UNIVERSITY OF CAMBRIDGE

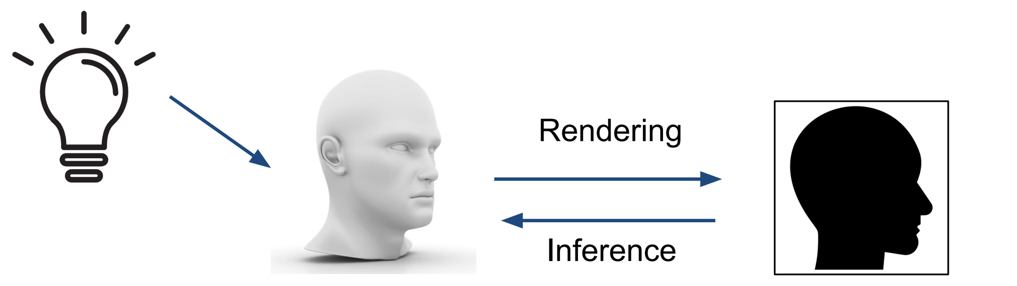
Scan the QR code to learn more

From Spectra to Biophysical Insights: End-to-End Learning with a Biased Radiative Transfer Model

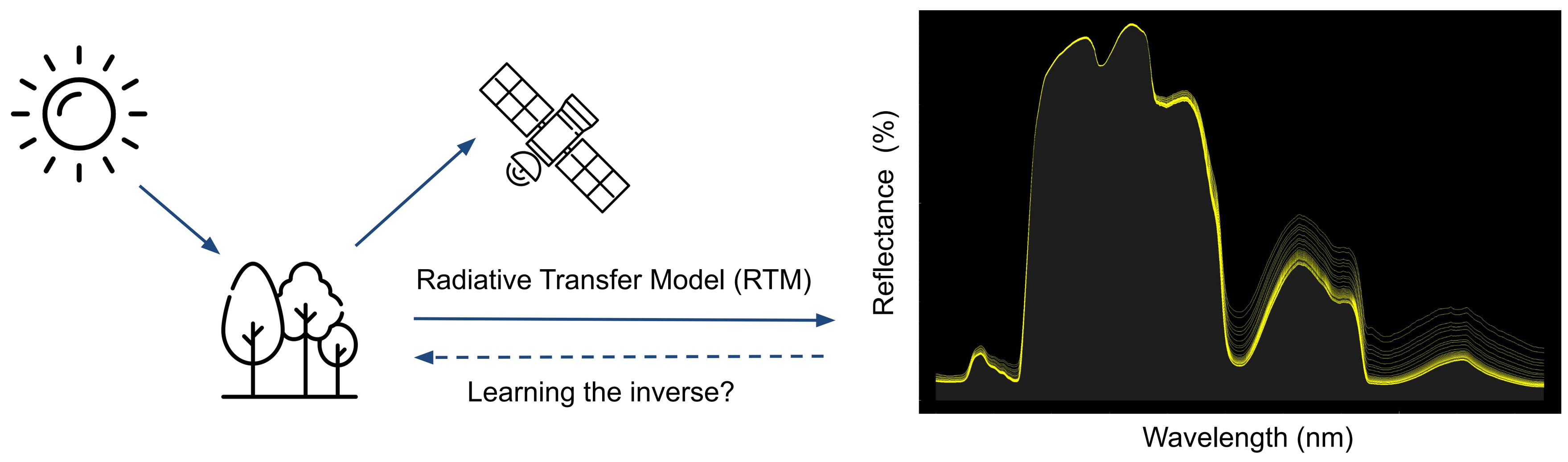
Yihang She [1], Clement Atzberger [2], Andrew Blake [1, 2] and Srinivasan Keshav [1]

[1] University of Cambridge, [2] Mantle Labs

1. Biophysical Representations of Forests



Differentiable Rendering
Can this concept be generalized to retrieve biophysical variables of forests?

