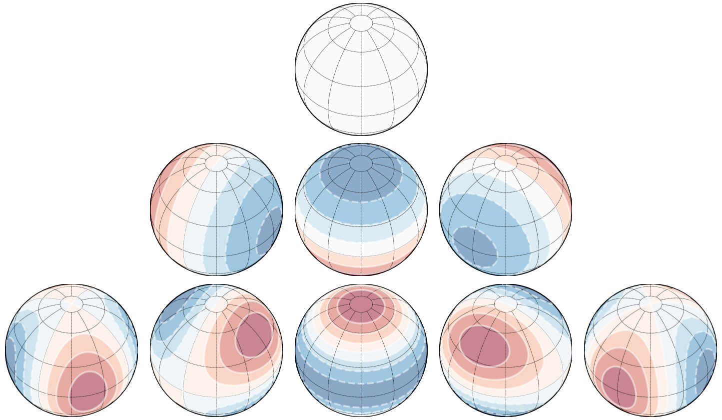
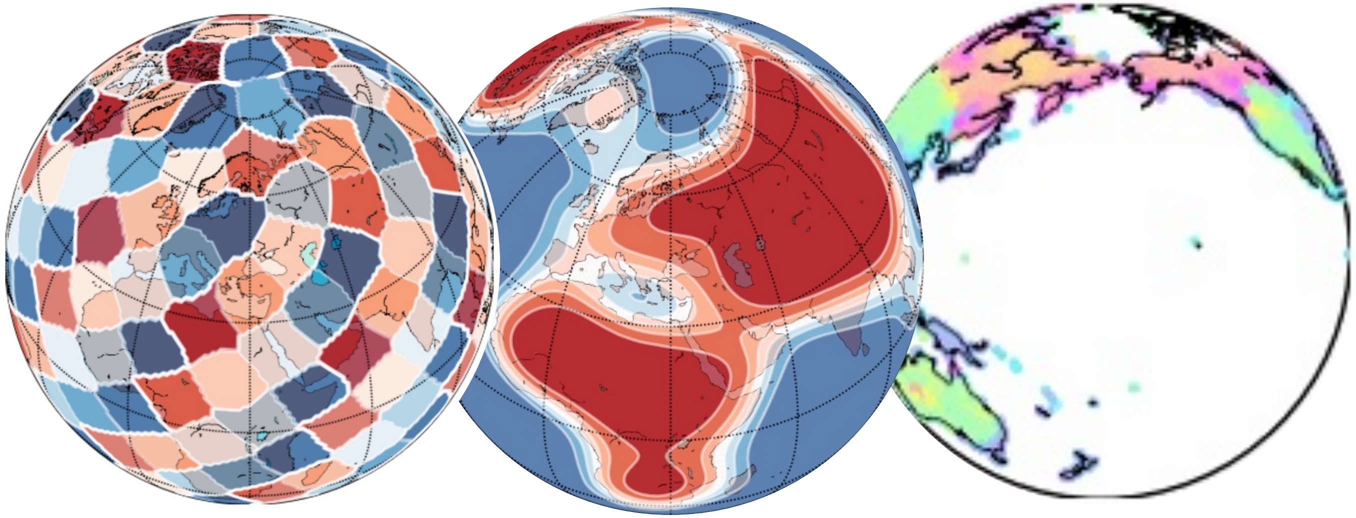
Geographic Location Encoding with Spherical Harmonics and Sinusoidal Representation Networks (Siren)

