Towards Ecological Network Analysis with Gromov-Wasserstein Distances

Kai M. Hung
Dept. of Computer Science
Rice University
Houston, TX 77005
kai.hung@rice.edu

Alex G. Zalles
Dept. of Computational & Applied Mathematics & Operations Research
Rice University
Houston, TX 77005
agz2@rice.edu

Ann E. Finneran
Dept. of Biosciences
Rice University
Houston, TX 77005
annie.finneran@rice.edu

Lydia Beaudrot
Dept. of Biosciences
Rice University
Houston, TX 77005
beaudrot@rice.edu

César A. Uribe
Dept. of Electrical & Computer Engineering
Rice University
Houston, TX 77005
cauribe@rice.edu

Abstract

Climate change is driving the widespread redistribution of species with cascading effects on predators and their prey. Formally comparing ecological interaction networks is a critical step towards understanding the impact of climate change on ecosystem functioning, yet current methods for ecological network analysis are unable to do so. We propose using the Gromov-Wasserstein (GW) metric for quantifying dissimilarity between ecological networks. We demonstrate the corresponding optimal transport plans of this distance can be interpreted as species functional alignment between food webs. Our results show that GW transport plans align species from different mammal communities consistent with ecological understanding. Furthermore, we illustrate extensions of the GW distance to notions of averages and factorization over ecological networks. Ultimately, we propose the foundation for a novel interpretable topological data analysis framework to inform future ecological research and conservation management.

1 Introduction

Predator-prey interactions play a critical role in the maintenance of biodiversity by promoting stability (Estes et al., 2011; Landi et al., 2018). However, climate change causes many species to shift their geographic ranges as they track preferred environmental conditions (Parmesan, 2006), thereby altering ecological communities and the interactions between species within them (Williams and Jackson, 2007). Investigating how food web topology varies along environmental gradients can inform our understanding of future food webs (Tylianakis and Morris, 2017), promoting interests in analyzing ecological networks from a data-driven graph machine learning perspective (Poisot et al., 2016; Tuia et al., 2022).

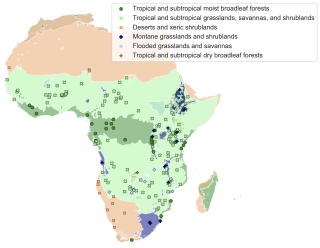

Figure 1: Location of the 170 mammal food webs across different biomes of Sub-Saharan Africa. See Appendix A.1 for details.

Despite widespread community interest in studying ecological networks, existing graph analysis and comparison methods (Tantardini et al., 2019) are either computationally intractable (Ghosh et al., 2018), not interpretable (Liu et al., 2022), dependent on extensive hyperparameter tuning and selection of motifs (e.g., graphlets (Sarajlić et al., 2016), graph kernels (Ghosh et al., 2018), and graph neural networks), or only applicable to same-sized

networks (Gera et al., 2018). Moreover, it remains an open problem to fully characterize the structural differences among ecological networks (Pellissier et al., 2018).

This paper proposes using the Gromov-Wasserstein (GW) distance to quantify dissimilarity between ecological networks. In the context of networks, the GW distance is a pseudometric (Chowdhury and Mémoli, 2019) with competitive empirical performance on tasks such as graph clustering (Chowdhury and Needham, 2021), graph partitioning (Xu et al., 2019), and graph factorization (Xu, 2020). More importantly, unlike current spectral methods based on graph Laplacians (Wilson and Zhu, 2008), GW compares graphs of different sizes, which is critically important for ecological networks because communities support different numbers of species. GW distances have corresponding optimal transport plans describing how the nodes are mapped between graphs. In the ecological context, this optimal transport plan can reveal corresponding functional roles between two species in different food webs. Furthermore, GW distance can be extended into a suite of distance-based machine learning algorithms onto ecological networks.

The main contributions of this paper are described as follows:

1. Measure graph characterization: Computing the GW distance requires a probability distribution representation of the graphs. We formalize the viable configurations, discuss their ecological relevance, and contrast their performance.

2. GW for topology representations: We present empirical evidence that the GW distance clusters synthetic graphs of varying sizes into groups corresponding to its topology: paths, cycles, and stars. Furthermore, we show that the GW distance produces geometrically intuitive transport plans and barycenters.

3. GW on food webs: We validate our proposed method by analyzing 170 Sub-Saharan mammal communities. We first illustrate how the optimal transport matrix successfully maps many species to themselves in a pair of food webs with many shared species. We also identify mismatched species based on their functional similarities as understood in ecology. Furthermore, we extend our analyses to Fréchet Mean and Graph Factorization over food webs, showcasing their applicability to ecology. We attached our results in the appendix due to space constraints.

2 Gromov-Wasserstein for Ecological Networks

In this section, we formally define configurations of probability distributions on graphs as required by computing the network GW distance. Their ecological interpretations and trade-offs follow the definition.