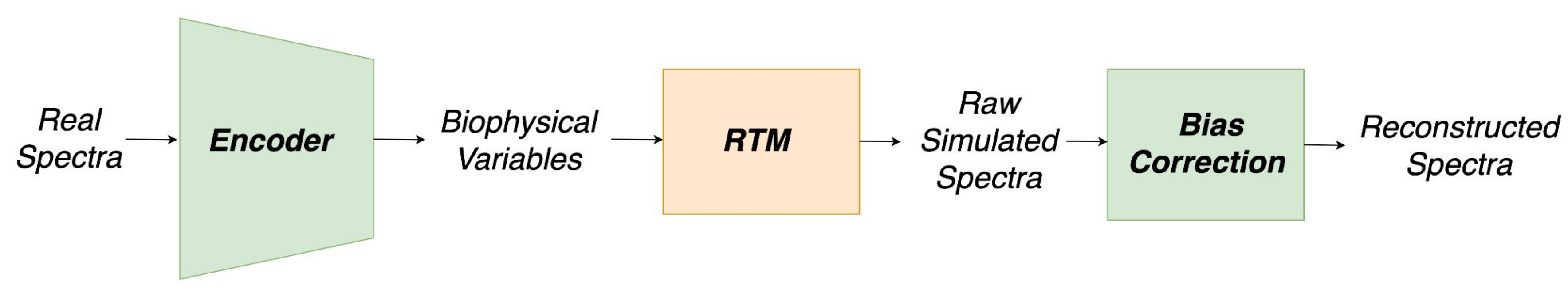


Challenges to invert the RTM

- Not uniquely invertible
- Systematic biases in simulated spectra
- Numerical model differentiability

State-of-the-Art: regressive neural network

Our Approach: learning the inverse with an auto-encoder



2. End-to-End Learning of the Inverse

INFORM — a forest RTM — and its variables to learn.

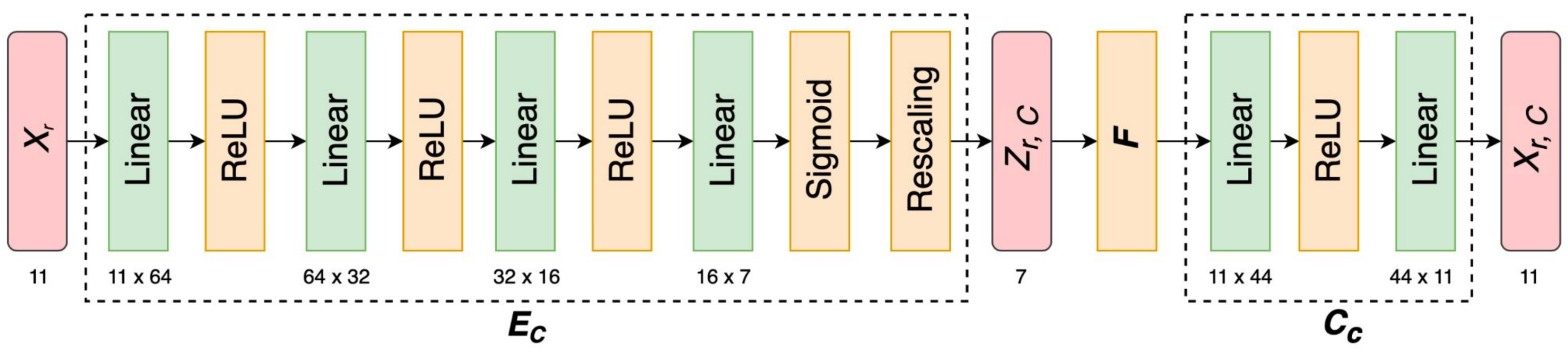
Sub-model	Leaf Model				Canopy Model	Forest Model			
Variable	Structure Parameter	Chlorophyll A+B	Water Content	Dry Matter	Leaf Area Index	Undergrowth LAI	Fractional Coverage	Tree Height	Crown Diameter
Acronym	N	cab	cw	cm	LAI	LAIu	fc	h	cd
Min	1	10	0.001	0.005	0.01	0.01	0.1	-	-
Max	3	80	0.02	0.05	5	1	1	-	-

Sentinel-2 data to use. Species and temporal information will be used to verify retrieved variables' plausibility .

