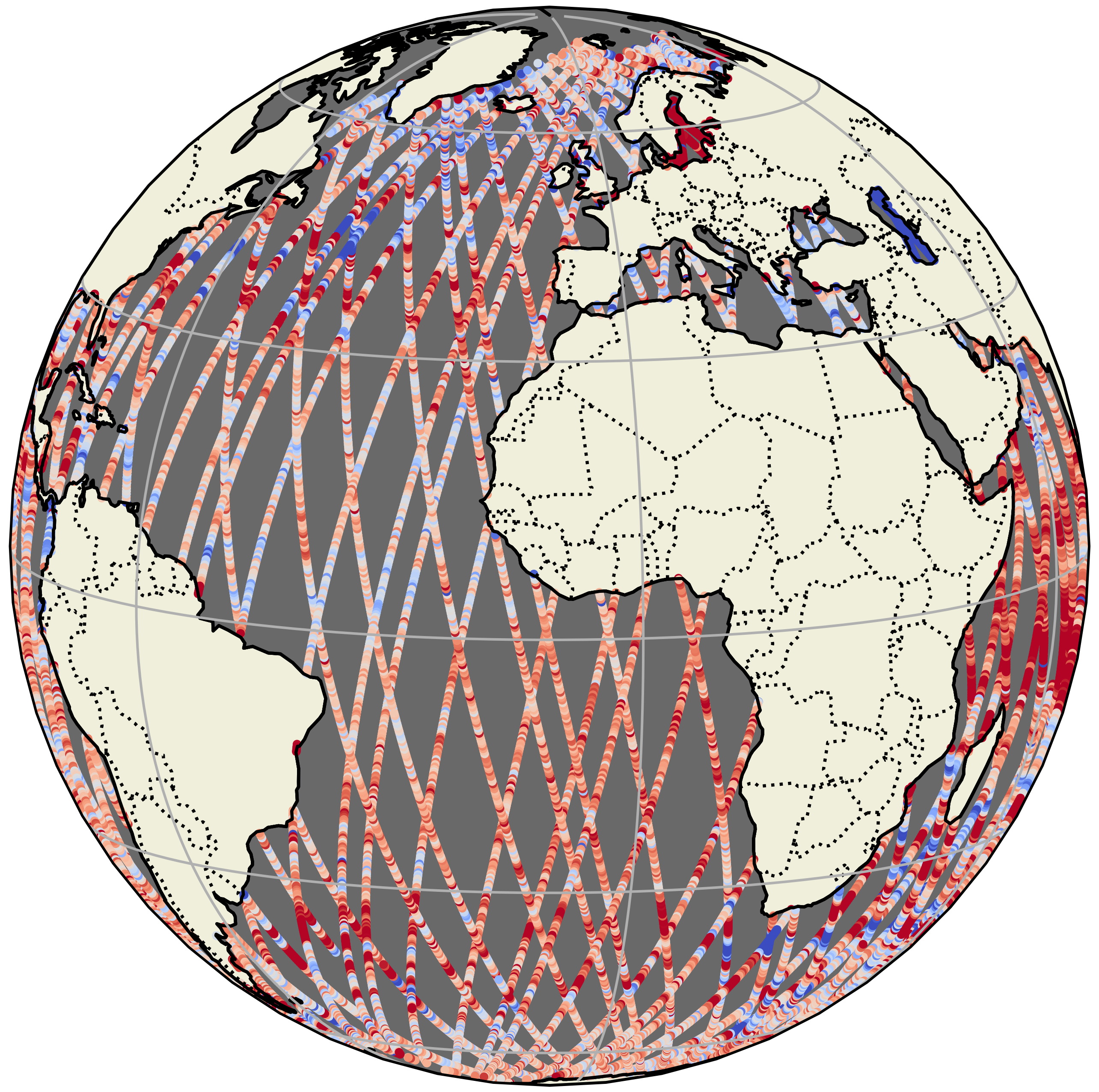
Deep Random Features for Interpolation of Spatiotemporal Data