2.1 Measure Networks with Ecological Interpretations

As an instance of optimal transport metrics (Peyré et al., 2019), the GW distance quantifies *minimal transportation cost* between probability distributions that live in different sample spaces (i.e., with different number of nodes in the context of graphs) (Mémoli, 2011). Applying GW distance to networks requires characterizing them as *measure networks*, i.e., a probability distribution over the network. Formally, given a simple directed graph $G = (V, E)$ with n nodes and m edges, we define the *measure network* as $\mathcal{M}_G := (p_G, w_G)$ where p_G is the node distribution and $w_G : V \times V \to \mathbb{R}$ is a bounded, real-valued edge weight function. Without loss of generality, we consider *uniform* and *degree* node distribution, as well as the *adjacency* and *shortest path* edge weight functions. Specifically, the *uniform* node distribution allocates each node $v \in V$ with a probability of $1/n$. The *degree* node distribution endows each node with a mass proportional to its number of neighbors, computed by normalizing the sum of neighbors for every $v \in V$. The *adjacency* formulation of w_G is binary depending on the existence of an edge between two nodes, akin to the adjacency matrix. If such a path exists, the *shortest path* formulation places $w_G(u, v)$ as the shortest path from node u to node v, and a large bounded real-value B otherwise. We will discuss their trade-offs and interpretation in the context of the GW distance below.

2.2 Discrete Network Gromov-Wasserstein Distance

Given two measure networks $\mathcal{M}_{G_1} = (p_{G_1}, w_{G_1})$ and $\mathcal{M}_{G_2} = (p_{G_2}, w_{G_2})$ with n_1 and n_2 nodes respectively, we define the *discrete p-network Gromov-Wasserstein distance* as

$$d_{\mathcal{N},p}(G_1, G_2) := \min_{\substack{T \in \mathbb{R}_{\geq 0}^{n_1 \times n_2}, \\ T\mathbb{1} = p_{G_1}, \\ T^\top \mathbb{1} = p_{G_2}}} \left\{ \sum_{u,u' \in V_1} \sum_{v,v' \in V_2} |w_{G_1}(u,u') - w_{G_2}(v,v')|^p T_{u,v} T_{u',v'} \right\} \quad (1)$$

where $T \in \mathbb{R}_{\geq 0}^{n_1 \times n_2}$ is a transport matrix between $\mathcal{M}_{G_1}$ and $\mathcal{M}_{G_2}$, i.e., T is a non-negative real-valued matrix satisfying the constraints $T\mathbb{1} = p_{G_1}$ and $T^\top \mathbb{1} = p_{G_2}$ where $\mathbb{1}$ is a vector of all 1's of appropriate dimension. Our formulation is a discretized distance variant proposed in Chowdhury and Mémoli (2019). In Eq. 1, we overload the notation by using $T_{u,v}$ to represent the mass transported from node u in graph G_1 to node v in graph G_2 as opposed to using their indices for ease of interpretation. For simplicity of exposition, we set $d_{gw} := d_{\mathcal{N},p}$ with $p = 1$ in the rest of this paper. The distance d_{gw} quantifies the minimal work required for mass transfer between nodes of two graphs. Notably, the "work required" measures obstruction to relational structures (defined by w_G), weighted by the importance of the involved nodes (governed by p_G). We refer the audience to (Chowdhury and Mémoli, 2019) for extended discussions on the theoretical properties of the d_{gw} pseudometric.

In the ecological context, entry (u, v) of T signals the correspondence between species u and v. Upon close examination of the GW distance formulation, one can observe that the node mass $p_G(v)$ dictates the priority of aligning node v well with respect to the graph topology captured by $w_G(v)$. The choice of the node distribution p_G then comes down to a preferred measure of species importance. With the *uniform* formulation, we do not differentiate between any species. With the *degree* formulation, we identify species importance through their relative interactions with respect to other species. The choice of edge weight function w_G governs the topological information to preserve, with *adjacency* formulation accounting for only direct predator-prey interactions and *shortest path* capturing relationships across hierarchies. In practice, the shortest path configuration is more computationally intensive than the adjacency, as ecological networks are usually directed and not strongly connected. On the contrary, if a higher-order relationship amongst species is of key interest, we recommend the shortest path configuration. Next, we will present empirical evidence for the applicability of the GW distance on synthetic graphs and real-world ecological networks.

3 Numerical Analysis on Food Webs

We demonstrate the ability of GW distance to capture well-defined topological structures. Our method can uncover relationships consistent with auxiliary ecological features through graph topology alone with experiments on complex, real-world ecological networks.

We perform our experiments in Python, using the `NetworkX` library by Hagberg et al. (2008) for generating graphs and the Python Optimal Transport (`POT`) library by Flamary et al. (2021) for computing d_{gw} and transport plans. Computing the d_{gw} efficiently per Eq. 1 requires solving a difficult non-convex optimization problem beyond the scope of our project. All computations of d_{gw} in this work utilize the widely available solver `ot.gromov.gromov_wasserstein` in `POT`.

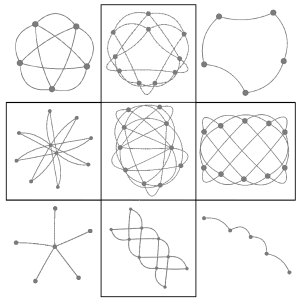

Figure 2: GW barycenters of support graphs depicted in the corner. The center graph is a barycenter of all supports, whereas the outer rim graphs are barycenters of their neighboring support graphs.

3.1 OT Plan, Barycenter, and MDS over Synthetic Graphs

We show both (1) the geometric interpretability of GW-based graph dissimilarity analysis and (2) empirical evidence that the GW distance differentiates families of synthetic networks.

