy : Identifying birds

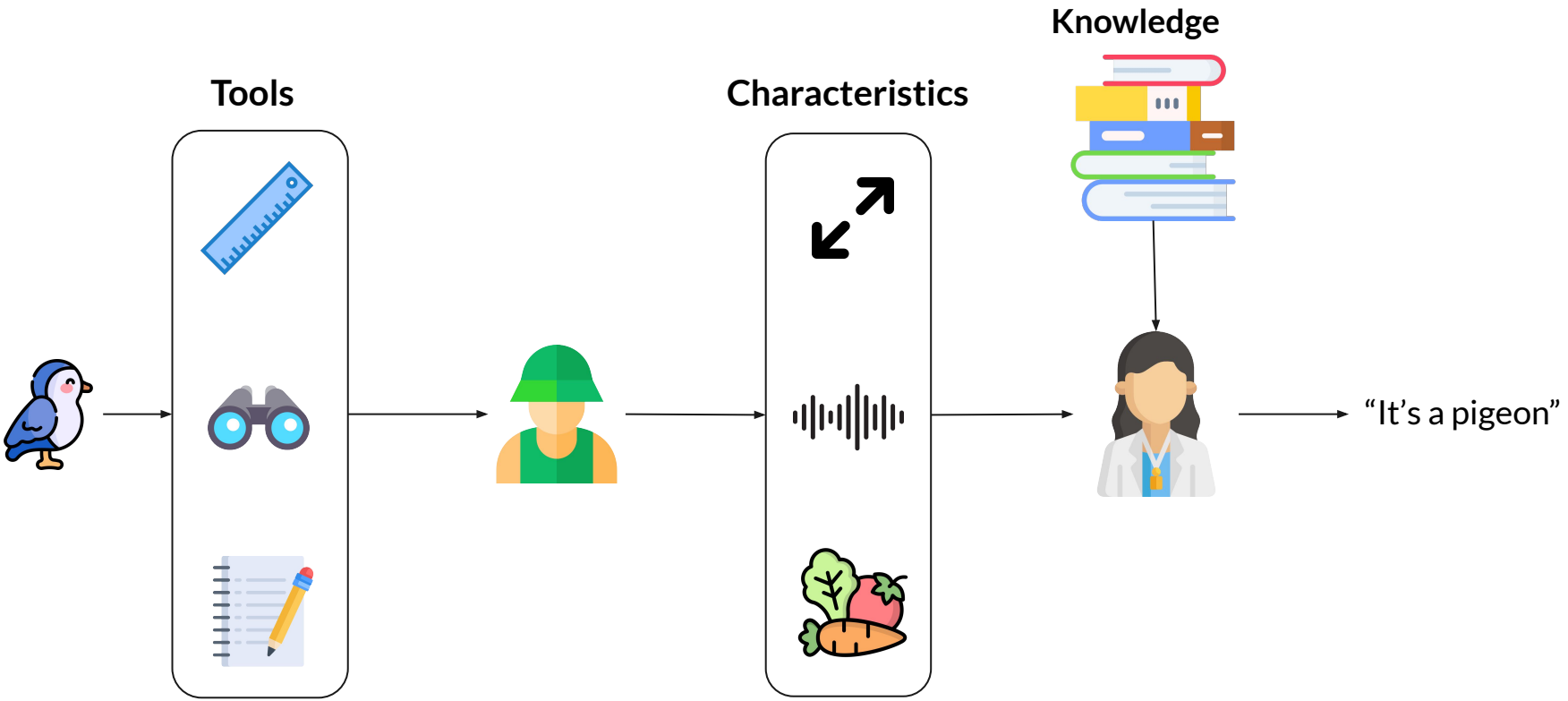


Technician



Expert

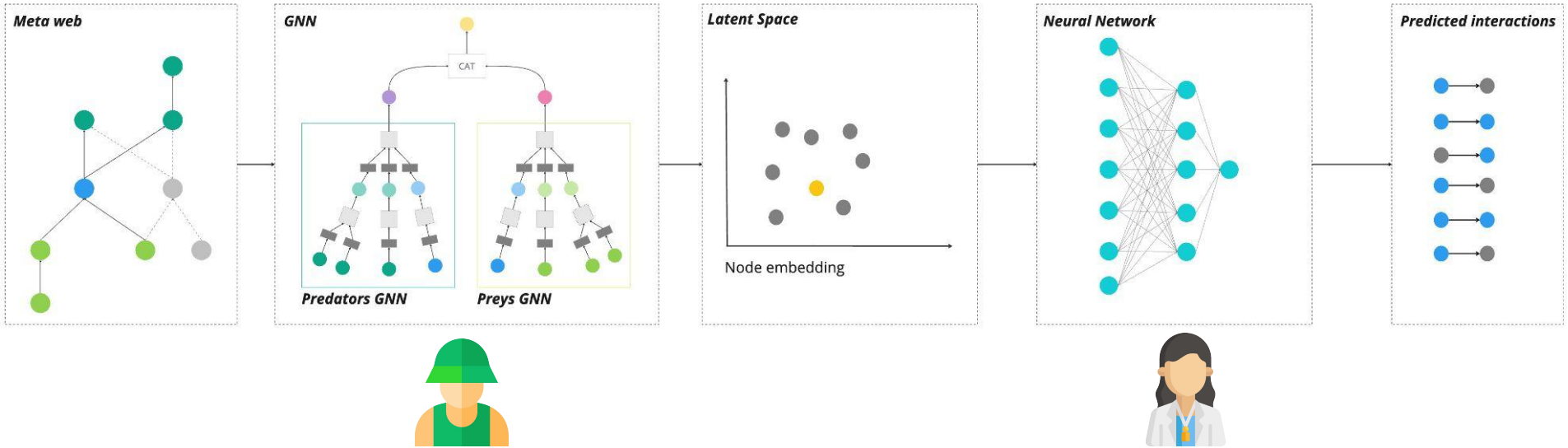
Analogy : Identifying birds



The model architecture - Graph Auto-Encoder (GAE)

Encoder
Extract relevant features

Decoder
Predict interactions



II. How well is my model working ?

Comparison with a reference model

Reference Model - Mechanistic Niche Model

Marine food webs are structured by the **body weight**

“Big animals eat smaller animals”

Node/Edge based verification

Log(BodyWeight)

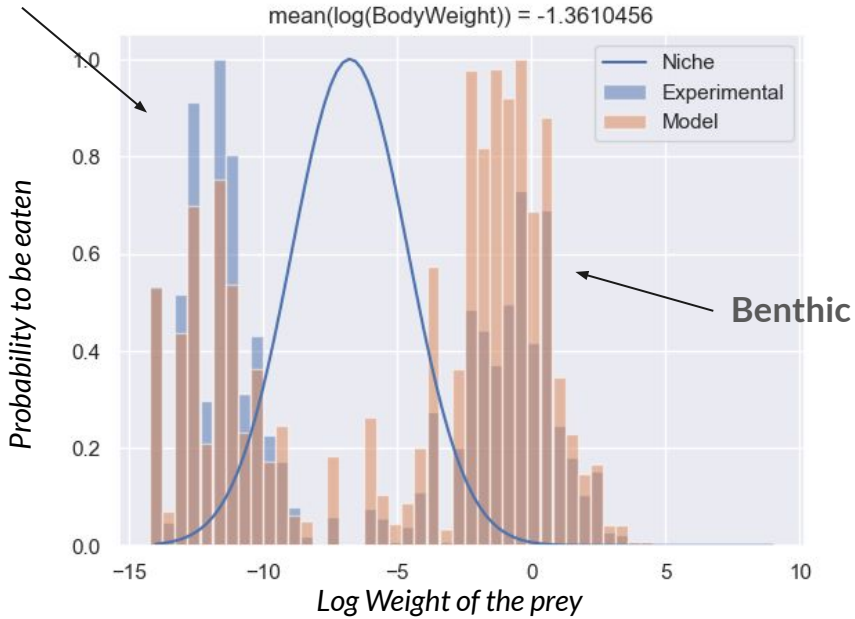
$$\mathbb{P}_\theta(u \to v) = \exp\left(-\frac{1}{2} * \frac{(w_u - \mu_\theta(w_v))^2}{\sigma_\theta(w_v)^2}\right)$$