Weibin Chen, Azhir Mahmood, Michel Tsamados, So Takao

Motivation

- Remote sensing observations are inherently sparse
- Exact Gaussian process has poor scalability
- Typical kernels are too simple

→ Struggles to learn from non-stationary fields



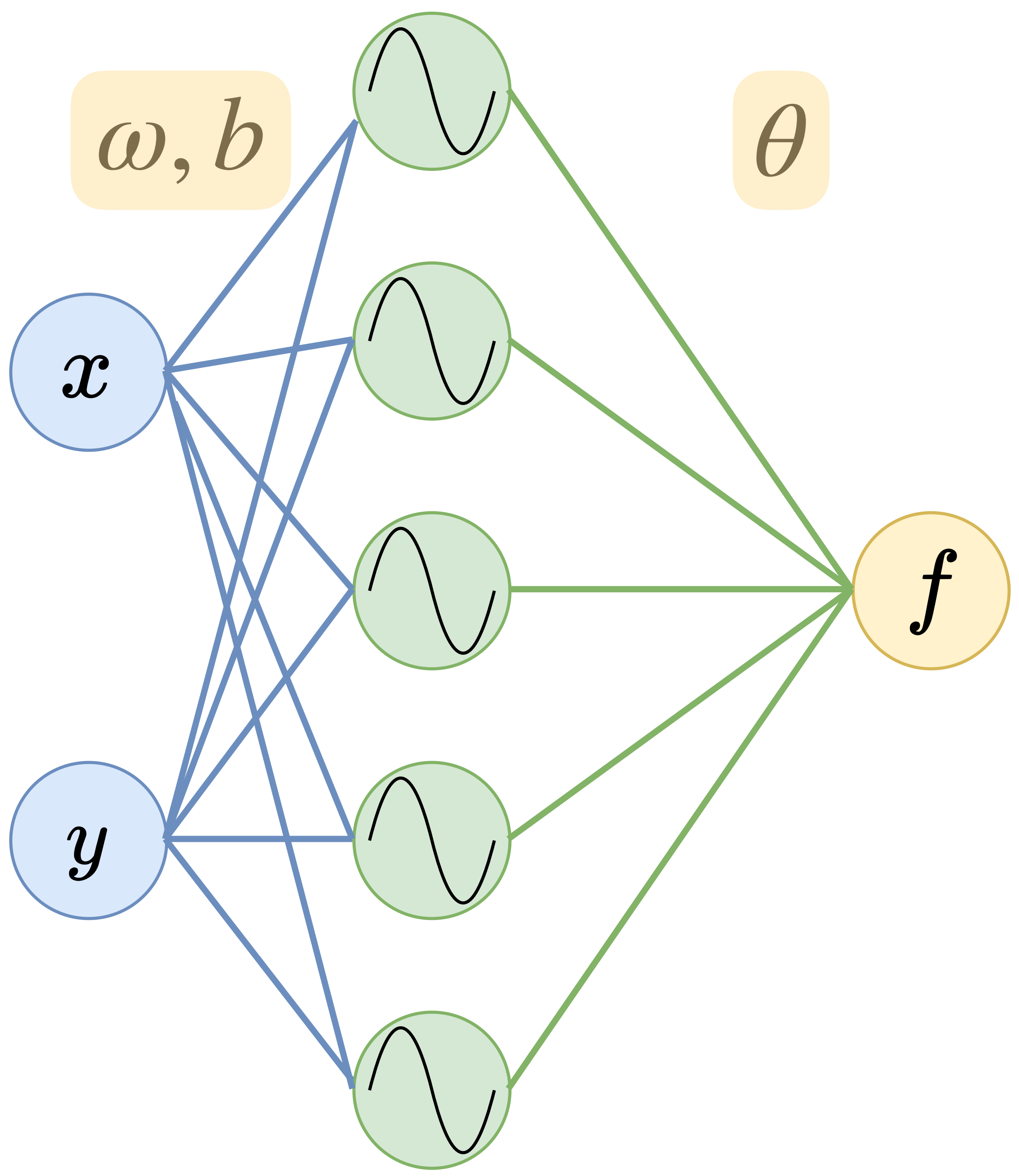
Random Fourier Features (RFFs)