Marc Rußwurm, Konstantin Klemmer,
Esther Rolf, Robin Zbinden, Devis Tuia

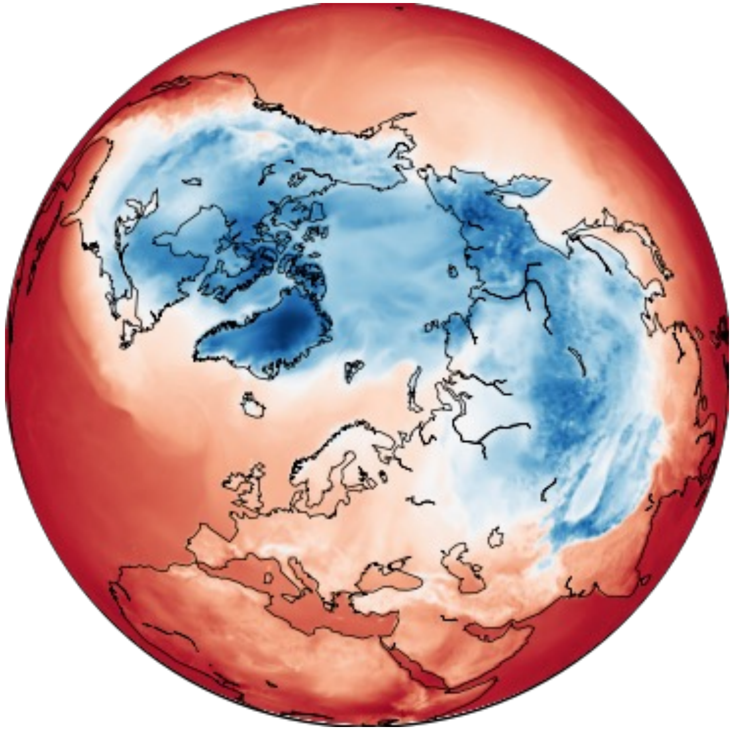




Research Fields using Geolocated Data

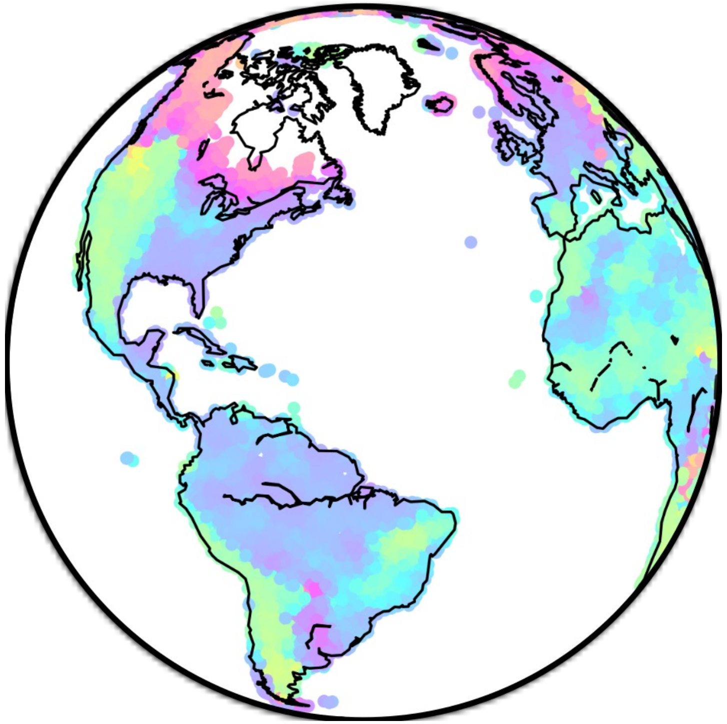
Compression of Weather Data

e.g., Huang & Hoefler (2023)



Encoding visual similarities by image geolocation

Klemmer et al., 2024, Cepeda et al., (2023).



Huang, L., & Hoefler, T. (2023). Compressing multidimensional weather and climate data into neural networks. ICLR 2023

Cepeda et al., (2023). "GeoCLIP: Clip-Inspired Alignment between Locations". NeurIPS 2023.

Klemmer, et al. (2024) "SatCLIP: Global, general-purpose location embeddings with satellite imagery."

Location Encoding in Species Identification

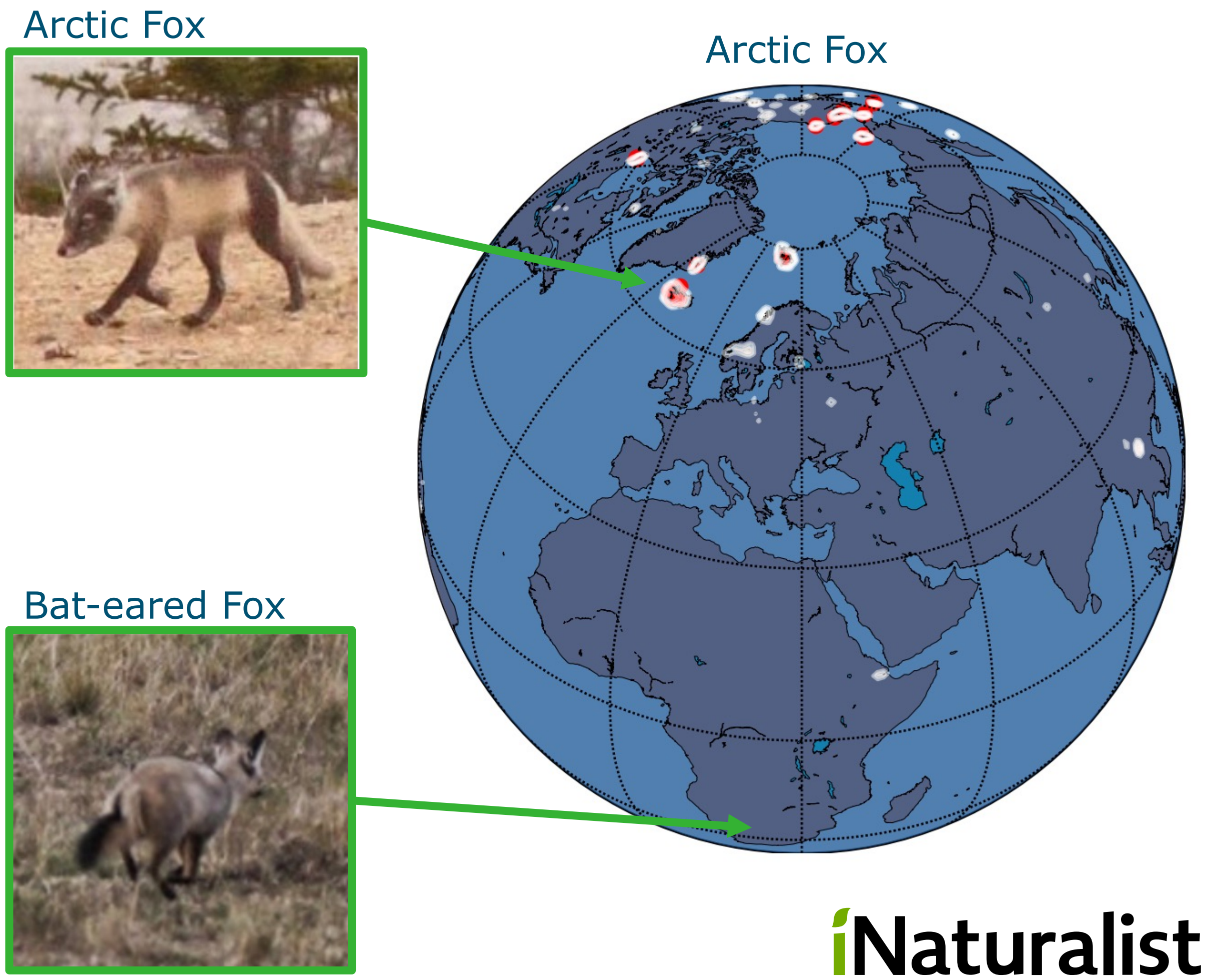
Image-only classification:

60% top-1 accuracy

with additional location information:

+9-13% top-1 accuracy

see Cole et al., 2023 specifically for species mapping with location encoders.



iNaturalist

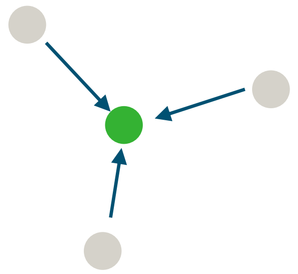
Figure inspired by Mai, et al., (2023)

Cole et al., (2023). Spatial implicit neural representations for global-scale species mapping. In ICML 2023
Mai, et al., (2023). CSP: Self-supervised contrastive spatial pre-training for geospatial-visual representations. In ICML 2023