Prey Predator

Comparison with the niche model

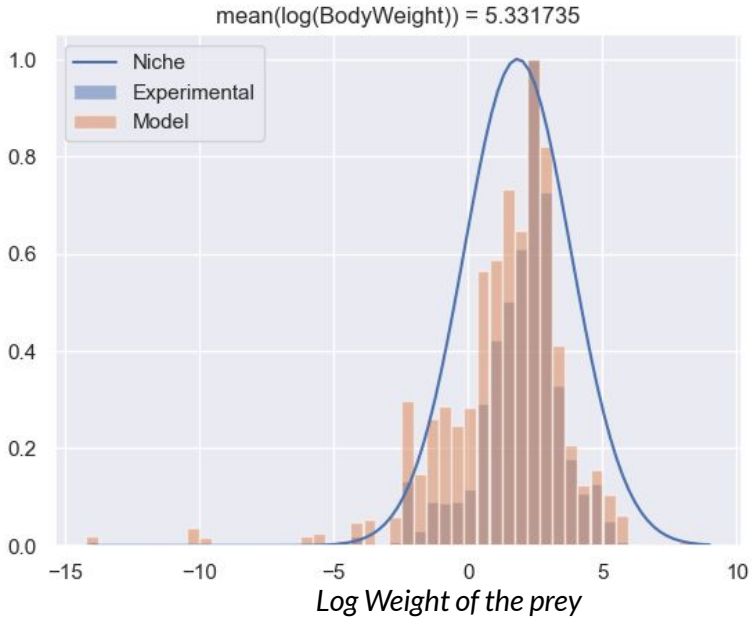
Small Predators Niche

Pelagic



Benthic

Big Predators Niche



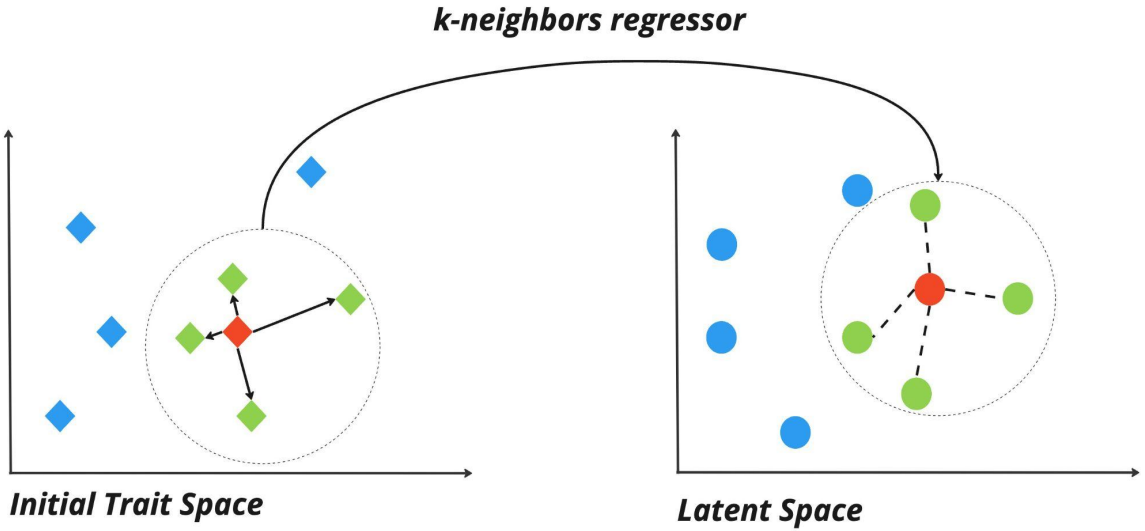
III. How to predict interactions of unknown species ?
Phylogenetic Transfer Learning

Phylogenetic Transfer Learning

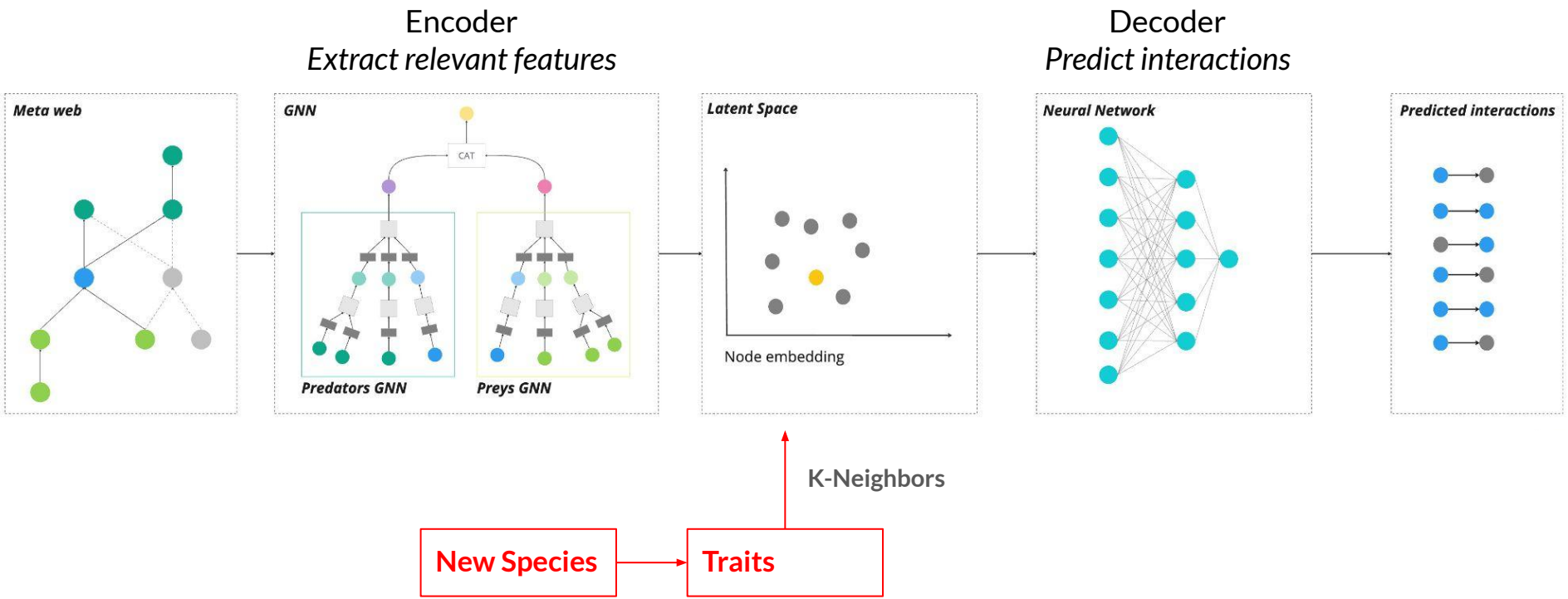
New species without any information about it's feeding behavior

Strong hypothesis :

Physically and taxonomically close species have similar feeding behaviors

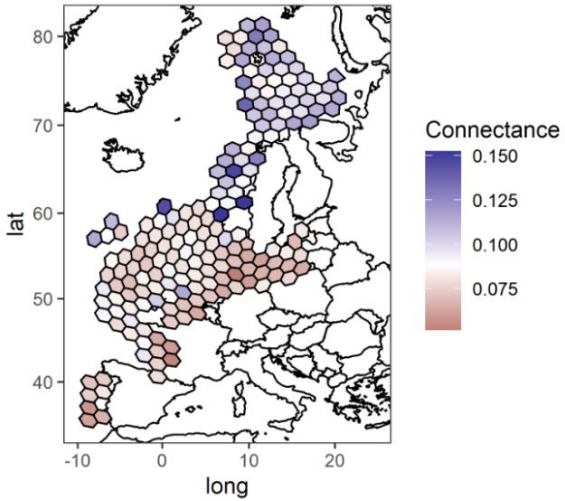


The model architecture - Graph Auto-Encoder (GAE)

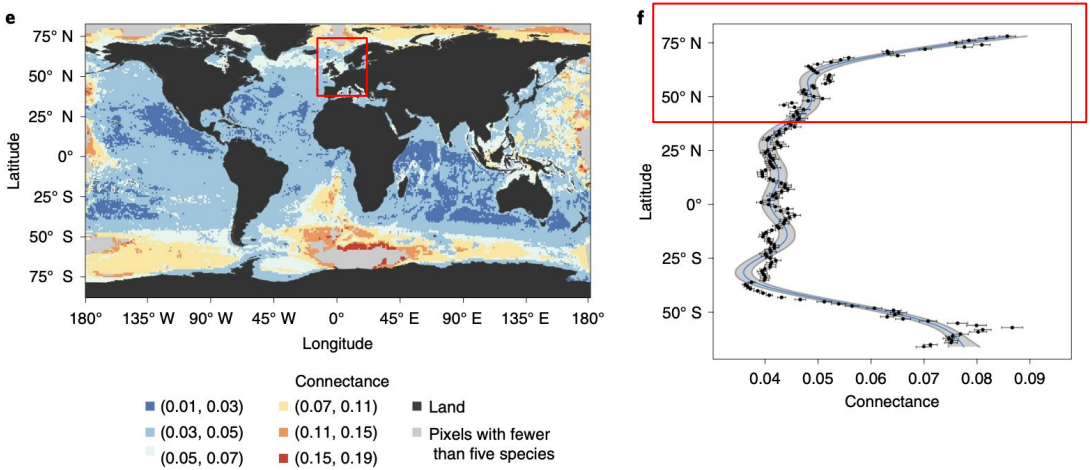


Spatial Patterns - *Network structure*

Our Paper



Albouy et al., « The Marine Fish Food Web Is Globally Connected ».



