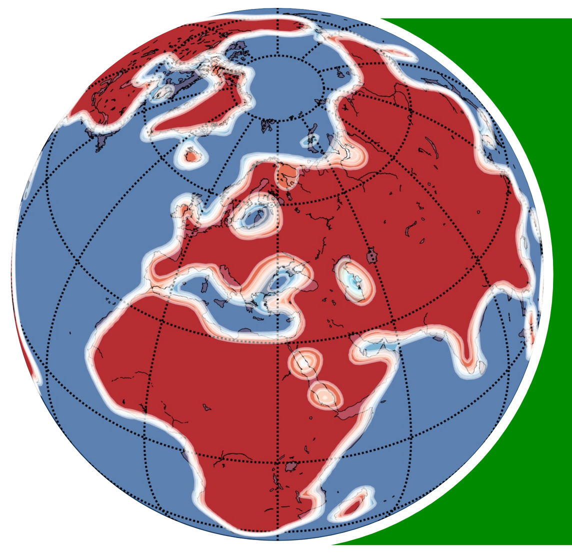
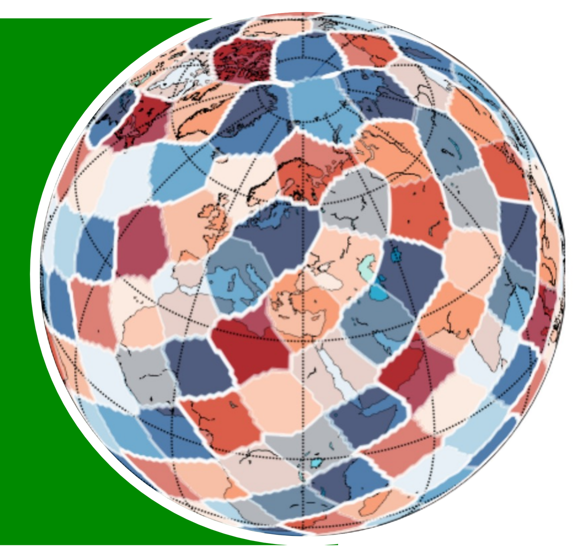


Geographic Location Encoding with Spherical Harmonics and Sinusoidal Representation Networks (Siren)

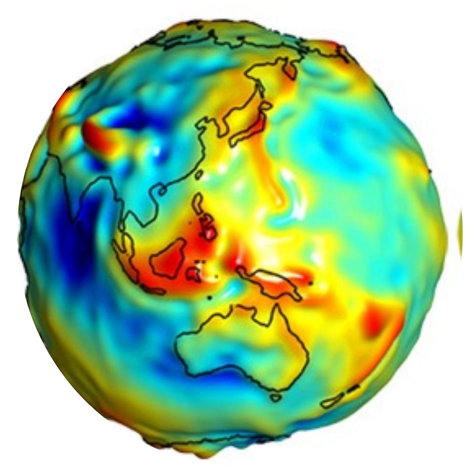


Marc Rußwurm, Konstantin Klemmer, Esther Rolf, Robin Zbinden, Devis Tuia

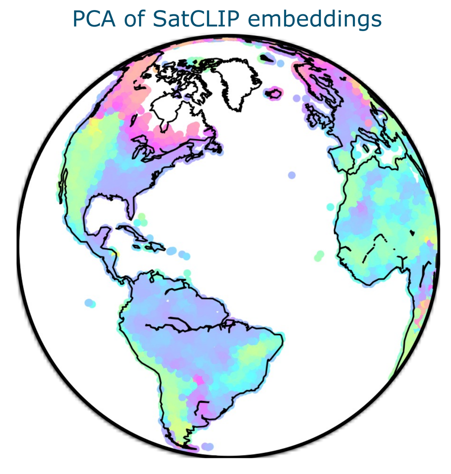
Why geographic location encoding?