- Any stationary kernel can be approximated using Random Fourier Features (RFFs)*

$$f(x) \approx \sum_{h=1}^{H} \theta_h \phi_h(x), \quad \theta_h \sim \mathcal{N}(0,1),$$

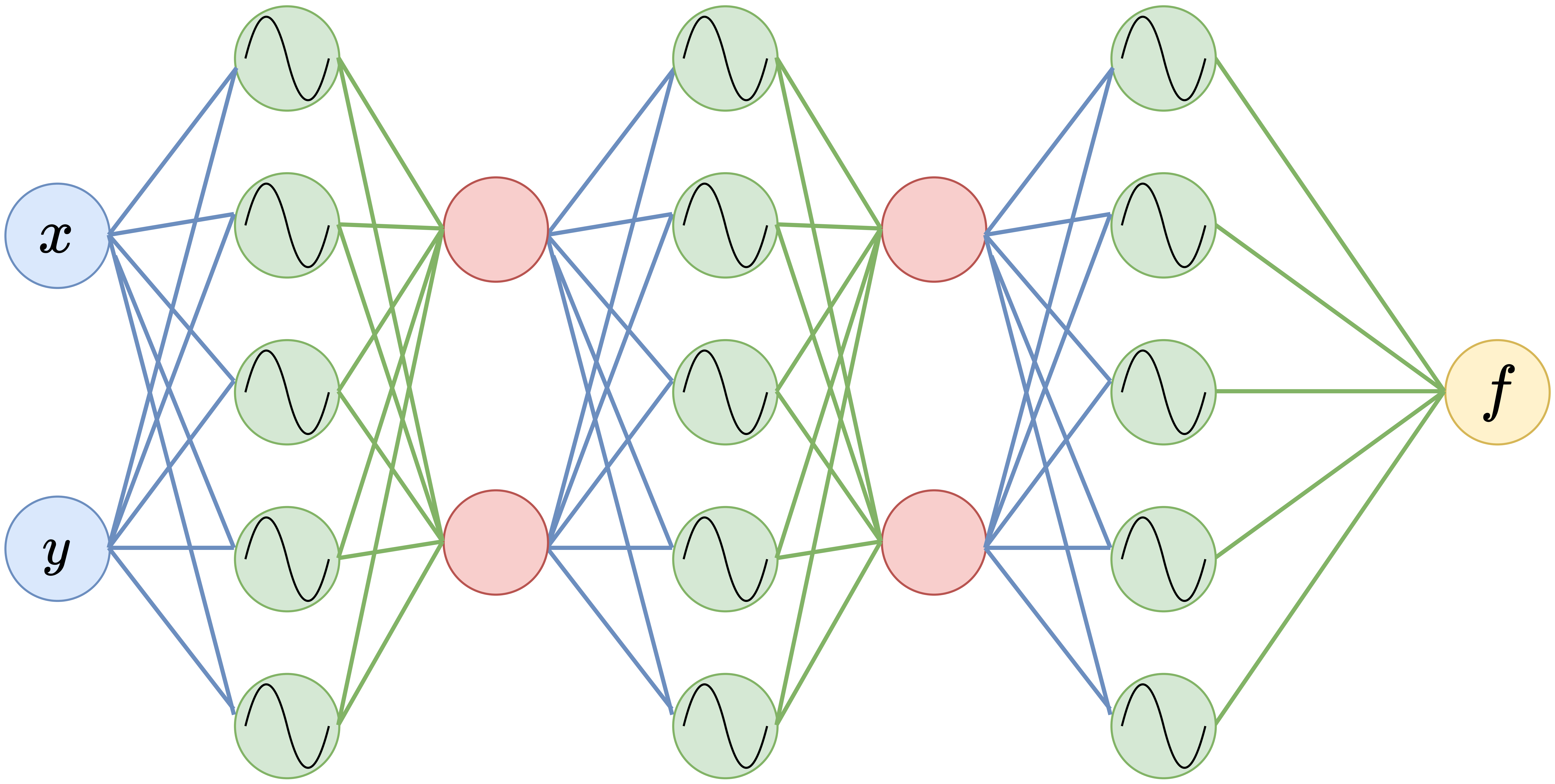
$$\text{where} \quad \phi_h(x) = \sqrt{2\sigma^2/H} \cos\left(\omega_h^\top x + b_h\right), \quad h = 1, \ldots, H,$$