Similar ecological patterns

Conclusion

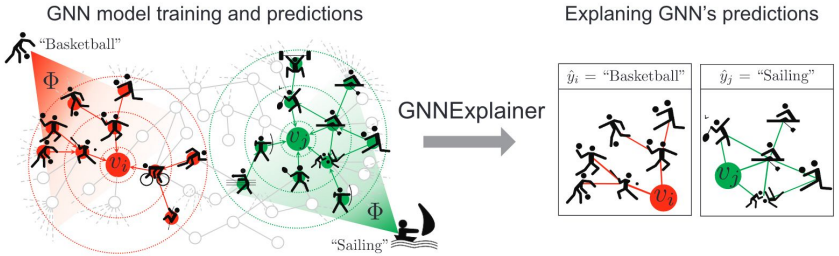
Key points

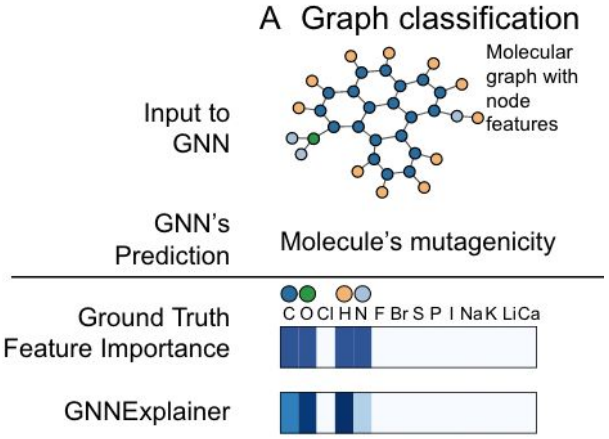
- Machine Learning -> data driven
- 2 steps :
  - Encoding : create a representation
  - Decoding : guessing the probability
- Reproduce Ecological Patterns
- Merge trait-based and network based approach
- Scalable
- Various applications
  - Link prediction
  - Introduction of new species
  - Prediction of unknown ecosystems

To go further

- Having a metric to compare networks
  - Fine-scale granularity
- Analysis of the inner mechanism of the model
  - Explainability
  - What are the important parameters ?
  - What and how much prior information do I need to measure to have good predictions ?
- Diffusion model
- Taking into account the hierarchical structure of the taxonomy
  - using hierarchical pooling
  - Improve the transfer learning metric

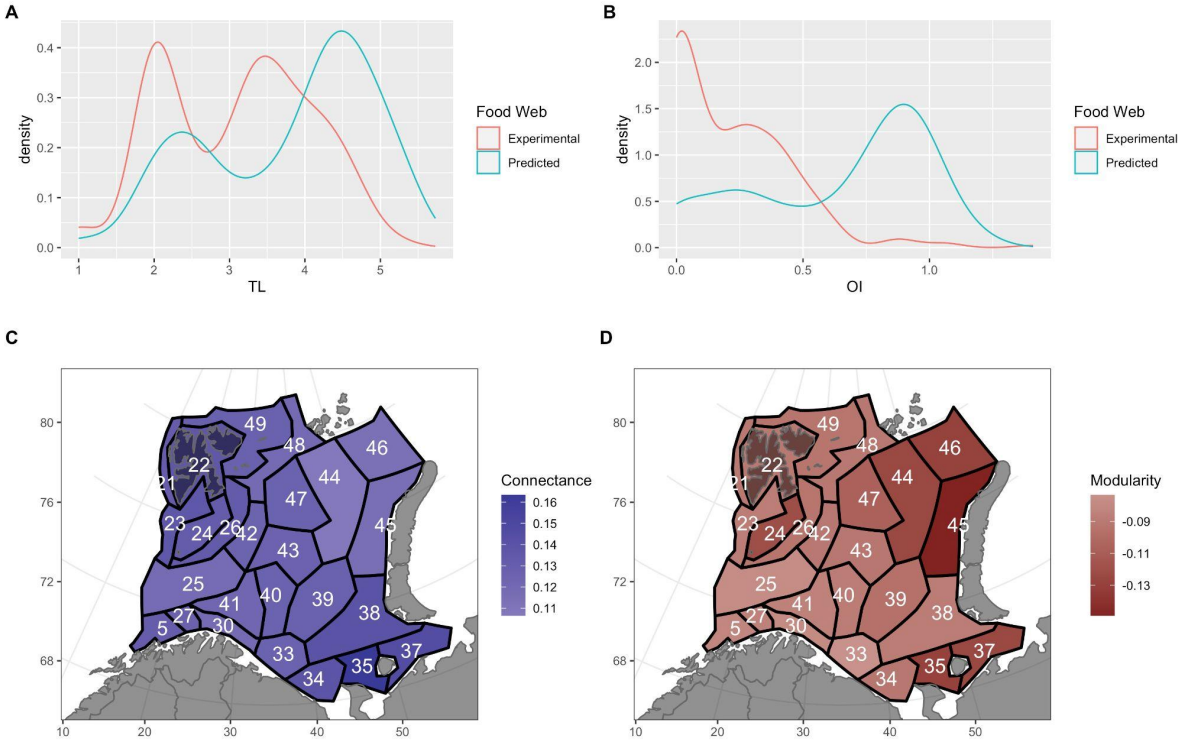
Ying et al.« GNNExplainer: Generating Explanations for Graph Neural Networks »