We begin by evaluating the optimal transport (OT) plan for d_{gw} on simple graph models according to Eq. 1. The OT plan between 25-node and 50-node path graphs in Fig. 7 suggests a geometrically intuitive stretching transportation under the network GW distance when moving from the shorter path graph to a longer one. For cycle graphs, the OT plan in Fig. 7 shows that the mass of each node is equally and sequentially distributed to a half-ring since every node shares identical relationships with its neighbors.

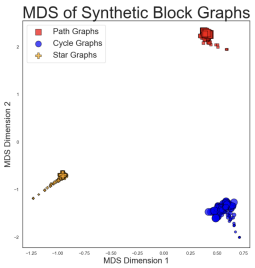

Figure 3: MDS of synthetic block graphs under d_{gw}. Size is scaled by the number of nodes.

Moving from pairwise graphs $\{G_1, G_2\}$ to a set of graphs $\{G_i\}$, we extend the notion of Fréchet mean to the graph space via the GW barycenter (Peyré et al., 2016). In our experiment, we generate 4 well-known graphs (complete, cycle, star, and path), configuring them with uniform p_G and adjacency w_G. Fig. 2 showcases the 10-node reconstructed barycenter graphs with the 4 simple graphs as their "support graphs" computed with the GW barycenter algorithm proposed in (Peyré et al., 2016) in the POT library. See Appendix A.3 for details.

Next, we show that the Multi-Dimensional Scaling (MDS) embeddings (Torgerson, 1952) of GW distance can separate graphs into different clusters consistent with their topological structure. We begin with cycle, path, and star graphs with 5 nodes. We then randomly sample 100 block sizes from 5 to 70, replacing each node in the original graphs with a fully connected block (see Fig. 8 for 4-block variants). We generated 100 networks of each family with different sizes and computed their pairwise d_{gw} with uniform p_G and shortest path w_G. Using these pairwise distances, we generate 2-D MDS embeddings for the graphs using the `Scikit-Learn` library (Pedregosa et al., 2011). In Fig. 10, this unsupervised approach, supported by the d_{gw}, perfectly generates clusters corresponding to the graph families.

3.2 Interpreting OT plan among Sub-Saharan food webs

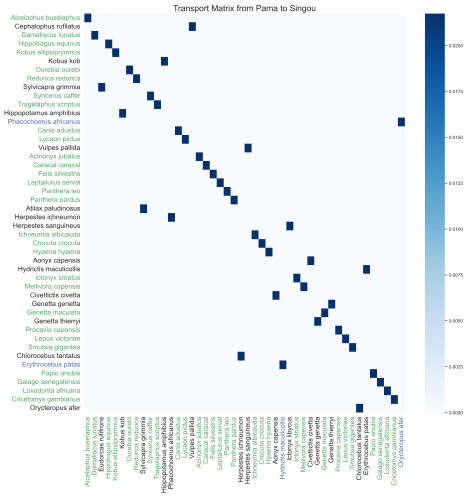

Figure 4: Optimal transport plan between mammal communities at Pama and Singou in Sub-Saharan Africa. Green highlights denote mass transported between shared species. Blue highlights surprising basal species mappings.

Our group constructed a dataset of 170 mammal food webs across Sub-Saharan Africa by extracting species-to-network mapping from mammal community composition data compiled by (Rowan et al., 2020) and predator-prey species interactions documented by (Kingdon, 2013). See Appendix A.1 for details on the dataset. We represent the ecological networks as a directed graph, extract the largest weakly connected component, and configure their corresponding network measure using the *uniform* node distribution and *shortest path* edge weight function. The former is chosen since we assume no prior information on species importance, and the latter is used to capture higher-order relationships among food webs.

From the 170 mammal food webs, we show the OT plan between the two closest networks with respect to the GW distance in Fig. 4. Overall, we found that OT plans worked very well for mapping large-bodied predator species to themselves when a species was found in both food webs. For example, this transport successfully mapped shared predator species between the two webs including lion (*Panthera leo*), leopard (*Panthera pardus*), cheetah (*Acinonyx jubatus*), etc. In some cases, the OT plans identified functionally equivalent but taxonomically different species such as caracal (*Caracal caracal*) for African wild dog (*Lycaon pictus*) in A.5.

Our OT plan produces surprising alignment among basal species, i.e., mammals that do not eat other mammals. For instance, it maps patas monkey (*Erythrocebus patus*) to spotted-necked otter (*Hydrictis maculicollis*), and warthog (*Phacochoerus africanus*) to aardvark (*Orycteropus afer*)). This limitation is an artifact of employing the shortest path distance since the shortest path from the leaf nodes to any other nodes is some large scalar B (1000 in our case).