Total Number of Spectra	Number of Individual Sites	Number of Dates	Number of Species
17962	1283	14	12

End-to-end pipeline to invert the RTM

- Integrating INFORM into an auto-encoder structure
- Bias correction function
- Making a fully differentiable INFORM assisted by GPT-4



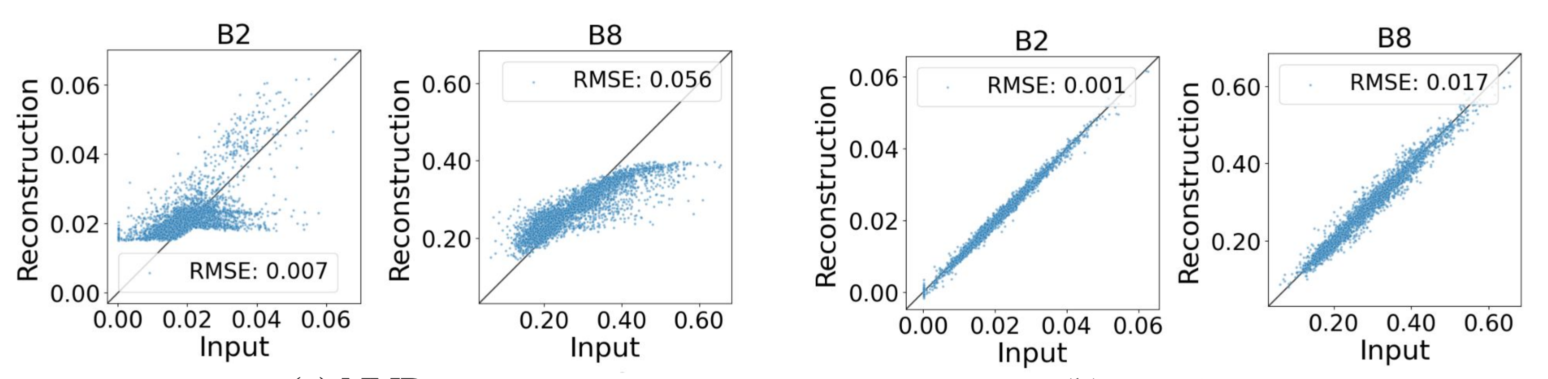
Our model (AE_RTM_corr) can do end-to-end learning of the inverse. Regressive neural network — a classical approach — serves as a **baseline (NNRegressor).**

3. Results

Bias correction. Significantly improved reconstruction loss.

AE_RTM_corr achieves **over ten times lower** MSE than the baseline NNRegressor.

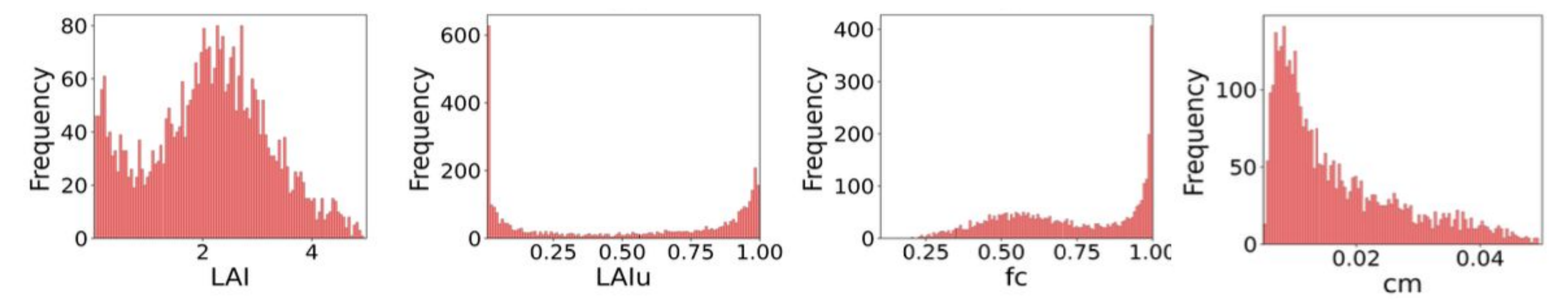
Model	Architecture	Dataset	MSE_{train}	MSE_{val}	MSE_{test}
AE_RTM_corr	$\mathbf{E}_C$+$\mathbf{F}$+$\mathbf{C}_C$	D_r	0.0210	0.0235	0.0217
NNRgressor	$\mathbf{E}_D$	D_r	-	-	0.6676