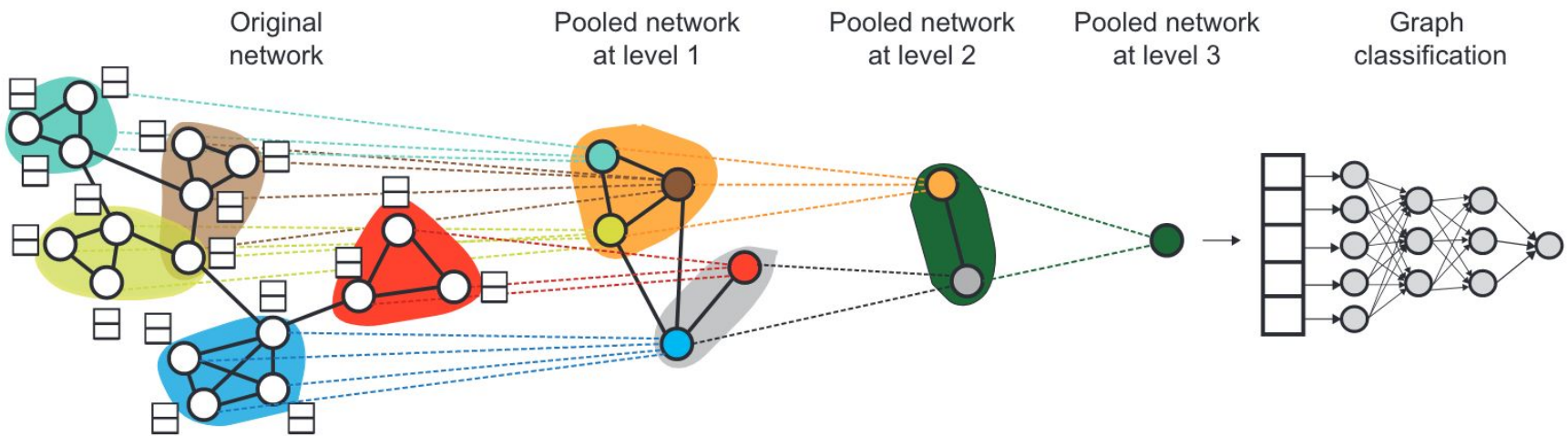


Supplementary Diapositives

Figure 7 - Network structure comparison



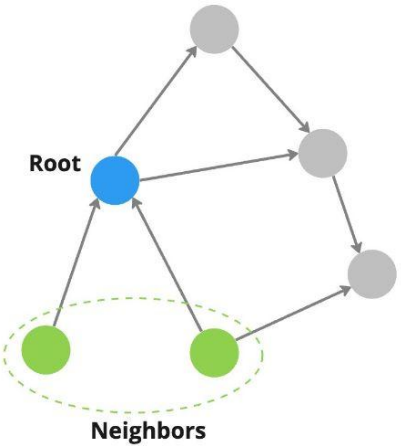
Hierarchical Pooling



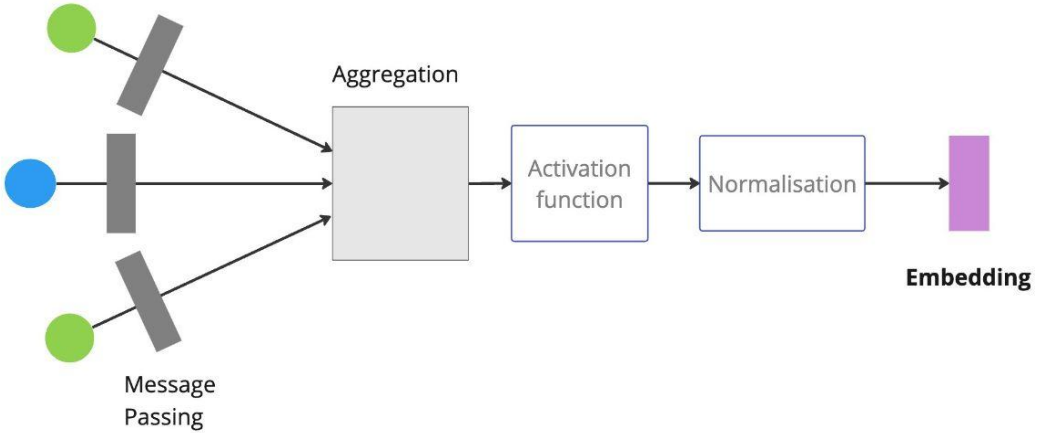
Ying et al., « Hierarchical Graph Representation Learning with Differentiable Pooling ».

Figure 2

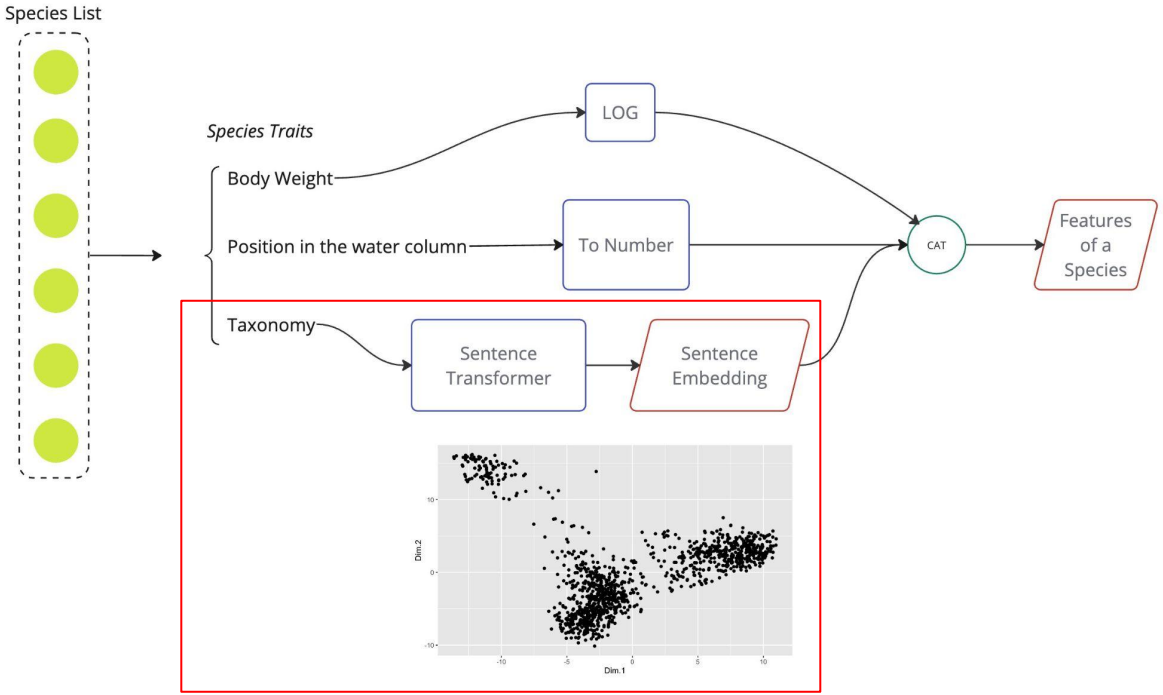
Graph



Graph Convolutional Network (GCN)



Initial Traits transformation



Kingdom	Animalia
Phylum	Arthropoda
Class	Hexanauplia
Order	Calanoida
Family	Acartiidae
Genus	Acartia
Species	Acartia

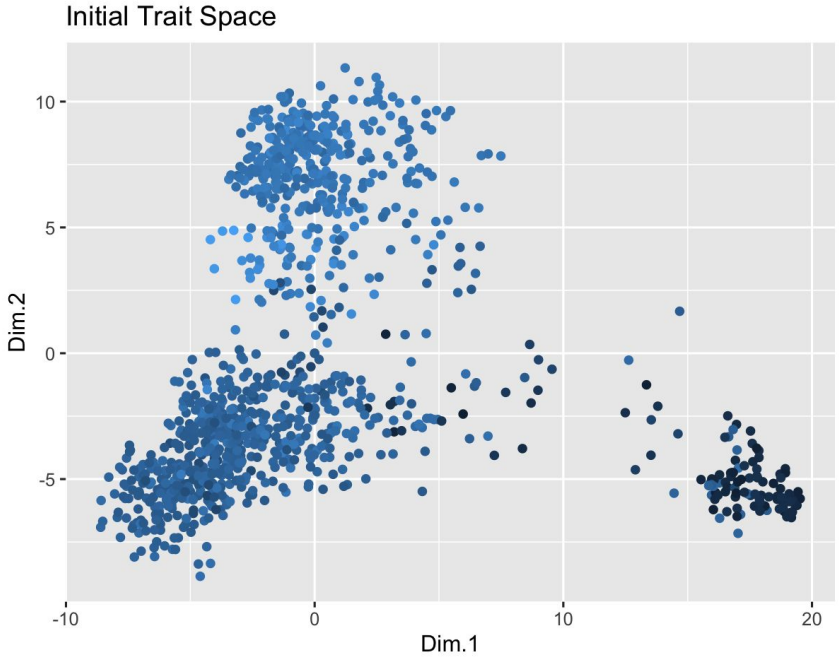
“Animalia ; Arthropoda ; Hexanauplia ; Calanoida ; Acartiidae ; Acartia ; Acartia”

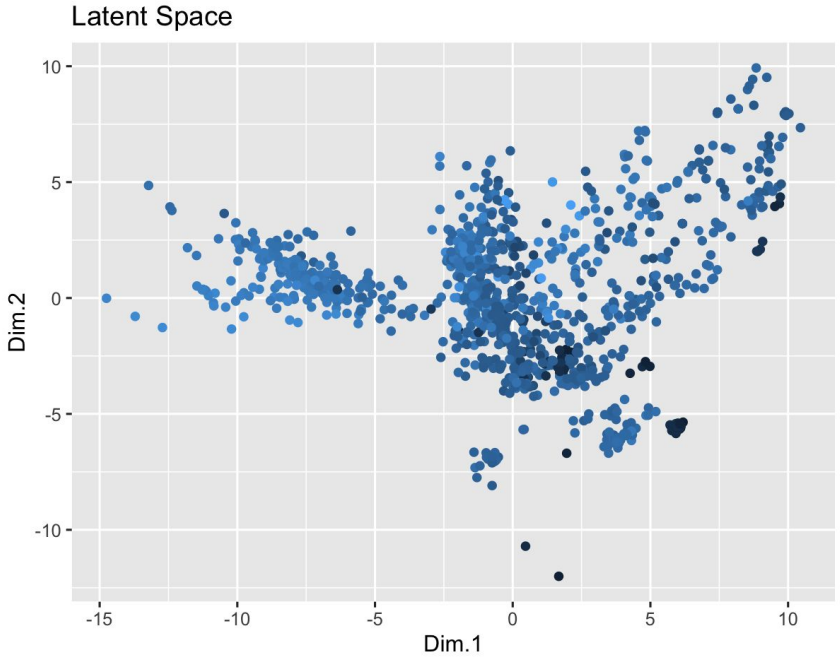
Table 1

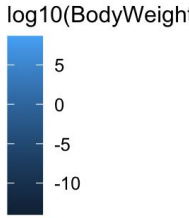
Food Web	Sensitivity	Specificity	Number of new links	Original Size
Weddell Sea	0.99	0.92	18260	15880
Barents Sea	0.94	0.86	7277	2461
Kongsfjorden	0.87	0.95	3441	1647
Reef	0.84	0.93	3718	3313
Metaweb	0.95	0.97	33830	23171