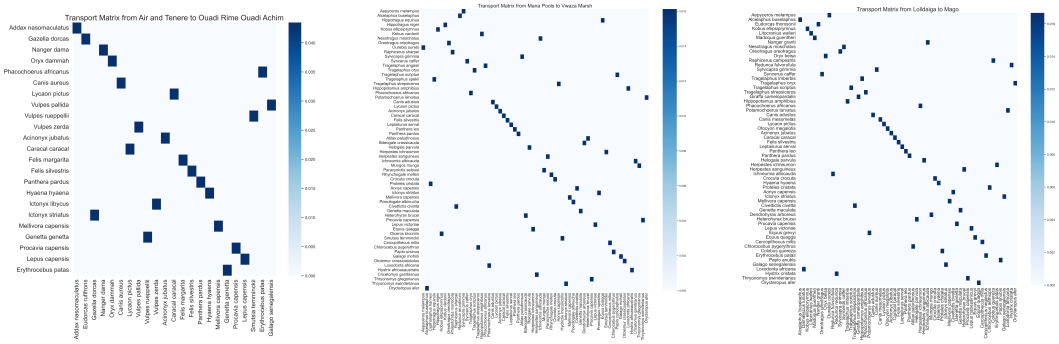

Figure 5: OT plans for the top-3 most similar networks across different biome types under d_{gw}. For detailed listing, refer to Table 1. Observe that many species are aligned to themselves across food webs as indicated in our analyses for Fig. 4.

Examining the nine most similar food webs under the d_{gw} (listed in Table 1) reveals that the list is dominated with webs from the *tropical and subtropical grasslands, savannas, and shrublands* biome. Fig. 5 demonstrates the top-9 OT plans outside of this biome. Detailed analyses of the top-9 most similar webs can be found in Appendix A.5. Further analyses of food webs across different environmental conditions are available in Appendix A.6 and A.7.

4 Conclusions and Future Work

We propose the Gromov-Wasserstein (GW) distance as a novel comparison method between ecological networks. One primary advantage of this distance relative to existing graph distances is the optimal transport matrix, which can provide insights into how parts of two network may be similar or dissimilar to one another. Our continuing work hope to further exploit this desirable property:

1. Uncertain species interactions: Ecological networks often have uncertain interactions, producing directed graphs with edge weights in the $[0, 1]$ range. The GW distance formulation can be extended to such settings naturally by defining the node distribution $p_G(v)$ of a node $v \in G$ as the sum of its edge weights, then normalized to satisfy the definition of p_G. In this setting, species with uncertain interaction will contribute less to the GW objective penalty, thereby reducing their impact on the GW distance.

2. Connecting species functional traits to topology: The use of functional traits (such as body mass and diet) to cluster species is widely adapted in ecology. Often, ecologists are interested in how these functional traits affect the composition of a food web. One natural approach is to compute correlations between the optimal transport plan and functional attribute dissimilarity matrix, thereby quantifying the consistency of select functional attributes between species and their soft matching under the GW distance. Furthermore, one could directly introduce functional attributes into the graph comparison via the Fused Gromov-Wasserstein distance proposed by Vayer et al. (2020).

Perhaps most exciting is the use of Gromov-Wasserstein distance as the ground metric for many classical learning algorithms, enabling a new suite of data-driven methods for studying ecological networks.

5 Acknowledgements

This work is supported by the U.S. National Science Foundation Grant No. 2213568.

References

J. H. Brown, O. Reichman, and D. W. Davidson. Granivory in desert ecosystems. *Annual Review of Ecology and Systematics*, 10(1):201–227, 1979.

N. P. Chacoff, D. P. Vázquez, S. B. Lomáscolo, E. L. Stevani, J. Dorado, and B. Padrón. Evaluating sampling completeness in a desert plant–pollinator network. *Journal of Animal Ecology*, 81(1):190–200, 2012.

S. Chowdhury and F. Mémoli. The gromov–wasserstein distance between networks and stable network invariants. *Information and Inference: A Journal of the IMA*, 8(4):757–787, 2019.

S. Chowdhury and T. Needham. Generalized spectral clustering via gromov-wasserstein learning. In *International Conference on Artificial Intelligence and Statistics*, pages 712–720. PMLR, 2021.

J. A. Estes, J. Terborgh, J. S. Brashares, M. E. Power, J. Berger, W. J. Bond, S. R. Carpenter, T. E. Essington, R. D. Holt, J. B. Jackson, et al. Trophic downgrading of planet earth. *science*, 333(6040):301–306, 2011.

R. Flamary, N. Courty, A. Gramfort, M. Z. Alaya, A. Boisbunon, S. Chambon, L. Chapel, A. Corenflos, K. Fatras, N. Fournier, L. Gautheron, N. T. Gayraud, H. Janati, A. Rakotomamonjy, I. Redko, A. Rolet, A. Schutz, V. Seguy, D. J. Sutherland, R. Tavenard, A. Tong, and T. Vayer. Pot: Python optimal transport. *Journal of Machine Learning Research*, 22(78):1–8, 2021.

R. Gera, L. Alonso, B. Crawford, J. House, J. Mendez-Bermudez, T. Knuth, and R. Miller. Identifying network structure similarity using spectral graph theory. *Applied network science*, 3:1–15, 2018.

S. Ghosh, N. Das, T. Gonçalves, P. Quaresma, and M. Kundu. The journey of graph kernels through two decades. *Computer Science Review*, 27:88–111, 2018.

A. Hagberg, P. Swart, and D. S Chult. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

J. Kingdon. *Mammals of Africa*. Bloomsbury London, 2013.

H. Kreft and W. Jetz. A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, 37(11):2029–2053, 2010.

P. Landi, H. O. Minoarivelo, Å. Brännström, C. Hui, and U. Dieckmann. Complexity and stability of ecological networks: a review of the theory. *Population Ecology*, 60:319–345, 2018.