Implicit Neural (Geo)Representations

Geospatial Interpolation Approaches

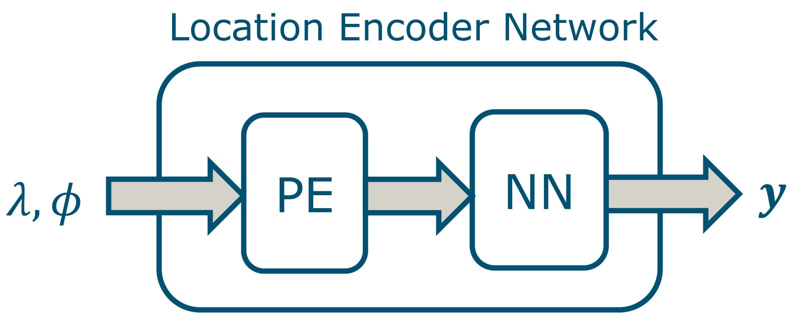
e.g., Kriging [Matheron, 1969]



scale poorly with large datasets $O(N^2)$ with N support (training) locations

Location Encoders as Implicit Neural Representations

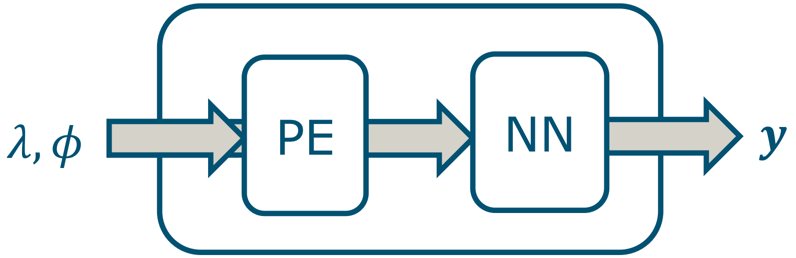
Summarized by Mai et al., 2023:



after being fitted on training data, they predict large datasets efficiently $O(N)$

Matheron, 1969. Le krigeage universel (Universal kriging)
Mai, G., et al., (2023). On the opportunities and challenges of foundation models for geospatial artificial intelligence.

Related Work: Location Encoders



Positional Embedding PE

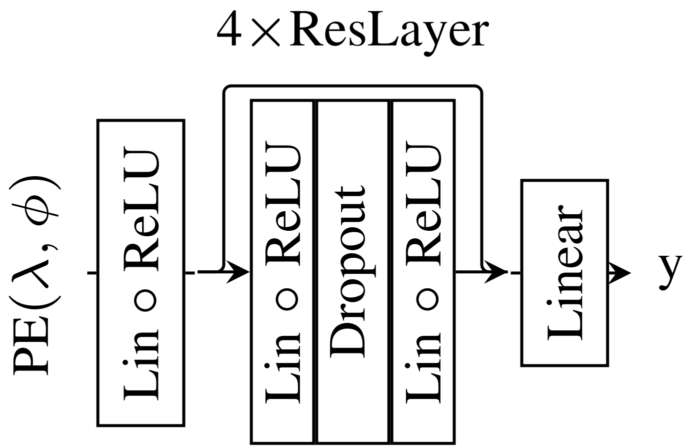
Sine-cosine transformations across scales similar to Fourier series.

Research on embeddings:

- Wrap [Aodha et al., 2019]
- Grid [Mai et al., 2020]
- Theory [Mai et al., 2020]
- SphereC [Mai et al., 2023]
- SphereM [Mai et al., 2023]

Neural Network NN

FcNet introduced by Aodha et al. (2019) and re-used in follow-up works

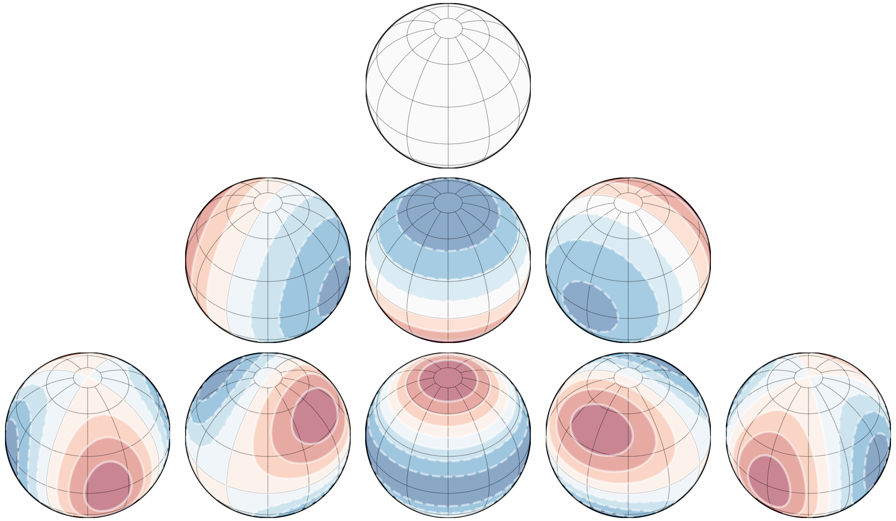


Mac Aodha, et al, (2019). Presence-only geographical priors for fine-grained image classification. In CVPR 2019.
Mai, G. et al., (2020) Space2Vec: Multi-Scale Representation Learning. In ICLR 2020
Mai, G. et al., (2023). Sphere2Vec: A general-purpose location representation learning. ISPRS Journal.