Opening the black box - understand the embedding

The Embedding - BodyWeight

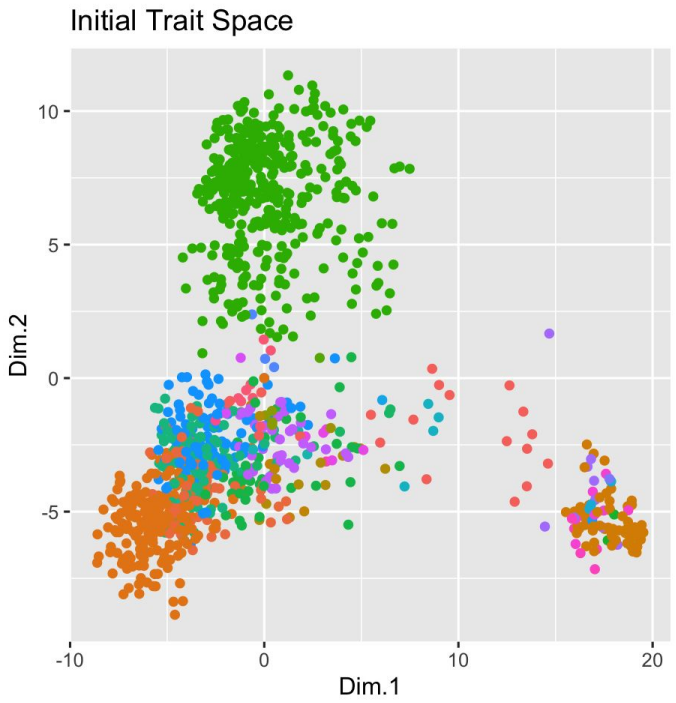


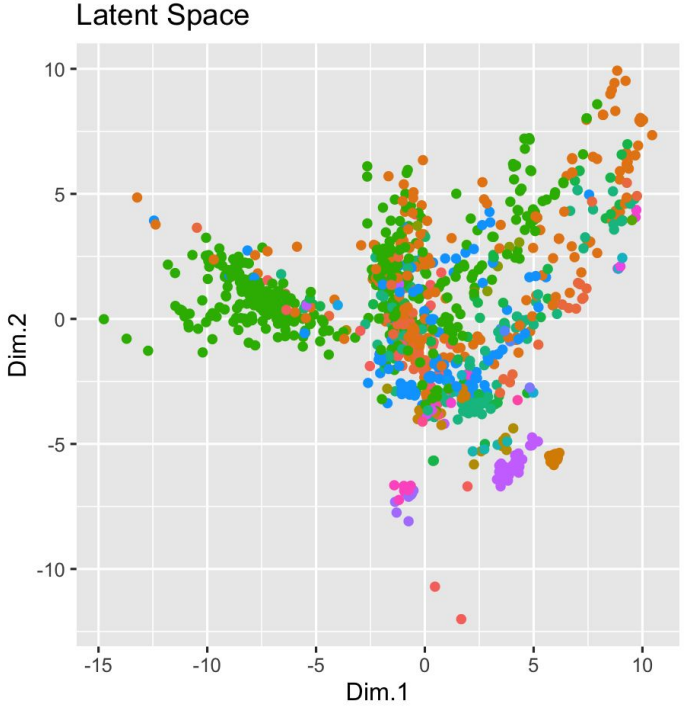




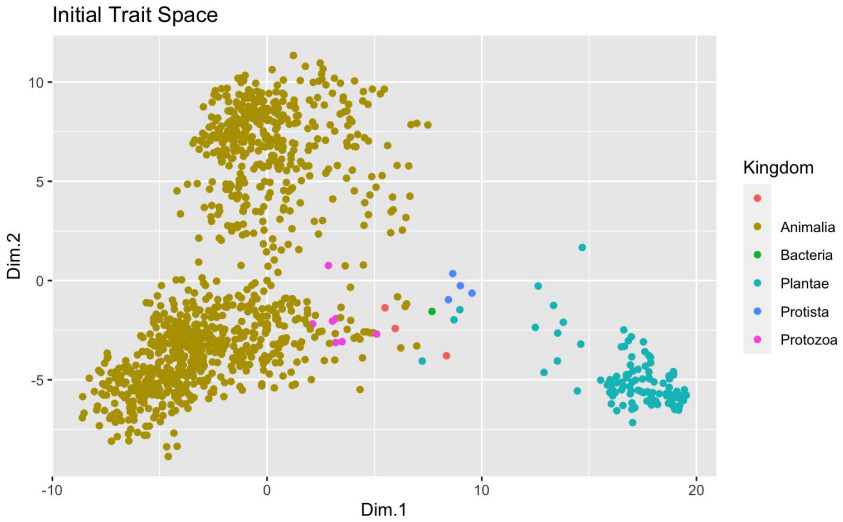
Change in structure

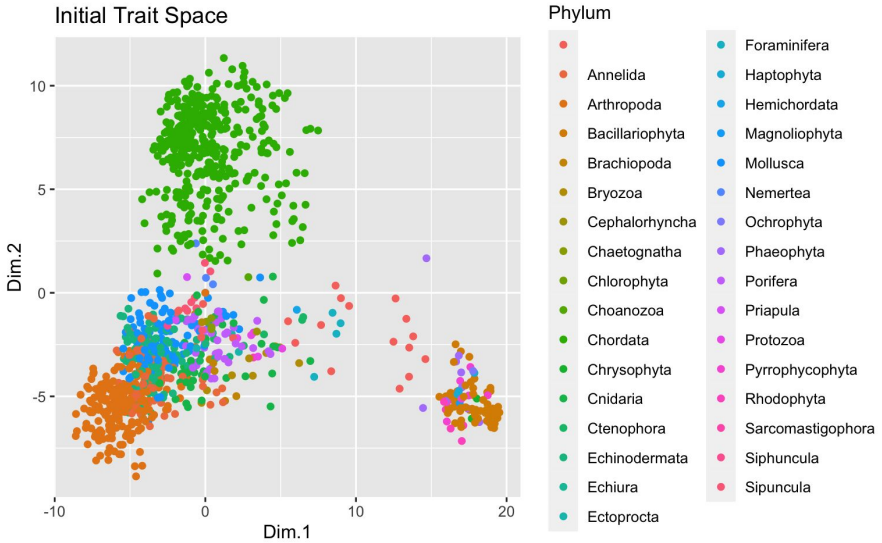
The Embedding - Taxonomy



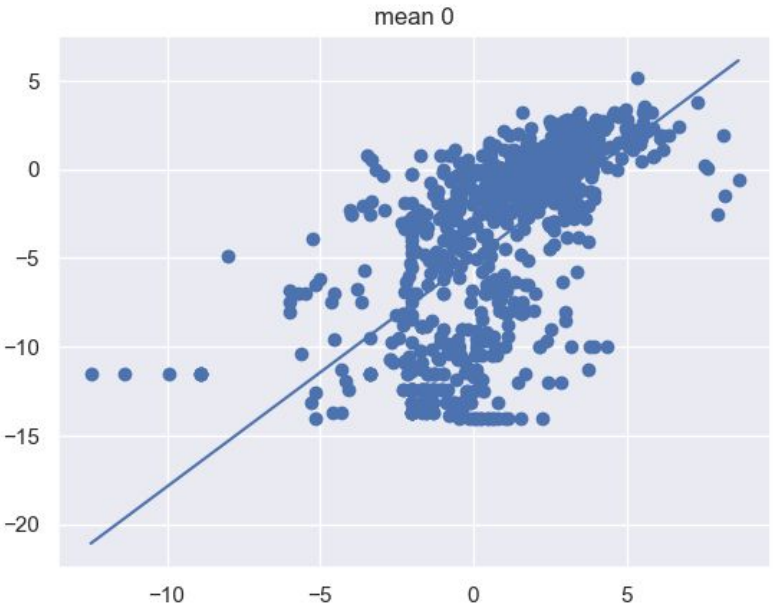


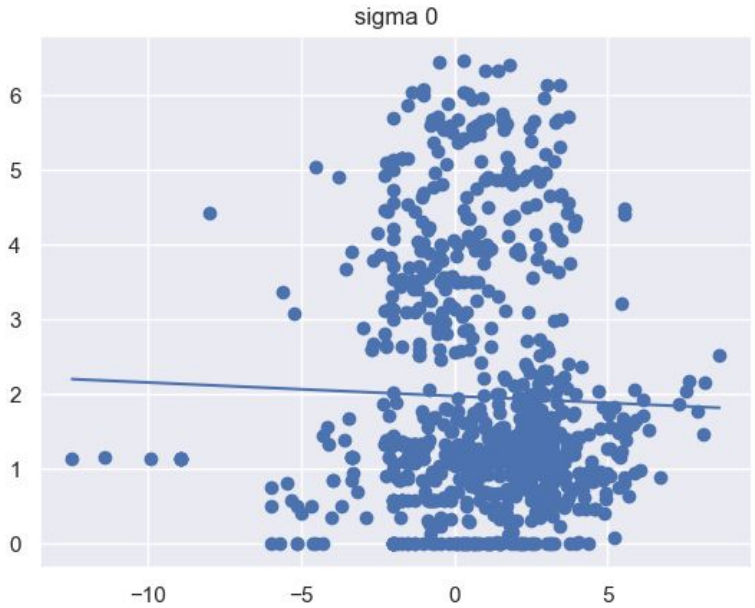
Initial traits space analysis





Niche Model regression





Figures of the article

Figure 1

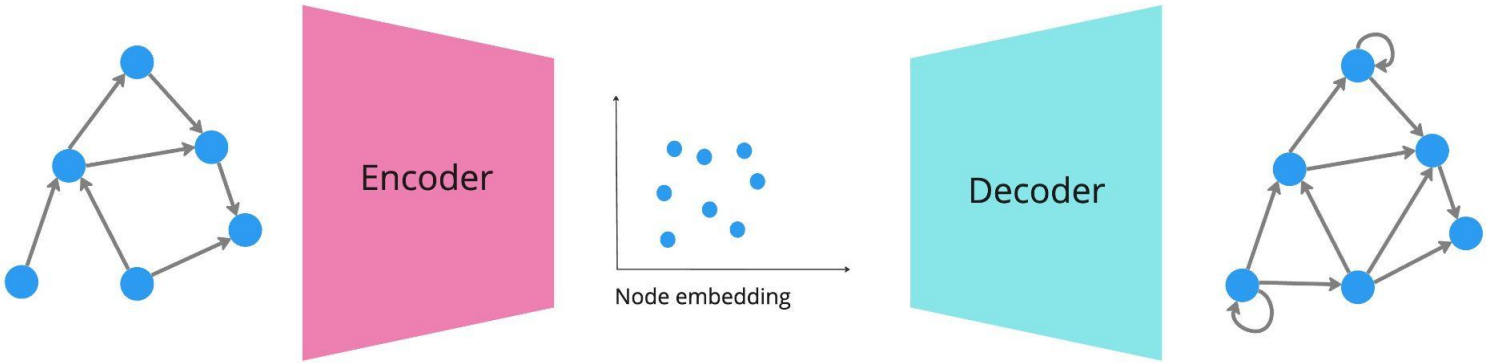
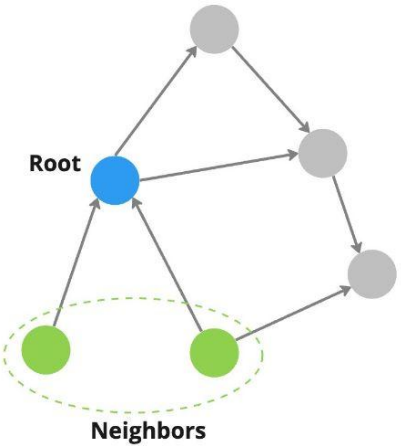


Figure 2

Graph



Graph Convolutional Network (GCN)