Good representations of geographical spaces are important for any application that integrates geolocated data:

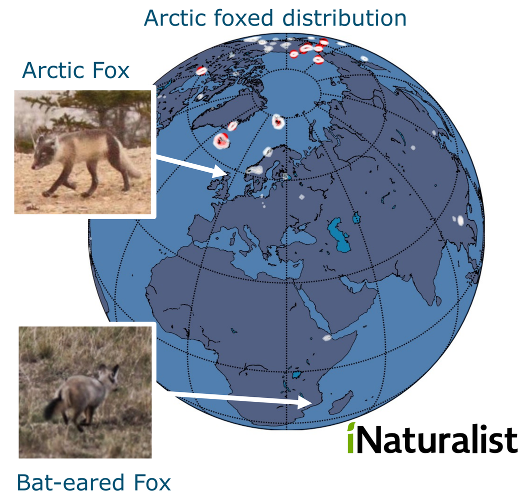
Compressing Weather Data
Huang & Hoefler, 2023 or
representing **Earth's Gravity**



Visual similarity encoding
Klemmer et al., 2024

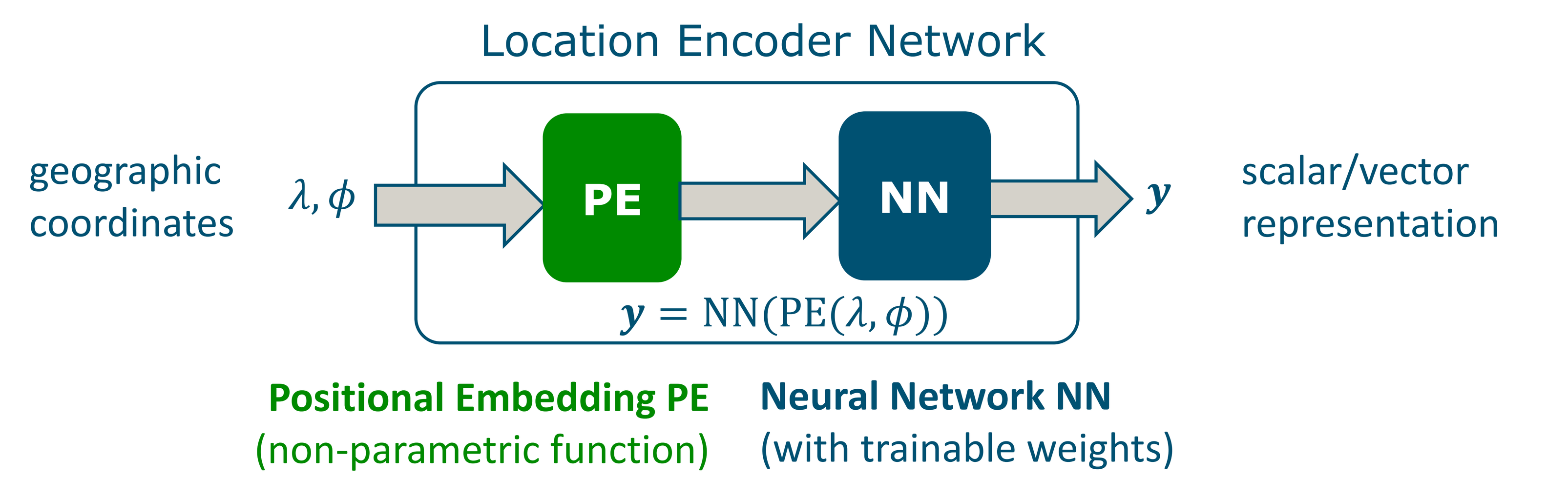


Species Mapping
Cole et al., 2023



What is location encoding?

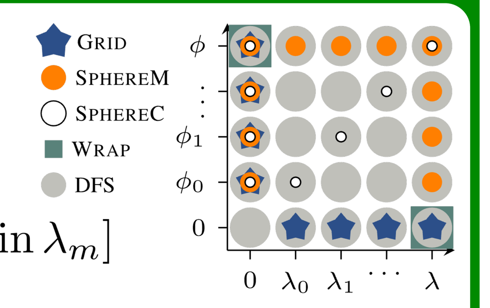
An implicit neural representation (INR) of geographic space



Related work: location encoders

Positional Embeddings: Double Fourier Sphere (DFS)

$$\mathrm{DFS}(\lambda, \phi) = \bigcup_{n=0}^{S-1} [\sin \phi_n, \cos \phi_n] \cup \bigcup_{m=0}^{S-1} [\sin \lambda_m, \cos \lambda_m] \cup \bigcup_{n=0}^{S-1} \bigcup_{m=0}^{S-1} [\cos \phi_n \cos \lambda_m, \cos \phi_n \sin \lambda_m, \sin \phi_n \cos \lambda_m, \sin \phi_m \sin \lambda_m]$$



DFS-encodings recently proposed: Wrap [Aodha et al., 2019], Grid, Theory [Mai et al., 2020], SphereC, SphereM [Mai et al., 2023]