Spherical Harmonics & Siren – Our Proposition

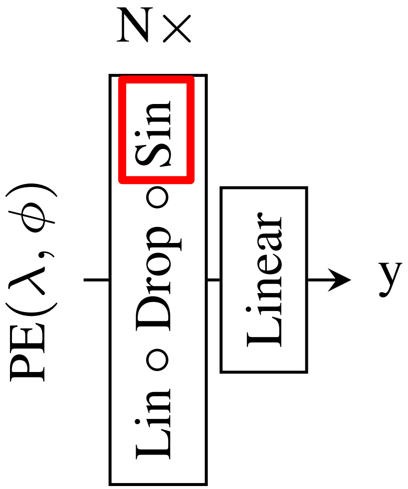
Positional Embedding PE

Spherical Harminic Basis Functions



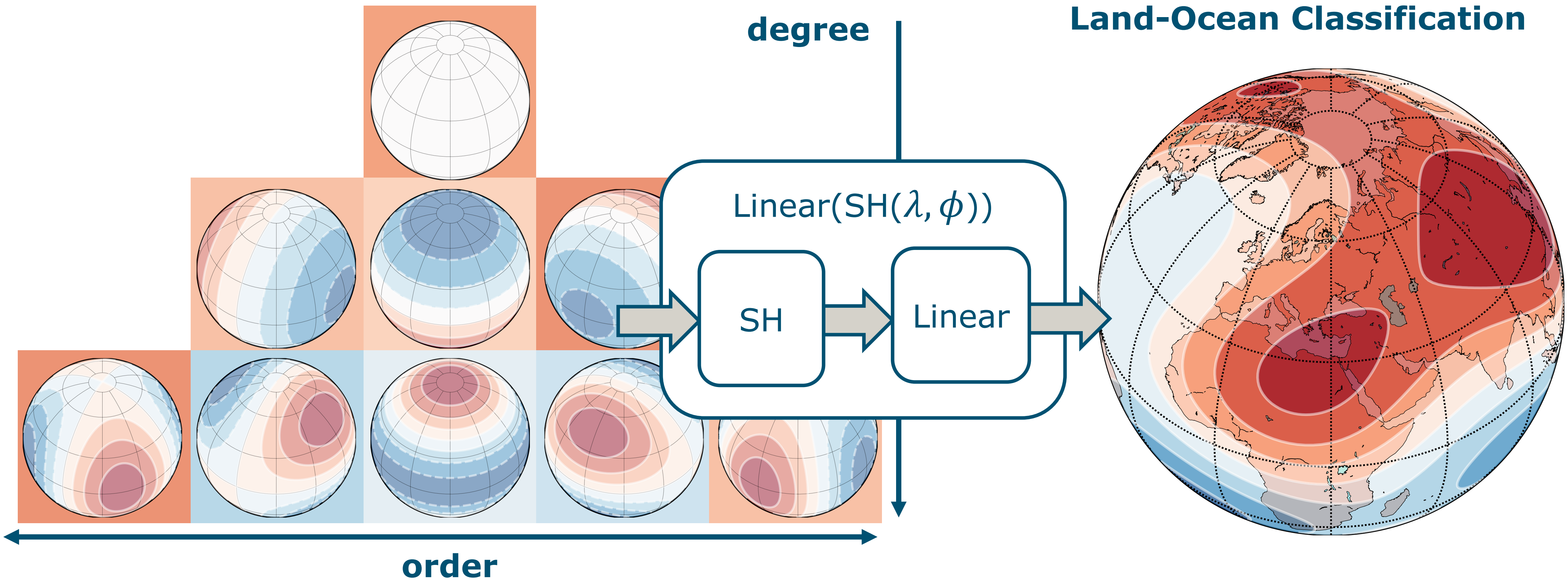
Neural Network NN

Sinusoidal Representation Networks (Siren) [Sitzmann et al., 2020]

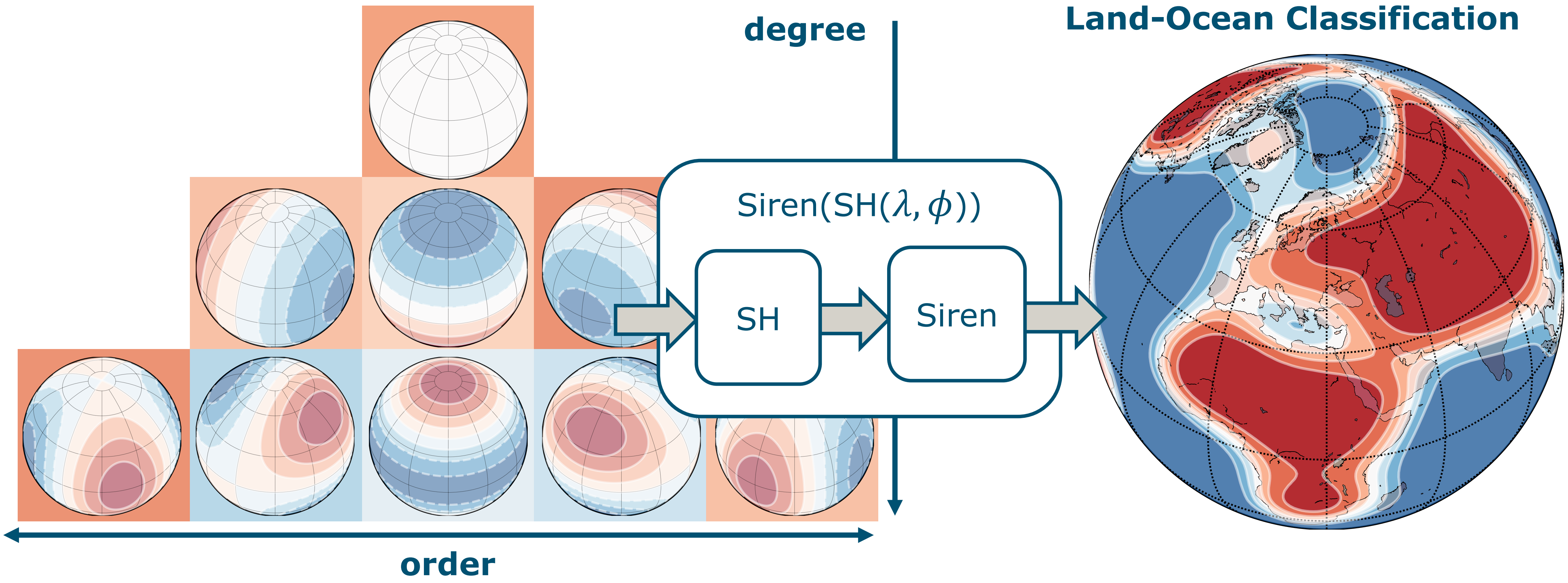


Sitzmann et al., (2020). Implicit neural representations with periodic activation functions. NeurIPS 2020