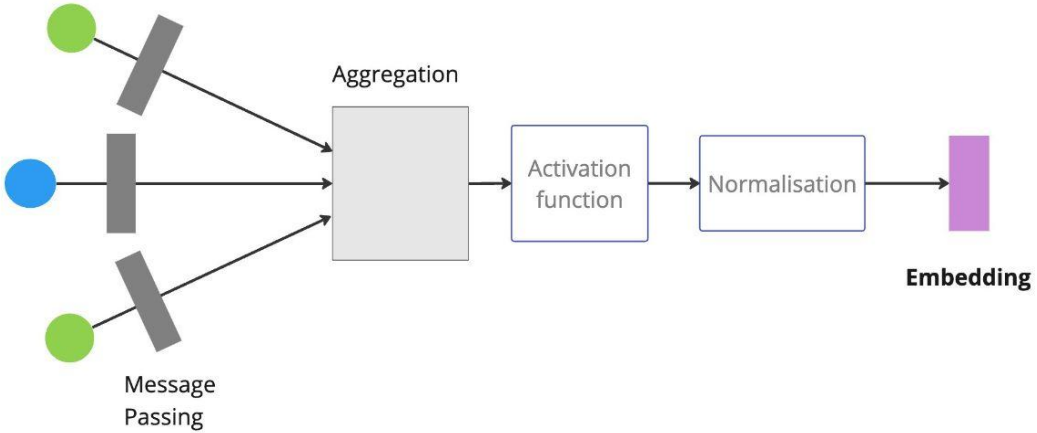


Figure 3

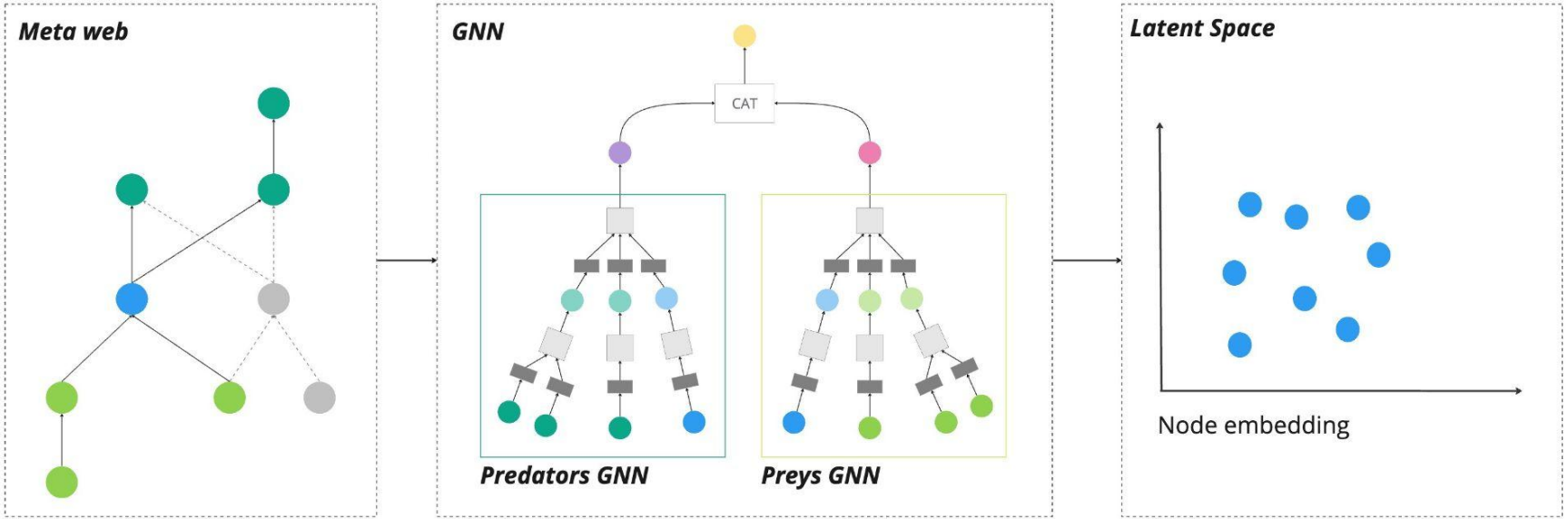

Meta web
GNN
CAT
Predators GNN
Preys GNN
Latent Space
Node embedding

Figure 4

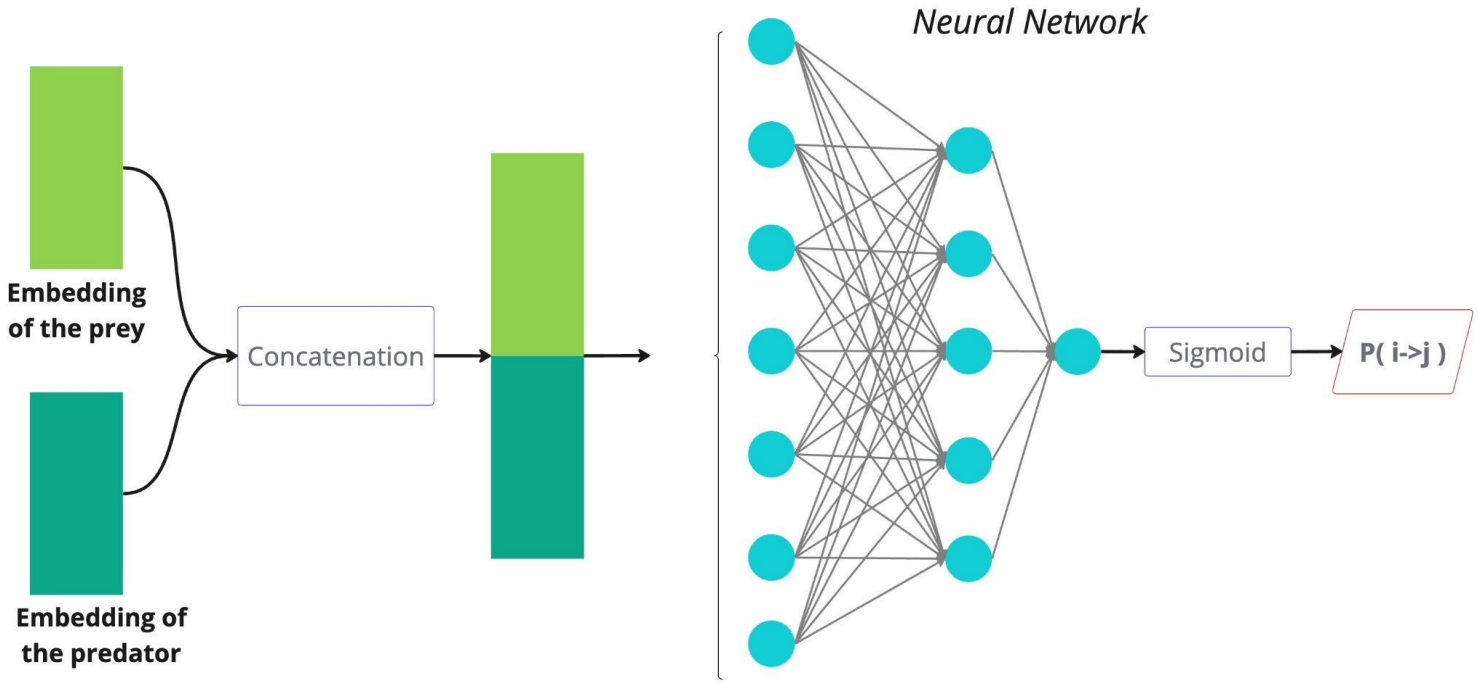
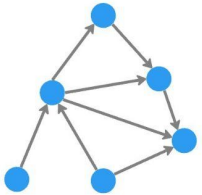
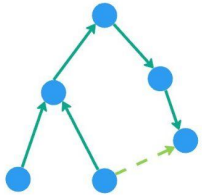

Neural Network
Embedding of the prey
Embedding of the predator
Concatenation
Sigmoid
P(i->j)

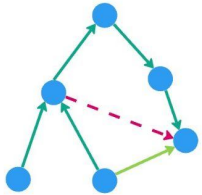
Figure 5



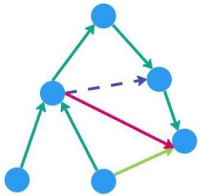
The original graph



(1) Training :
Use training message edges to predict the training message edges & training supervision edges.



(2) Validation :
Use training message edges & training supervision edges to predict the validation edges.



(3) Test :