➡ Assume rectangular (not spherical) domain of longitude and latitude

Neural Networks: ReLU Networks

The fully-connected net (FcNet) with ReLU activations by Aodha et al., 2019 is commonly used (e.g., in Cole et al., 2023)

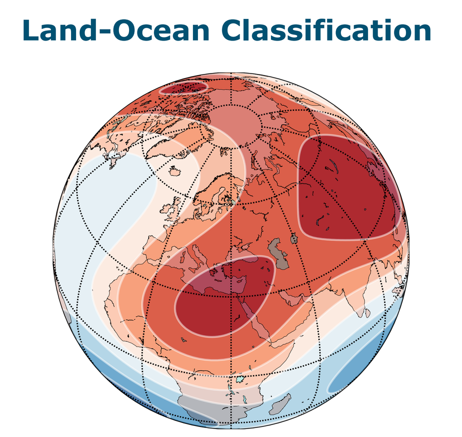
➡ ReLU networks are not optimal for INRs [Sitzmann et al., 2020]

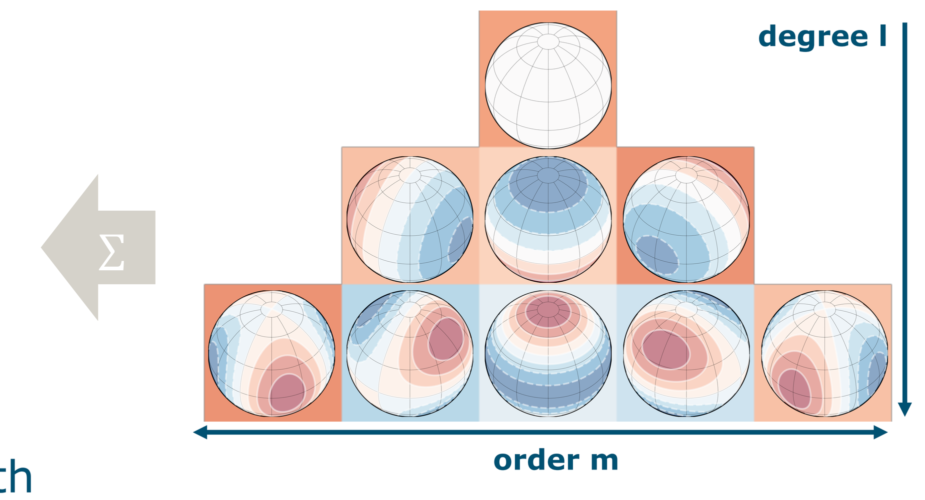
Our Proposition: $\underbrace{\mathrm{Siren}}_{\text{as NN}}(\underbrace{\mathrm{SH}}_{\text{as PE}}(\lambda, \phi))$

Positional Embedding with Spherical Harmonic functions

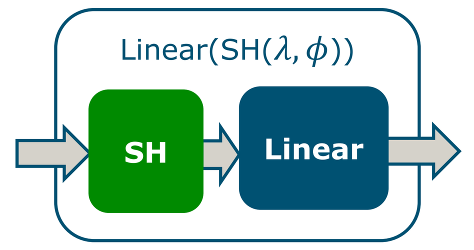
Weighted Spherical Harmonics (SH) approximate arbitrary functions on the sphere ...

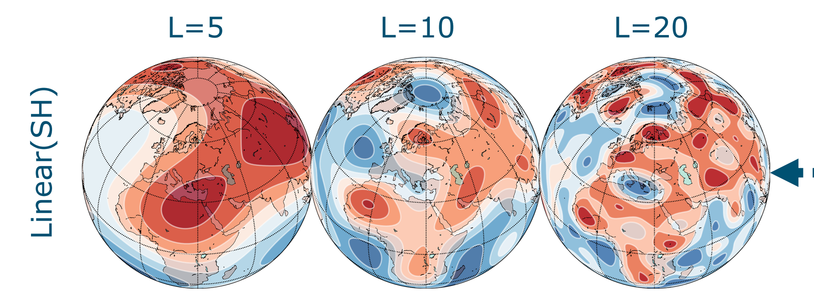
$$f(\lambda, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} w_l^m Y_l^m(\lambda, \phi)$$