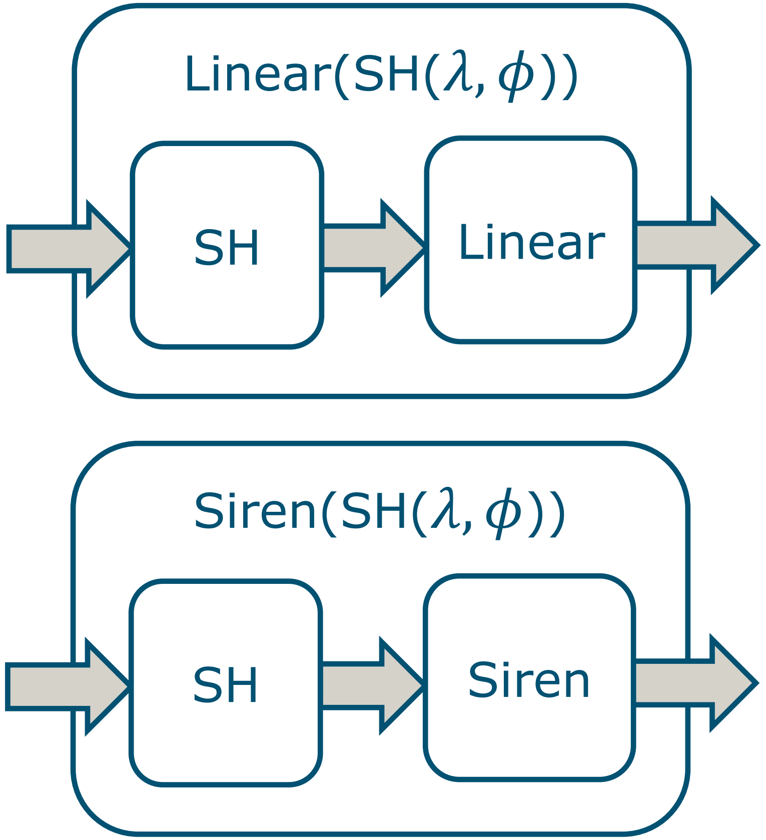
Spherical Harmonic (SH) Basis Functions

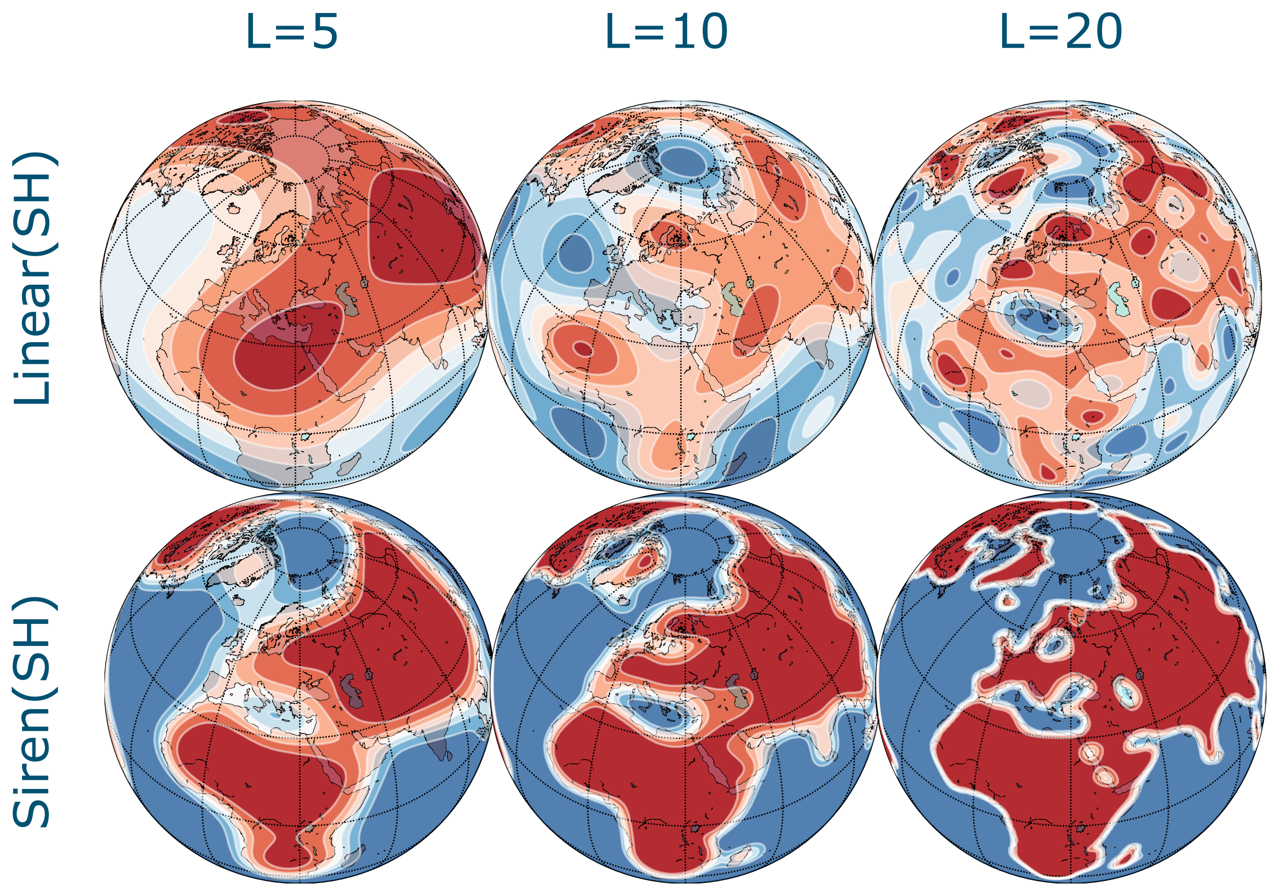


Spherical Harmonics & Siren – Our Proposition

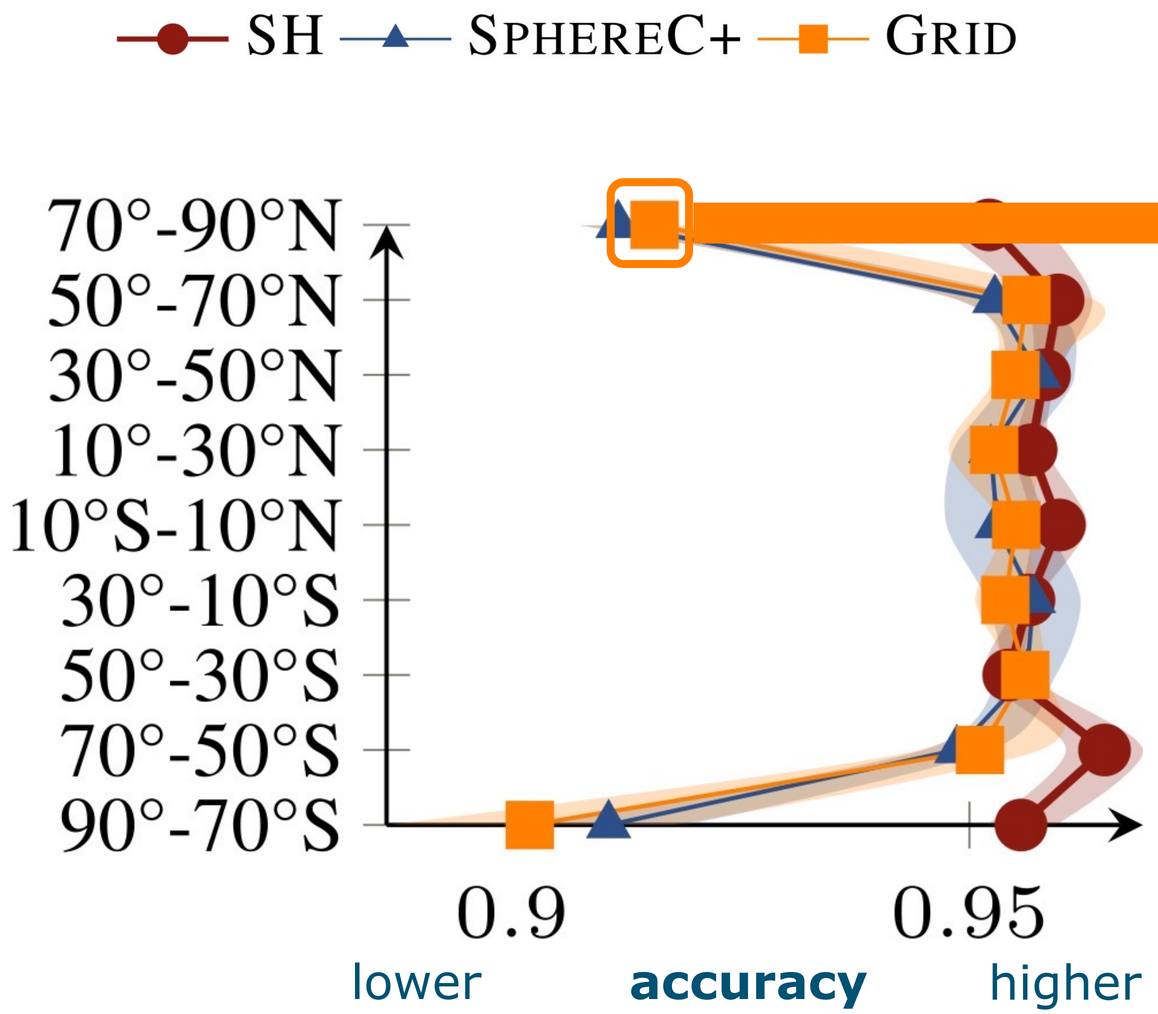


Spherical Harmonics (SH) with a Linear Layer and Siren

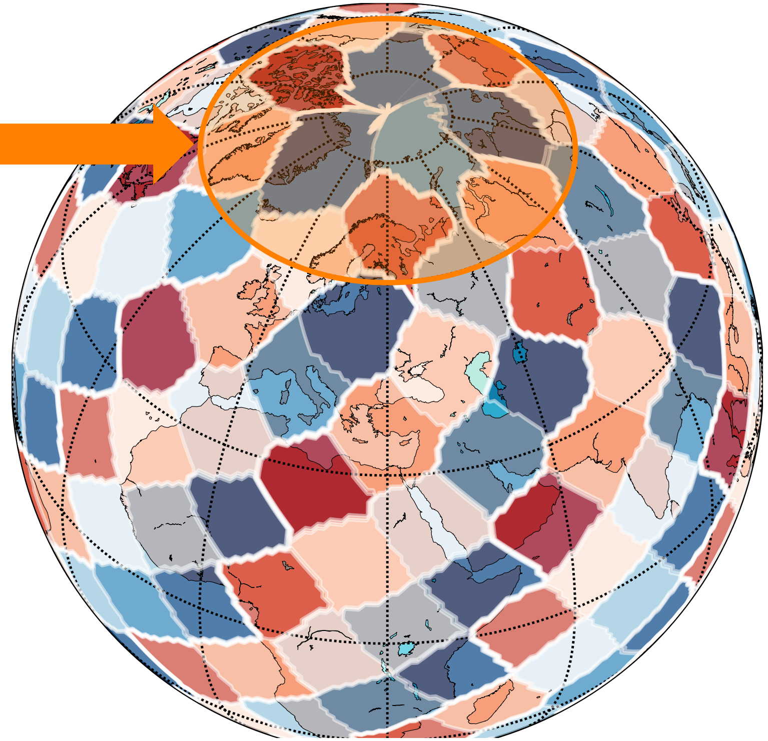




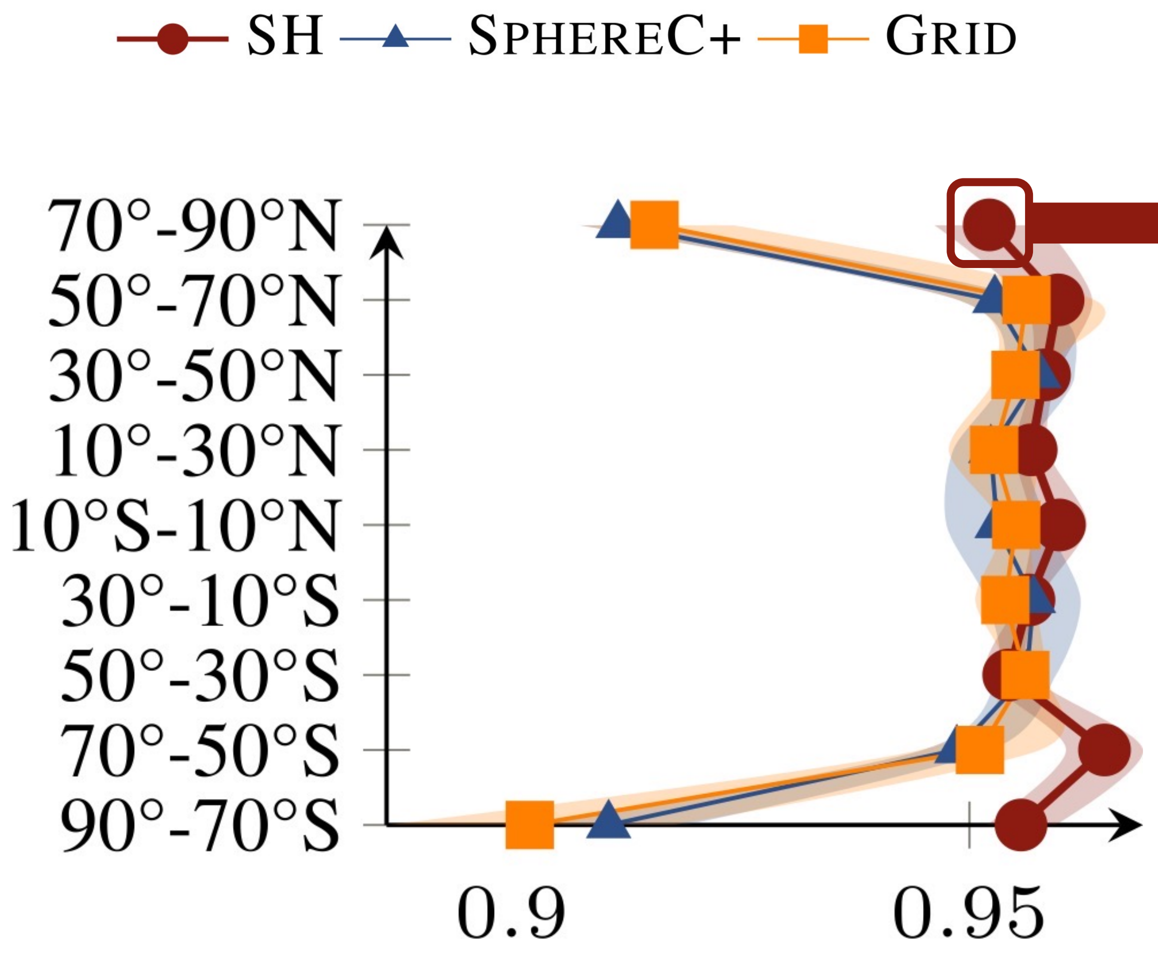