Use training message edges & training supervision edges & validation edges to predict the test edges.

Figure 6

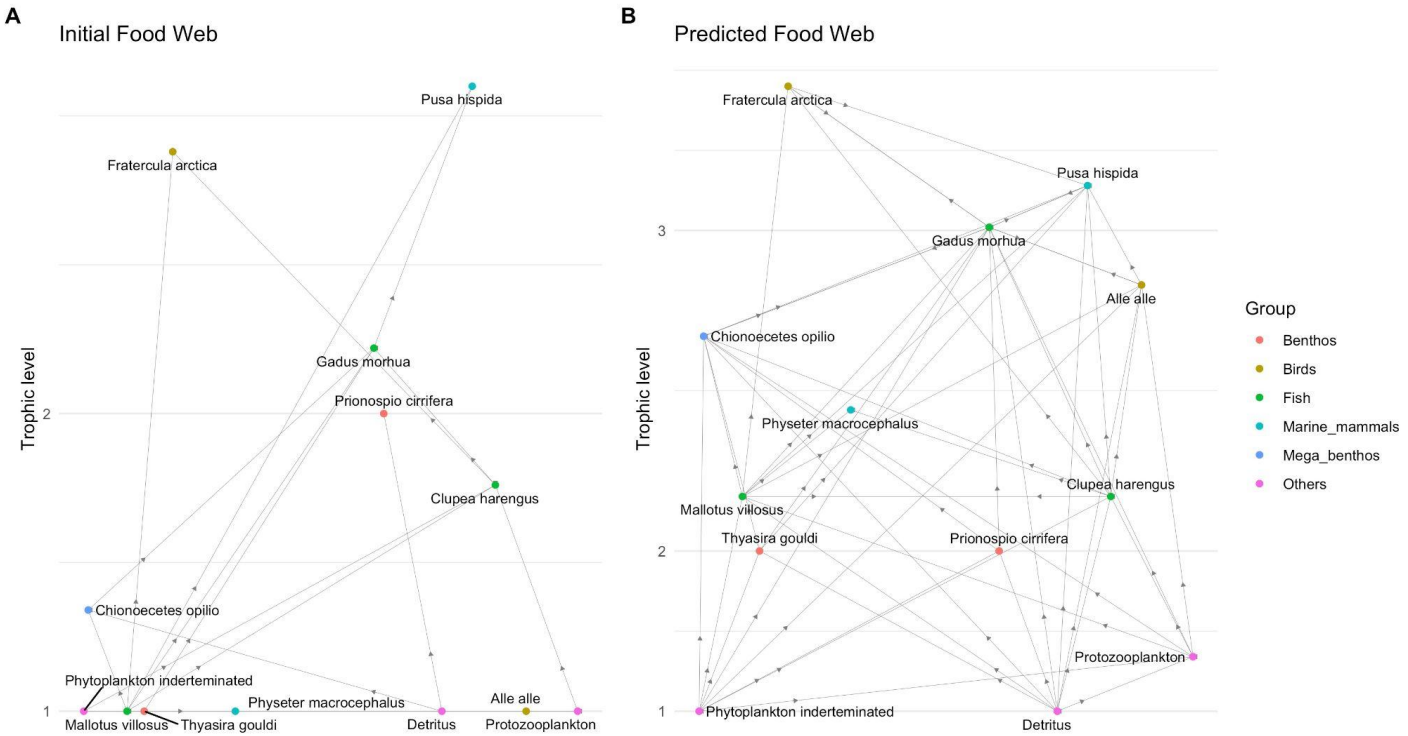

A
Initial Food Web
Trophic level
1
2
Pusa hispida
Fratercula arctica
Gadus morhua
Prionospio cirrifera
Clupea harengus
Chionoecetes opilio
Phytoplankton inderteminated
Physeter macrocephalus
Alle alle
Mallotus villosus
Thyasira gouldi
Detritus
Protozooplankton
B
Predicted Food Web
Trophic level
1
2
3
Fratercula arctica
Pusa hispida
Gadus morhua
Alle alle
Chionoecetes opilio
Physeter macrocephalus
Clupea harengus
Mallotus villosus
Thyasira gouldi
Prionospio cirrifera
Protozooplankton
Phytoplankton inderteminated
Detritus
Group
Benthos
Birds
Fish
Marine_mammals
Mega_benthos
Others