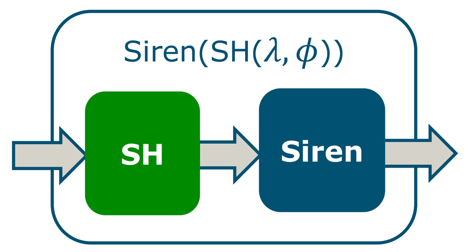
... like a location encoder with Spherical Harmonic Positional Embedding and a Linear Layer as "NN"

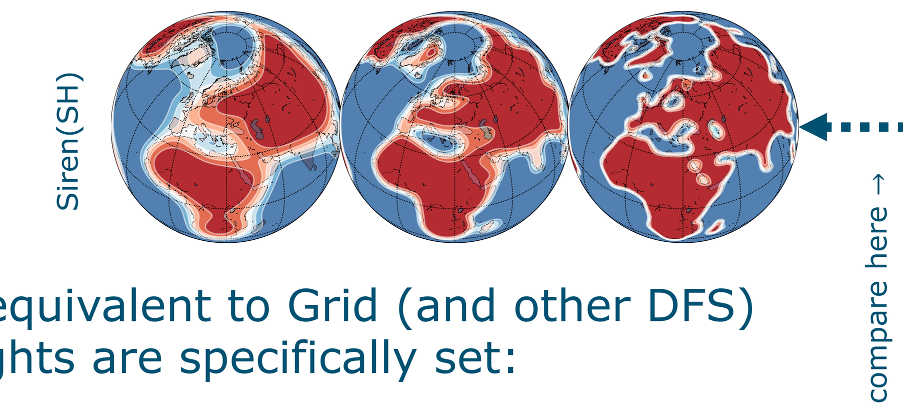




Sinusoidal Representation Networks (Siren)

with sine activation functions





We show: a 1-layer Siren is equivalent to Grid (and other DFS) embeddings when some weights are specifically set:

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

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

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

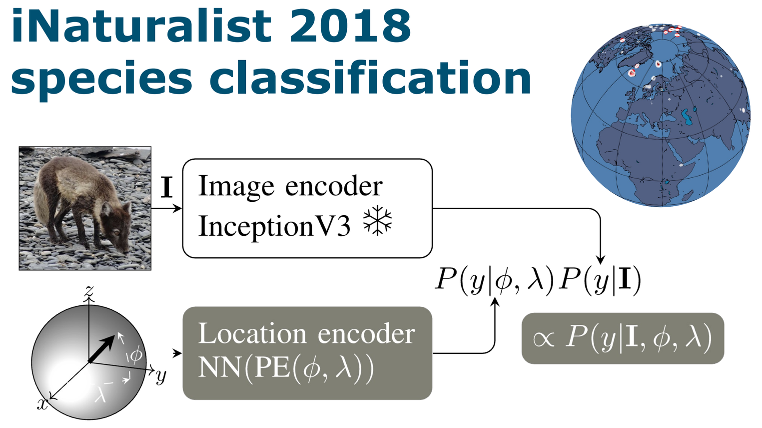
both are equivalent if
1. $b_h = b_{h+2} = \frac{\pi}{2}$
2. $w_{h+1}^\lambda = w_{h+1}^\lambda = w_{h+2}^\phi = w_{h+3}^\phi = \frac{1}{\alpha_s}$