Longitudinal Accuracy – Checkerboard Classification



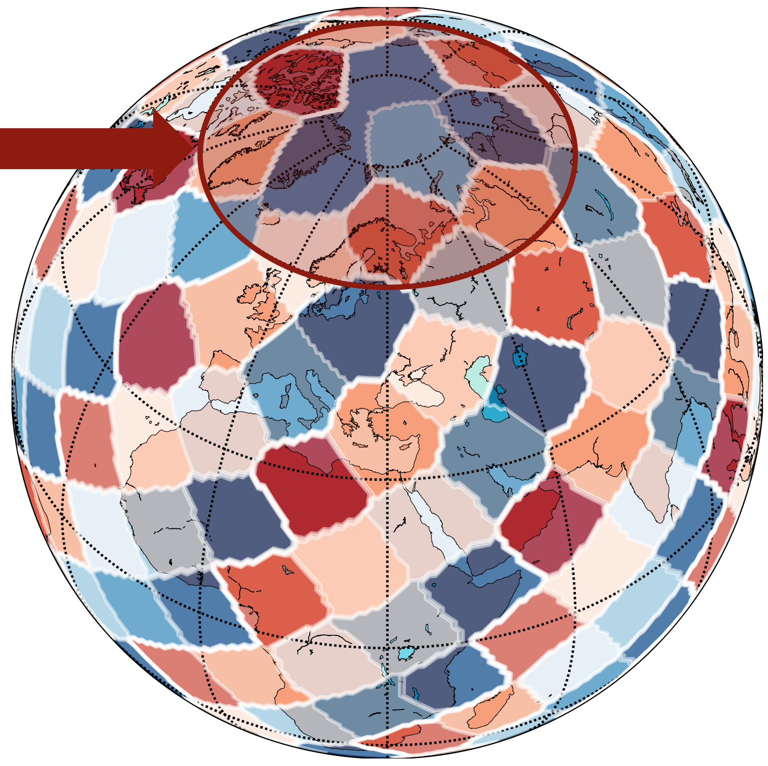
Siren(Grid)



Longitudinal Accuracy – Checkerboard Classification



Siren(SH)



Quantitative Results

Land-Ocean classification accuracy

PE ↓ NN →	Linear	FcNet	SirenNet
Direct	71.4 ± 0.0	90.3 ± 0.7	95.1 ± 0.3
Cartesian3D	70.5 ± 3.5	92.7 ± 0.3	92.8 ± 0.3
Wrap	74.4 ± 0.3	93.2 ± 0.3	95.2 ± 0.2
Grid	81.7 ± 0.1	95.1 ± 0.1	95.5 ± 0.2
Theory	86.9 ± 0.1	94.9 ± 0.2	95.5 ± 0.1
SphereC	79.6 ± 0.2	95.0 ± 0.3	95.2 ± 0.1
SphereC+	84.6 ± 0.2	95.3 ± 0.1	95.5 ± 0.1
SphereM	74.0 ± 0.0	89.1 ± 0.1	88.3 ± 0.4
SphereM+	81.9 ± 0.2	92.1 ± 0.3	93.7 ± 0.1
SH (ours)	$\underline{94.4 \pm 0.1}$	$\mathbf{95.9 \pm 0.1}$	$\underline{95.8 \pm 0.1}$