Figure 7

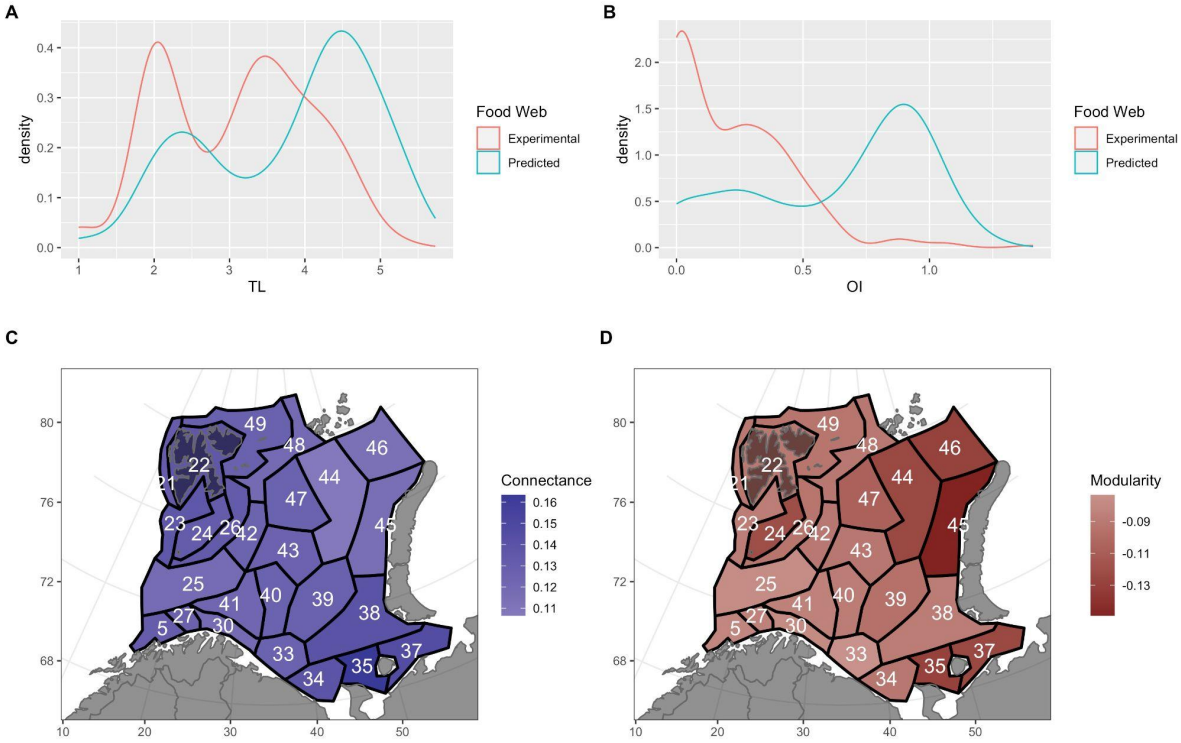

A
density
TL
Food Web
Experimental
Predicted
B
density
OI
Food Web
Experimental
Predicted
C
Connectance
0.16
0.15
0.14
0.13
0.12
0.11
D
Modularity
-0.09
-0.11
-0.13

Figure 8

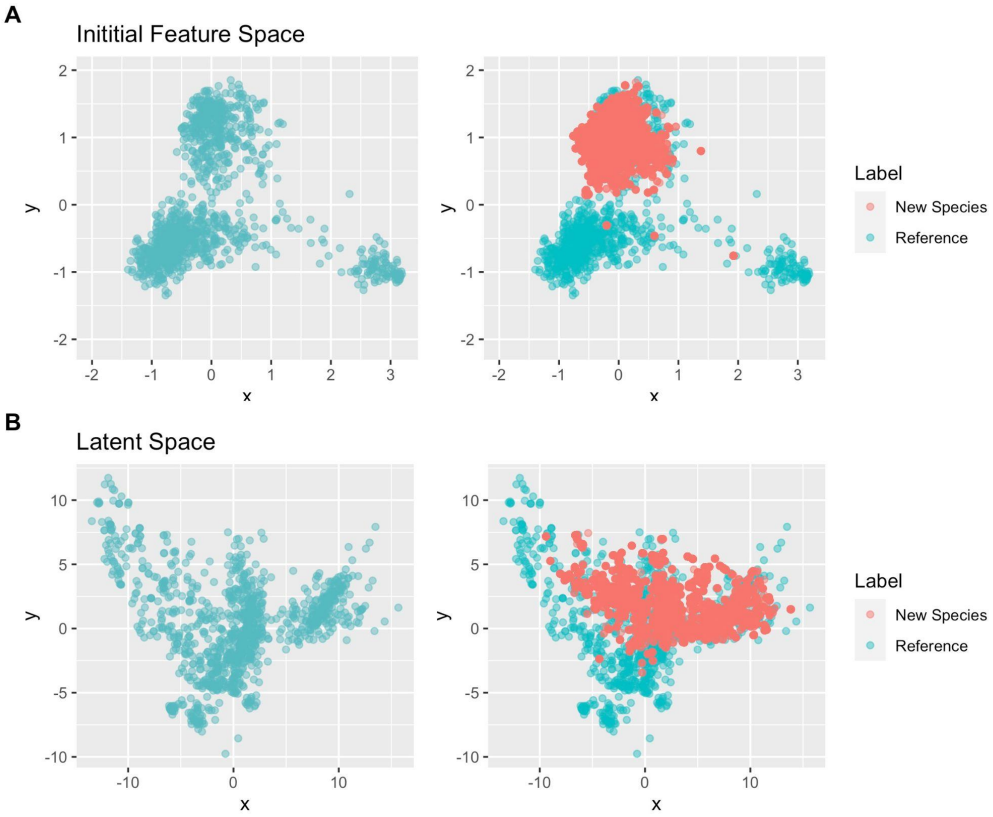

A
Initial Feature Space
B
Latent Space
x
y
Label
New Species
Reference

Figure 9

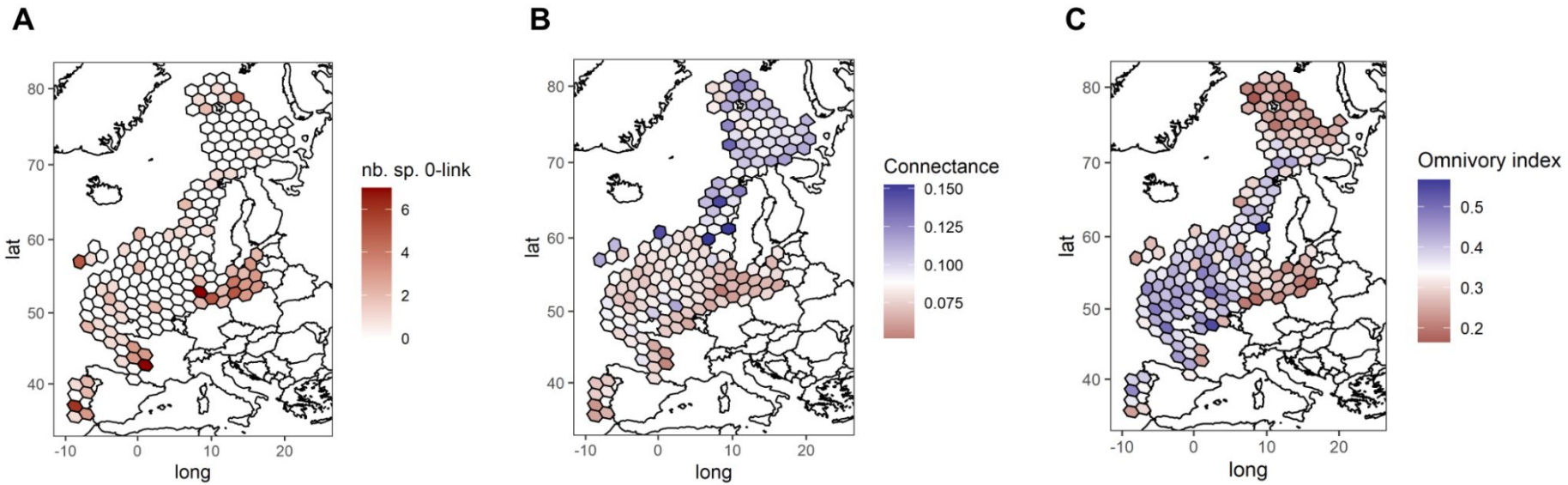

A
B
C
lat
long
nb. sp. 0-link
Connectance
Omnivory index
80
70
60
50
40
-10
0
10
20
6
4
2
0
0.150
0.125
0.100
0.075
0.5
0.4
0.3
0.2

Table 1

Food Web	Sensitivity	Specificity	Number of new links	Original Size
Weddell Sea	0.99	0.92	18260	15880
Barents Sea	0.94	0.86	7277	2461
Kongsfjorden	0.87	0.95	3441	1647
Reef	0.84	0.93	3718	3313
Metaweb	0.95	0.97	33830	23171