N. Liu, Q. Feng, and X. Hu. Interpretability in graph neural networks. *Graph Neural Networks: Foundations, Frontiers, and Applications*, pages 121–147, 2022.

J. Looman and J. Campbell. Adaptation of sorensen's k (1948) for estimating unit affinities in prairie vegetation. *Ecology*, 41(3):409–416, 1960.

F. Mémoli. Gromov–wasserstein distances and the metric approach to object matching. *Foundations of computational mathematics*, 11:417–487, 2011.

C. Parmesan. Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.*, 37:637–669, 2006.

M. Pascual and J. A. Dunne. *Ecological networks: linking structure to dynamics in food webs*. Oxford University Press, 2006.

F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830, 2011.

L. Pellissier, C. Albouy, J. Bascompte, N. Farwig, C. Graham, M. Loreau, M. A. Maglianesi, C. J. Melián, C. Pitteloud, T. Roslin, et al. Comparing species interaction networks along environmental gradients. *Biological Reviews*, 93(2):785–800, 2018.

G. Peyré, M. Cuturi, and J. Solomon. Gromov-wasserstein averaging of kernel and distance matrices. In *International conference on machine learning*, pages 2664–2672. PMLR, 2016.

G. Peyré, M. Cuturi, et al. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.

T. Poisot, D. B. Stouffer, and S. Kéfi. Describe, understand and predict. *Functional Ecology*, 30(12):1878–1882, 2016.

J. Rowan, L. Beaudrot, J. Franklin, K. E. Reed, I. E. Smail, A. Zamora, and J. M. Kamilar. Geographically divergent evolutionary and ecological legacies shape mammal biodiversity in the global tropics and subtropics. *Proceedings of the National Academy of Sciences*, 117 (3):1559–1565, 2020.

A. Sarajlić, N. Malod-Dognin, Ö. N. Yaveroğlu, and N. Pržulj. Graphlet-based characterization of directed networks. *Scientific reports*, 6(1):35098, 2016.

M. Tantardini, F. Ieva, L. Tajoli, and C. Piccardi. Comparing methods for comparing networks. *Scientific reports*, 9(1):17557, 2019.

W. S. Torgerson. Multidimensional scaling: I. theory and method. *Psychometrika*, 17(4): 401–419, 1952.

D. Tuia, B. Kellenberger, S. Beery, B. R. Costelloe, S. Zuffi, B. Risse, A. Mathis, M. W. Mathis, F. van Langevelde, T. Burghardt, et al. Perspectives in machine learning for wildlife conservation. *Nature communications*, 13(1):792, 2022.

J. M. Tylianakis and R. J. Morris. Ecological networks across environmental gradients. *Annual Review of Ecology, Evolution, and Systematics*, 48:25–48, 2017.

T. Vayer, L. Chapel, R. Flamary, R. Tavenard, and N. Courty. Fused gromov-wasserstein distance for structured objects. *Algorithms*, 13(9):212, 2020.

J. W. Williams and S. T. Jackson. Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, 5(9):475–482, 2007.

R. C. Wilson and P. Zhu. A study of graph spectra for comparing graphs and trees. *Pattern Recognition*, 41(9):2833–2841, 2008.

H. Xu. Gromov-wasserstein factorization models for graph clustering. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 6478–6485, 2020.

H. Xu, D. Luo, and L. Carin. Scalable gromov-wasserstein learning for graph partitioning and matching. *Advances in neural information processing systems*, 32, 2019.

H. Xu, J. Liu, D. Luo, and L. Carin. Representing graphs via gromov-wasserstein factorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(1):999–1016, 2022.

A Appendix

A.1 Dataset description

Our dataset comprises mammal food webs from 170 protected regions across Sub-Saharan Africa. In particular, we leverage published terrestrial community composition data from (Rowan et al., 2020), which contains 258 mammal species from 12 orders and 33 families. The mammal community data were collected from field surveys, species lists, and databases and then cross-referenced with International Union for Conservation of Nature (IUCN) Red List range maps and standardized to IUCN Red List taxonomy. The communities contain only mammals weighing $\geq$ 500 grams because data for small taxa are far less available and less reliable. To construct the food webs from mammal communities, we utilize the common meta-web approach (Pascual and Dunne, 2006) for large-scale spatial studies. In particular, we establish connections between species based on predator-prey species interaction documented in (Kingdon, 2013). Fig. 1 visualizes the mammal food webs over the African continent, with labels corresponding to biome categories. Fig. 6 contains kernel density estimations of the node, edge, and aggregated degree distribution across all 170 constructed food webs.

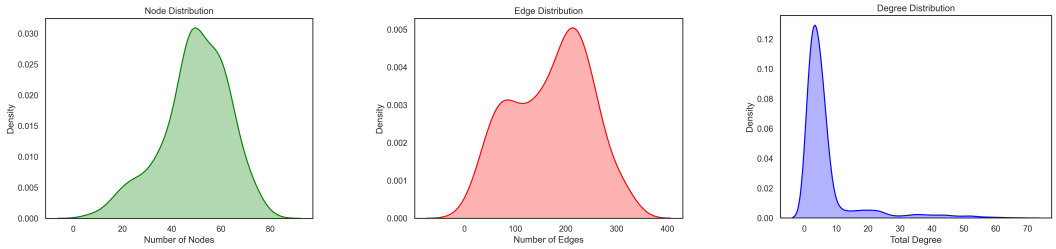

Figure 6: The kernel density estimation plots of graph-level statistics across the 170 mammal communities in 1. (Left) Node per graph distribution. (Center) Edge per graph distribution. (Right) Degree per node distribution across all graphs.