iNaturalist 2018 accuracy improvement

PE ↓ NN →	Linear	FcNet	SirenNet
Direct	-5.9 ± 0.1	$+9.3 \pm 0.3$	$+12.1 \pm 0.1$
Cartesian3D	$+0.8 \pm 0.2$	$+11.8 \pm 0.1$	$+12.0 \pm 0.1$
Wrap	-0.1 ± 0.1	$\underline{+12.1 \pm 0.1}$	$+12.1 \pm 0.1$
Grid	$+11.2 \pm 0.1$	$+11.8 \pm 0.2$	$+11.6 \pm 0.4$
Theory	$+11.5 \pm 0.0$	$+10.8 \pm 0.0$	$+11.4 \pm 0.1$
SphereC	$+11.2 \pm 0.1$	$+12.0 \pm 0.2$	$\mathbf{+12.3 \pm 0.1}$
SphereC+	$+11.1 \pm 0.2$	$+11.5 \pm 0.3$	$+10.3 \pm 0.4$
SphereM	$+7.2 \pm 0.2$	$+11.3 \pm 0.2$	$+10.6 \pm 0.6$
SphereM+	$+11.6 \pm 0.1$	$+12.0 \pm 0.1$	$+10.7 \pm 0.2$
SH (ours)	$+10.5 \pm 0.1$	$+12.0 \pm 0.0$	$\mathbf{+12.3 \pm 0.2}$

image-only: 59.2% top-1 accuracy with encoder NN(PE) ↑

Spherical Harmonics work well with all NNs

Siren is work well with all PEs

1-Layer Siren as learned Grid embeddings

$$\textsc{SirenNet}(\phi, \lambda) \; = \; \sin(\mathbf{W}[\phi, \lambda]^T + \mathbf{b}) = \bigcup_{h=1}^{H} [\sin(w_h^{\lambda}\lambda + w_h^{\phi}\phi + b)]$$

set $w_h^{\phi} = w_{h+1}^{\phi} = w_{h+2}^{\lambda} = w_{h+3}^{\lambda} = 0$ and $\mathrm{b}_{\mathrm{h+1}} = \mathrm{b}_{\mathrm{h+3}} = 0$

$$\textsc{GridSiren}(\lambda, \phi) = \bigcup_{h=0,4,\ldots}^{H-1} [\sin(w_h^{\lambda}\lambda + b_h), \sin(w_{h+1}^{\lambda}\lambda), \sin(w_{h+2}^{\phi}\phi + b_{h+2}), \sin(w_{h+3}^{\phi}\phi)]$$

both are equivalent if

1. $b_h = b_{h+2} = \frac{\pi}{2}$
2. $w_{h+1}^{\lambda} = w_{h+1}^{\lambda} = w_{h+2}^{\phi} = w_{h+3}^{\phi} = \frac{1}{\alpha_s}$

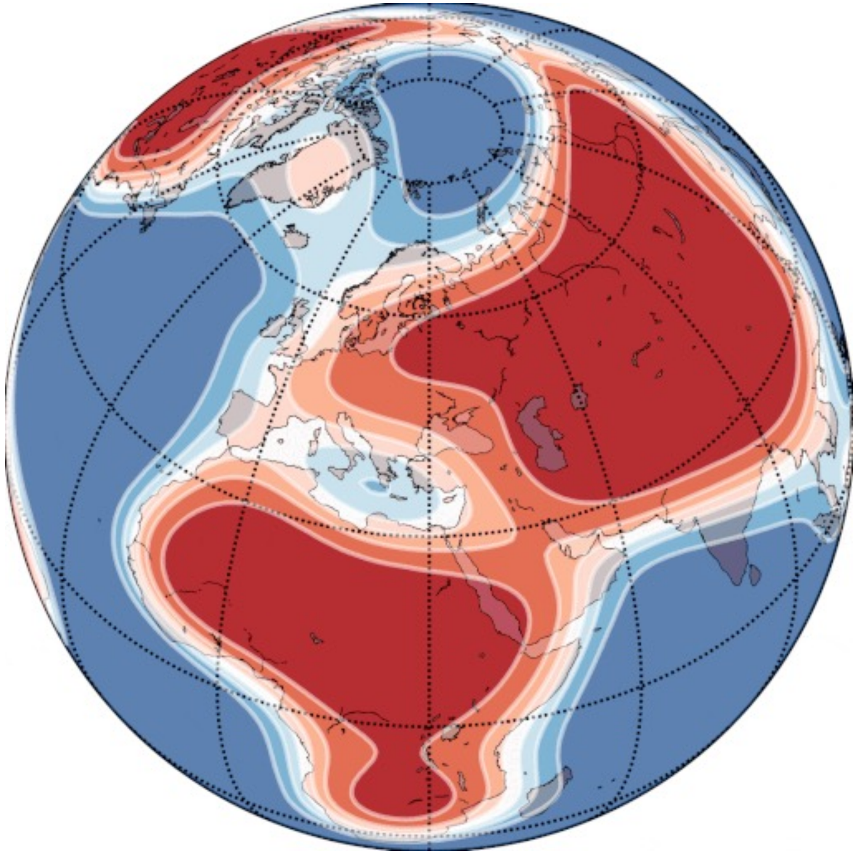
$$\textsc{Grid}(\lambda, \phi) = \bigcup_{s=0}^{S-1} \left[\sin(\frac{\lambda}{\alpha_s} + \frac{\pi}{2}), \sin(\frac{\lambda}{\alpha_s}), \sin(\frac{\phi}{\alpha_s} + \frac{\pi}{2}), \sin(\frac{\phi}{\alpha_s})\right]$$

Conclusions & Takeaways

Geographic Location Encoding with Spherical Harmonics and Sinusoidal Representation Networks

We can recommend:

1. Siren as Neural Network for **any location encoding problem** and
2. Spherical Harmonic basis functions for **global geographic problems** where the spherical geometry matters



Code: github.com/marccoru/locationencoder
Contact: marc.russwurm@wur.nl

WAGENINGEN UNIVERSITY & RESEARCH HARVARD UNIVERSITY EPFL Microsoft