$$\text{with } \omega_h \sim p(\omega) \text{ and } b_h \sim U([0, 2\pi])$$



*Rahimi et al. "Random features for large-scale kernel machines." *NeurIPS 2007*

Deep Random Features (DRF)



Spherical Random Features

$$f(z) \approx \sum_{i=1}^{M} \theta_i \, \phi_i(z), \quad \theta_i \sim \mathcal{N}(0,1)$$

Euclidean case ($z = x \in \mathbb{R}^d$)

$$\phi_i(x) = \sqrt{\frac{2\sigma^2}{M}} \cos\left(\omega_i^\top x + b_i\right)$$

$$\omega_i \sim p(\omega), \quad b_i \sim U([0, 2\pi])$$

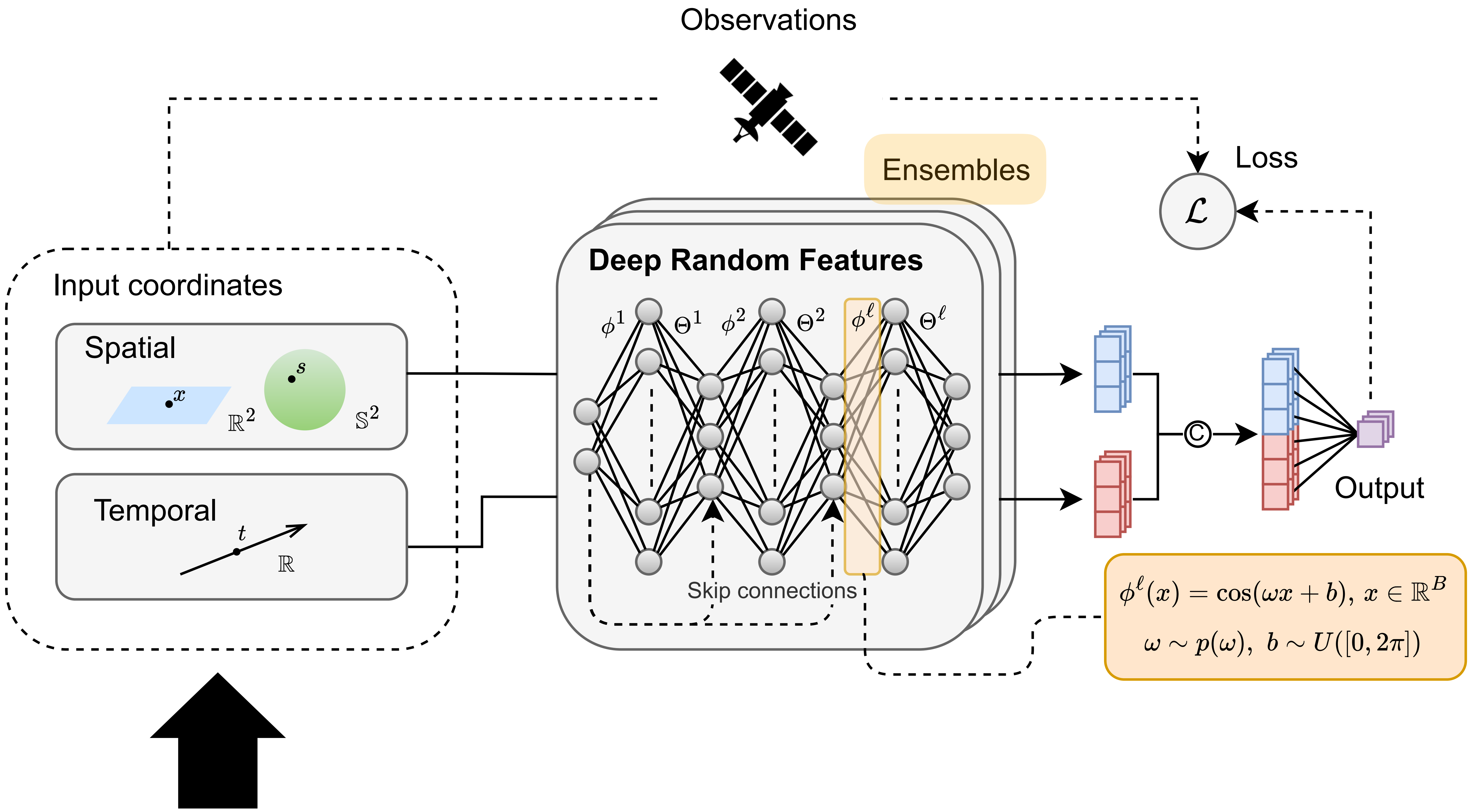
Spherical case ($z = s \in \mathbb{S}^2$)