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

Quantitative Comparisons

1. Siren works with all PEs including direct embedding
2. Spherical Harmonics works with all NNs even a linear layer

iNaturalist 2018 species classification

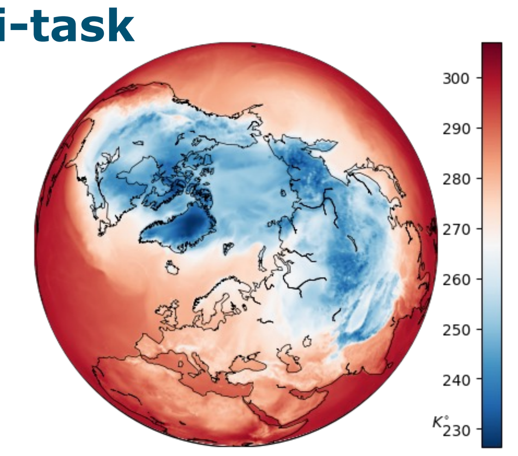


Evaluation: Accuracy increase over image-only

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) ↑

ERA5 multi-task regression of 8 climate variables



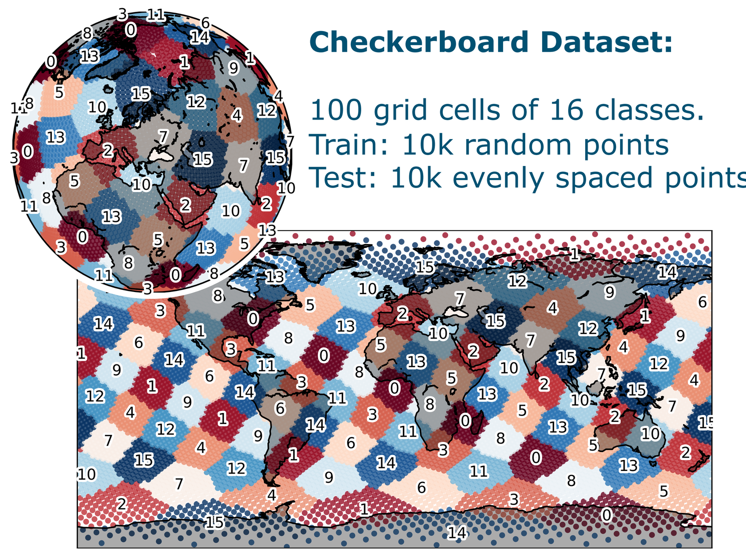
Evaluation: averaged MSE over test points

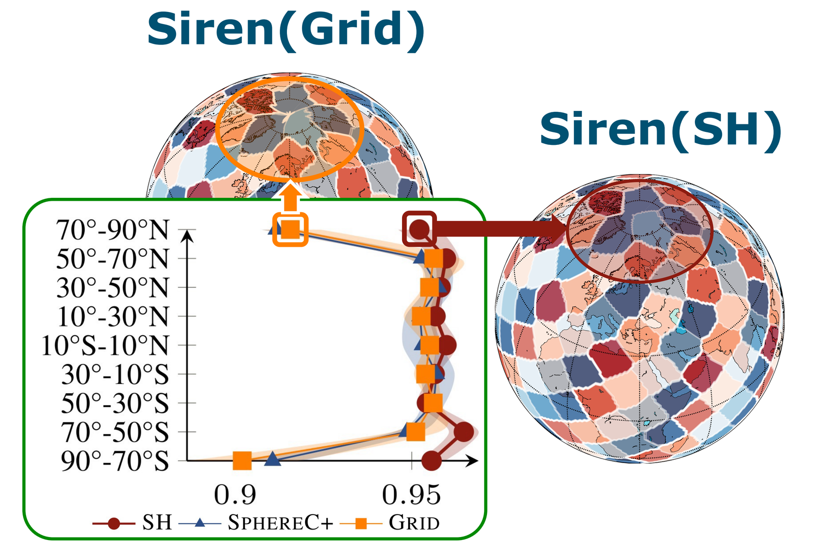
PE ↓ NN →	Linear	FcNet	SirenNet
Direct	27.19 ± 0.08	7.83 ± 1.06	1.62 ± 0.10
Cartesian3D	24.18 ± 0.00	4.23 ± 0.25	1.57 ± 0.11
Wrap	13.26 ± 0.02	4.13 ± 0.25	1.89 ± 0.07
Grid	9.83 ± 0.01	1.51 ± 0.04	2.37 ± 0.13
Theory	9.24 ± 0.01	1.61 ± 0.05	2.99 ± 0.10
SphereC	20.03 ± 0.02	1.92 ± 0.06	1.95 ± 0.09
SphereC+	8.50 ± 0.02	1.38 ± 0.03	1.97 ± 0.08
SphereM	26.68 ± 0.13	3.51 ± 0.08	5.89 ± 0.68
SphereM+	9.94 ± 0.02	1.54 ± 0.07	2.75 ± 0.09
SH (ours)	$\underline{1.39 \pm 0.02}$	$\mathbf{0.58 \pm 0.02}$	$\underline{1.19 \pm 0.04}$

Qualitative Experiments

1. Spherical Harmonics remain accurate on the poles

Checkerboard Dataset:
100 grid cells of 16 classes.
Train: 10k random points
Test: 10k evenly spaced points





2. Siren increases the resolution with fewer harmonics L

Takeaway & Conclusion

Takeaway: Spherical Harmonics provide the best representations for global-scale problems on our spherical planet.

Concretely, 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

WAGENINGEN UNIVERSITY & RESEARCH HARVARD UNIVERSITY EPFL Microsoft ICLR

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

video paper code