A.2 Path and cycle optimal transport plans

For this experiment, we used the `NetworkX` library in Python to generate path and cycle graphs with 25 and 50 nodes respectively. Then, we configure them into measure networks using the uniform and shortest path settings. In Fig. 7, we visualize the optimal transport plan between the respective graphs of the same class, yet with varying sizes.

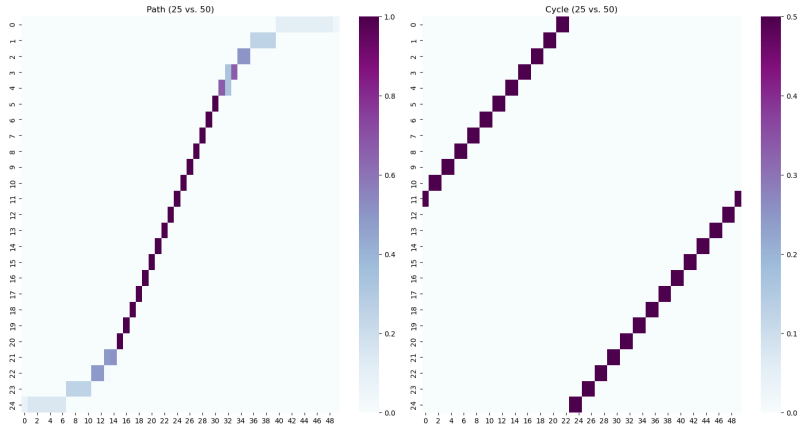

Figure 7: OT plans for d_{gw} computation between 25-node and 50-node path graphs (left) and cycle graphs (right). The axes are node IDs. Nodes in the path graph are labeled from end-to-end, whereas the cycle graph is labeled consecutively starting with an arbitrary node.

A.3 Block graphs

For this experiment, we compute the d_{gw} between base graphs and n-block graphs for $n = 4$ across the star, path, and cycle families. The construction of an n-block graph is formed by replacing each node in the base graph with n fully connected nodes, while forming the original connections using the first and last added nodes.

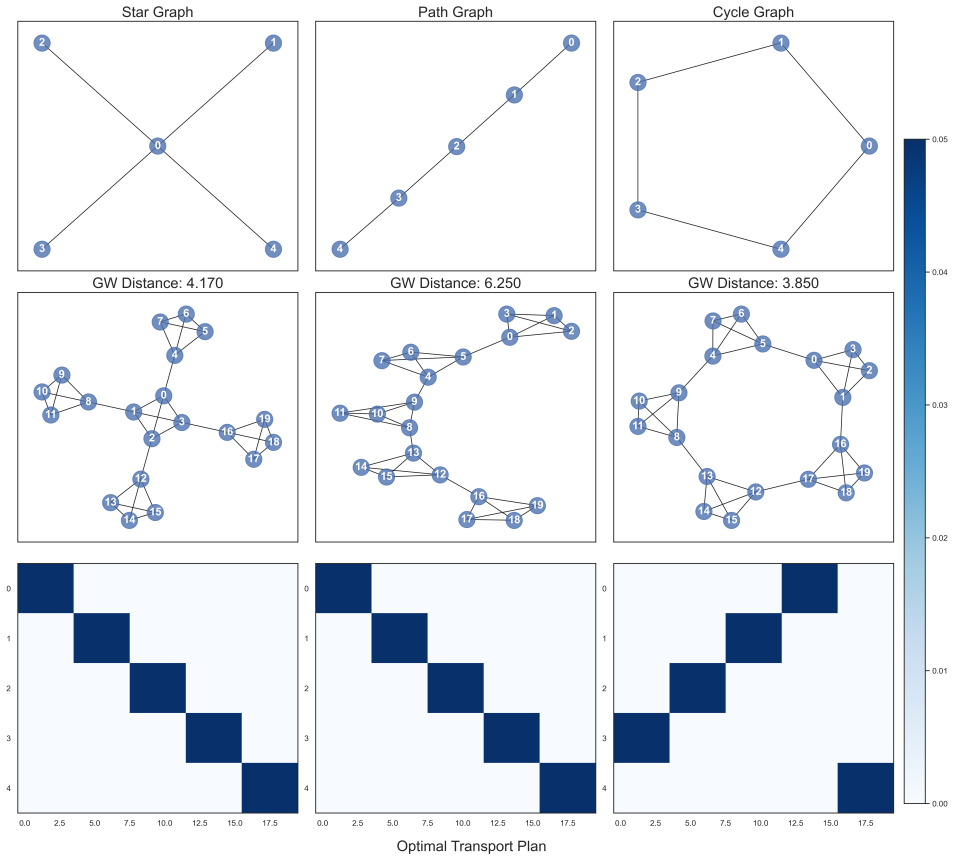

Figure 8: A comparison of star, path, and cycle graphs (top) with their corresponding block graphs with a block size of 4 (middle). The optimal transport plans show that the nodes in the base graphs are transported to their corresponding blocks in the block graphs (bottom).