$$\phi_i(s) = \sqrt{\frac{1}{M} c_{\omega_i} \, \mathcal{G}_{\omega_i}^{1/2}(d_{\mathbb{S}^2}(s, b_i))}$$

$$\omega_i \sim \mathrm{Multinomial}(C_\Phi^{-1}\Phi(\lambda_1), \ldots, C_\Phi^{-1}\Phi(\lambda_J))$$

$$b_i \sim U(\mathbb{S}^2)$$

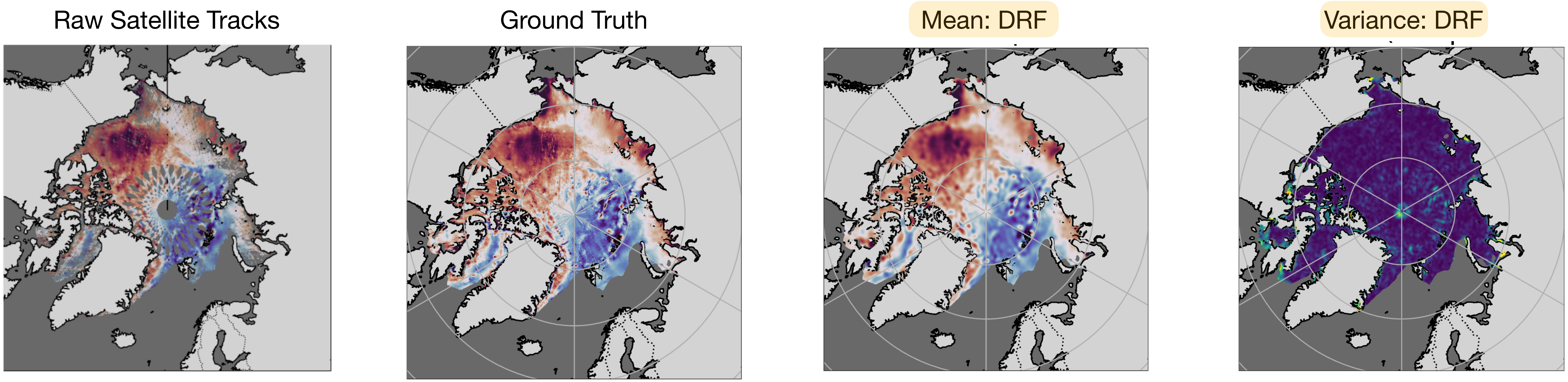
General training settings



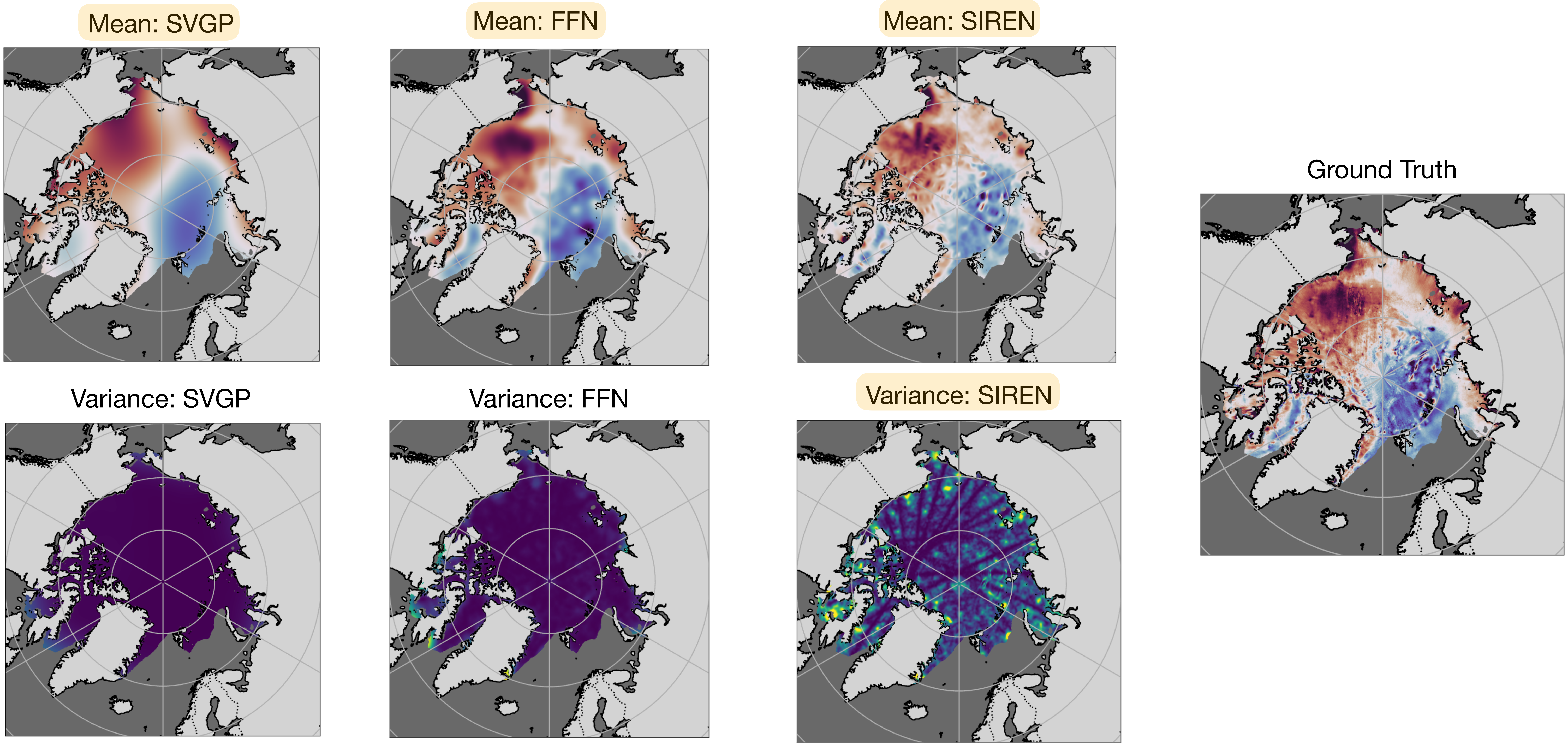
- **Deep ensembles** for uncertainty quantification
- **Inner loop**: Train ensemble of models
- **Outer loop**: Tune hyperparameters via Bayesian optimisation

Results

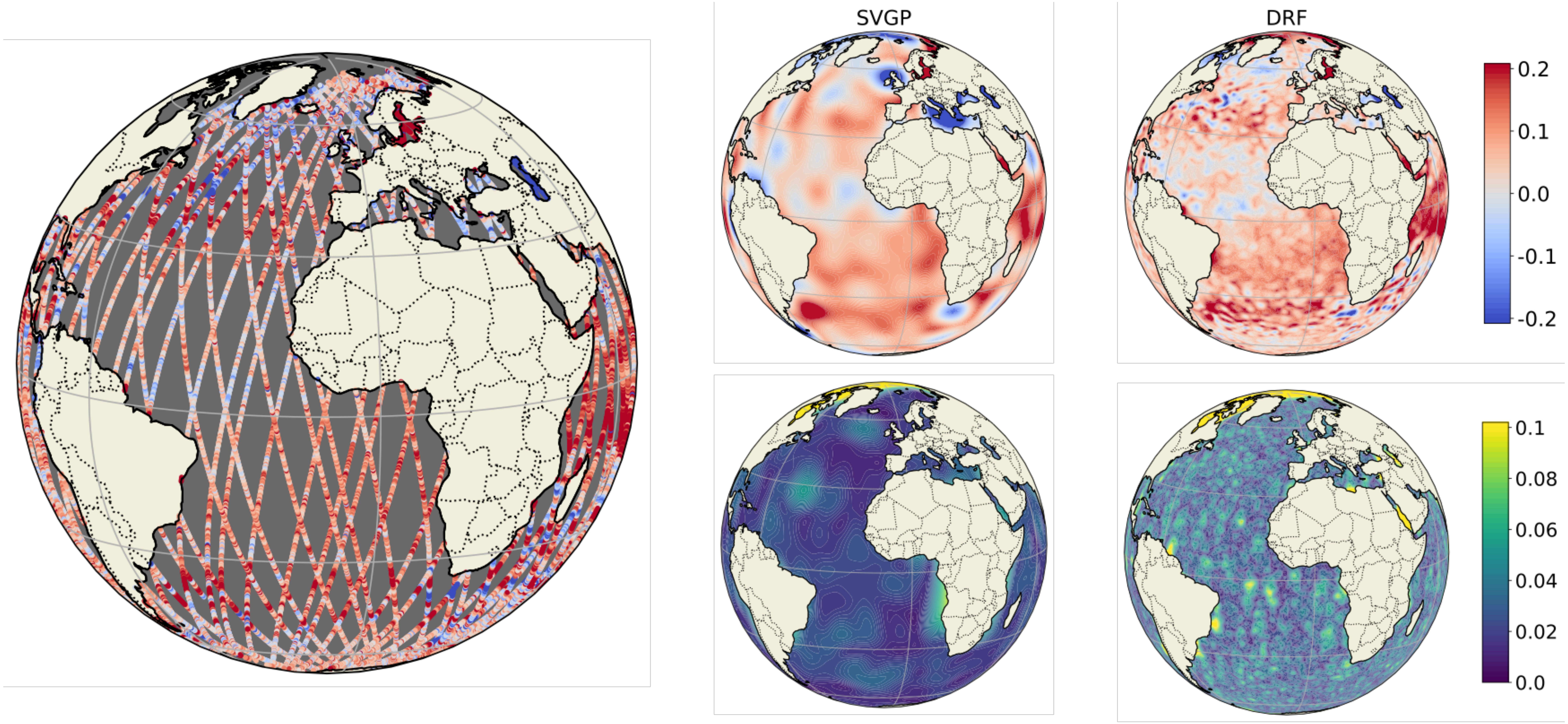
Local scale interpolation of mean sea surface



Comparisons with baselines



Results: Global scale



Conclusion

+ Scalability

+ Learn non-stationary fields

- Hyperparameter tuning is challenging