A.4 Graph reconstruction from GW barycenters

Using the Python Optimal Transport (POT) library, we can compute the GW barycenter using the algorithm proposed in (Peyré et al., 2016). Given a list of N graphs $(G_i)_{i=1}^N$, we first configure them into their corresponding measure networks $(\mathcal{M}_{G_i})_{i=1}^N$. We also specify a weight vector $\lambda = (\lambda_i)_{i=1}^n$ such that $\lambda \in \boldsymbol{\Delta}(N)$ where $\boldsymbol{\Delta}(N)$ is the probability simplex of dimension N. Along with a specified dimension n_b, we can compute the GW barycenter $B \in \mathbb{R}^{n_b \times n_b}$ using the `ot.gromov.gromov_barycenters` function in POT. Under these configurations, one can interpret B as the $w_{G'}$ for a new graph G'.

In the case that $(\mathcal{M}_{G_i})_{i=1}^N$ is configured with a *adjacency* edge weight, the GW barycenter can be interpreted as an adjacency matrix. However, B contains values in the $\mathbb{R}$ space, and we must first map it into binary, i.e., all entries are either 0 (no edge) or 1 (has an edge) to convert this barycenter into a proper adjacency matrix. In this work, we set thresholds to produce graphs as seen in Fig. 2. Our work leaves room for future explorations into graph reconstruction from GW barycenters.

A.5 Most similar food webs by d_{gw}

Beyond examining the two most similar food webs under the d_{gw}, we also present the top-9 most similar networks in Table 1. From this table, we observe that many of the most similar communities are from the tropical and subtropical grasslands and savannas biome. We also observe 3 pairs of communities from different biomes and present their OT plan in Fig. 5.

Site 1	Site 2	d_{gw}
Pama	Singou	1414.01
Singou	Tamou	1414.02
Pama	Tamou	1886.60
Air and Tenere	Ouadi Rime Ouadi Achim	1886.60
Mana Pools	Vwaza Marsh	2595.25
Lolldaiga	Mago	2772.26
Banhine	Mudumu	2865.89
Mudumu	Okavango Delta	3726.56
Kourtiagou	Madjoari	3773.17

Table 1: Top 9 most similar food webs under d_{gw} with uniform and shortest path configuration. All regions are from the tropical and subtropical grasslands, savannas, and shrublands biome with the exception of Air and Tenere, Lolldaiga, Vwaza Marsh, and Okavango Delta. In particular, Air and Tenere is from the desert biome; Lolldaiga is from the tropical moist broadleaf forest, Vwaza Marsh from montane grasslands and savannas; and Okavango Delta is from the flooded grasslands and savannas biome.

A.6 Pairwise d_{gw} among sub-saharan mammal food webs

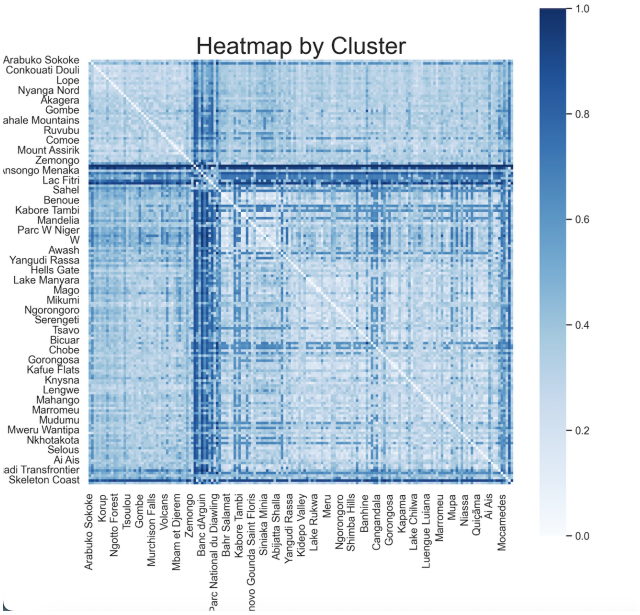

Figure 9: Heatmap of pairwise d_{gw} ordered by cluster groups obtained from Sorensen-Dice coefficient, a measure of species similarity between sites.

In this section, we show the pairwise d_{gw} heatmap with rows and columns grouped by clusters computed using the Sorensen-Dice coefficient on species similarity between sites as in (Looman and Campbell, 1960).

$$SD(G_1, G_2) := \frac{2|V_1 \cap V_2|}{|V_1| + |V_2|} \quad (2)$$

The Sorensen-Dice coefficient between graphs G_1 and G_2 is a measure of shared species, i.e. the node sets V_1 and V_2. The clusters are then obtained using the `recluster` package in R, which creates consensus trees in a resampling approach. Specifically, we used the function `recluster.multi` and the unweighted pair group method with arithmetic mean (UPMGA) and 1000 bootstrap iterations. UPMGA has been shown to work well across taxonomic groups in (Kreft and Jetz, 2010). We examined our clusters created with 94% of dissimilarity explained and adjusted according to the following guidelines 1) Clusters could not be nested within each other. 2) Sites unassigned to a cluster were assigned to the closest geographic cluster. 3) Sites that did not satisfy above conditions were excluded (N=2).

In Fig. 9, we identify lightened blocks on the heatmap diagonals. This indicates partial correspondence between food web topology captured by the d_{gw} and shared species presence. For instance, a bright region in the middle corresponds to the d_{gw} distance between Abijata-Shalla, Parc W Niger, and Yangudi Rassa National Parks. These are all sites of the savanna and grassland regions. On the contrary, a set of mammal communities vastly differs from others (indicated by the dark blue shaded "cross" in the middle of the heatmap). Sites in this region includes the Banc d'Arguin and Lac Fitri national parks, which are coastal land and landmass within a freshwater lake respectively. Our heatmap suggests that the topological structure of food webs, given species similarity, are more consistent for regions with diverse species than those with less. This is consistent with the ecological understanding that removal of species from a species-rich biome typically inflict less impact due to functional redundancy than those with fewer species.

Next, we utilized MDS to visualize the 170 mammal foods in 2-D space. We observe no obvious clusters from our MDS visualization of the mammal food web pairwise d_{gw} in Fig. 10. However, we notice that almost all communities in the tropical and subtropical moist broadleaf forests are clustered together. On the contrary, desert biome graphs are scattered around the outskirts of the produced cluster. This is consistent with the observed uniqueness of the desert biome induced by the scarcity of resources (Brown et al., 1979; Chacoff et al., 2012). It also suggests that the topological structure between food webs of regions with diverse species may be more alike, whereas those with lower species varieties tend to vary more.

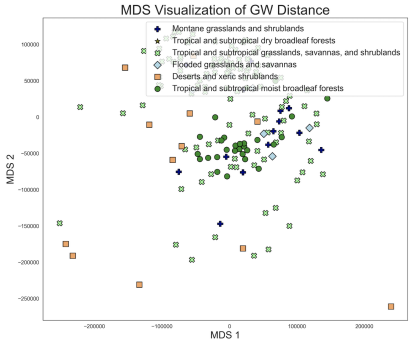

Figure 10: Multi-dimensional scaling visualization of pairwise d_{gw} labeled by biome types.

A.7 Gromov-Wasserstein Barycenter & Graph Factorization

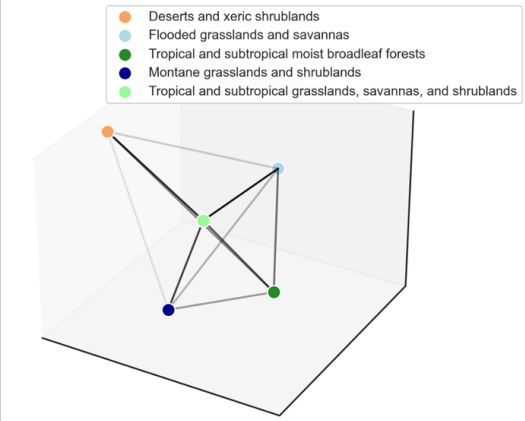

Figure 11: Barycenter per biome visualized in 3D space. Each point represents the GW barycenter of a biome, with the transparency of an edge denoting the d_{gw} between two barycenters (lighter is closer).

Lastly, we apply the GW factorization (GWF) algorithm from (Xu et al., 2022) on our measure networks with the number of graph atoms set to 80 and the size of the atom generated uniformly over the range of the node distribution from the food webs themselves. In addition to the graph atoms, the algorithm also learns an 80-dimensional embedding for every ecological network, corresponding to their weight to each of the 80 graph atoms. In Fig. 12, we plot the first 2 components TSNE learned from the atom weight embeddings with sampled graph atoms across the ecological networks.

Beyond analysis of pairwise d_{gw}, we computed the GW barycenter of mammal communities per biome type shown in Fig. 11. In this plot, we observe that the barycenter of tropical and subtropical grasslands, savannas, and shrublands lie in the center and dominates the positions of the other biome barycenters. This is consistent with the relatively large sample of mammal communities of this biome type in our dataset and their scattered positioning in Fig. 10 and Fig. 12 respectively. Furthermore, this plot shows that the desert biome barycenter is positioned farthest away from the other tropical biomes. We omit the tropical and subtropical dry broadleaf forest in this analysis since there are very few food webs of this type present in our dataset, resulting in a non-informative barycenter.

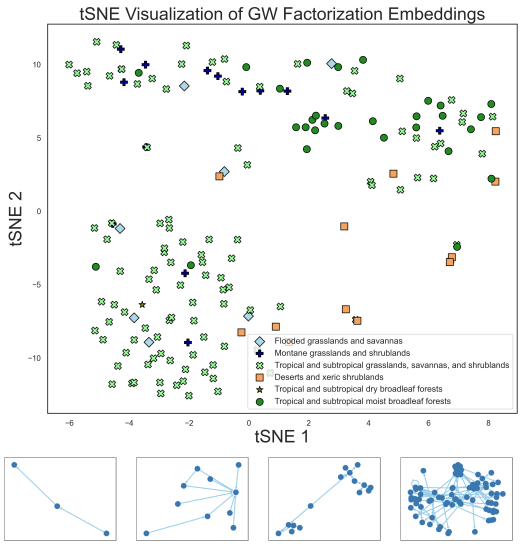

Figure 12: (Top) T-SNE visualization of GWF embeddings labeled by biome type. (Bottom) Top-4 significant graph atoms.