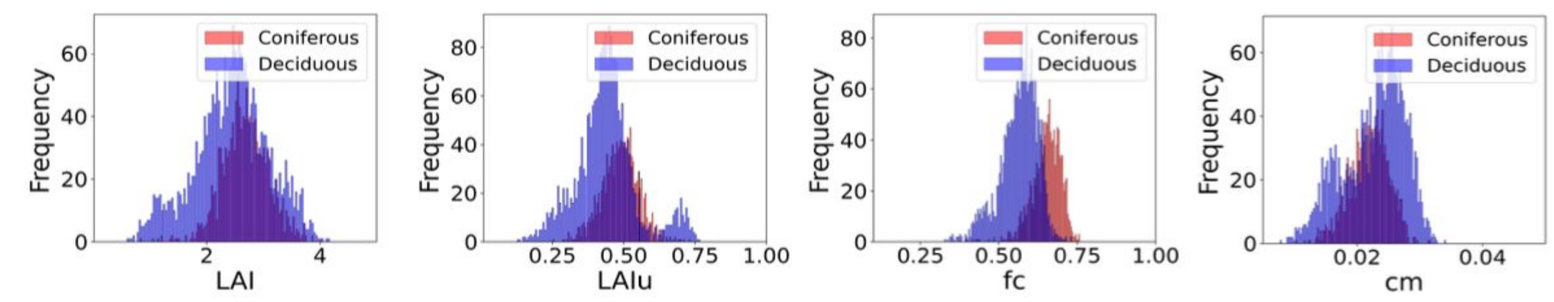
(a) NNRegressor (b) AE_RTM_corr

Superior reconstruction accuracy for AE_RTM_corr, illustrated by spectral band.

Biophysical variables. More plausible recovery.

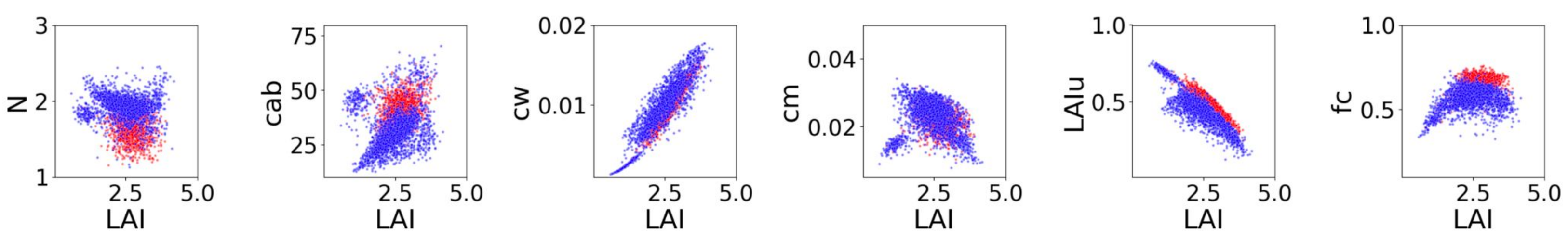


(a) NNRegressor

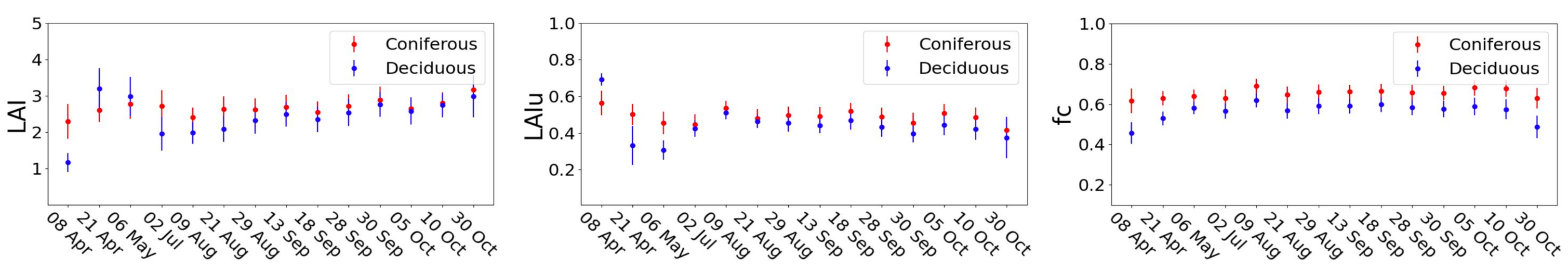


(b) AE_RTM_corr

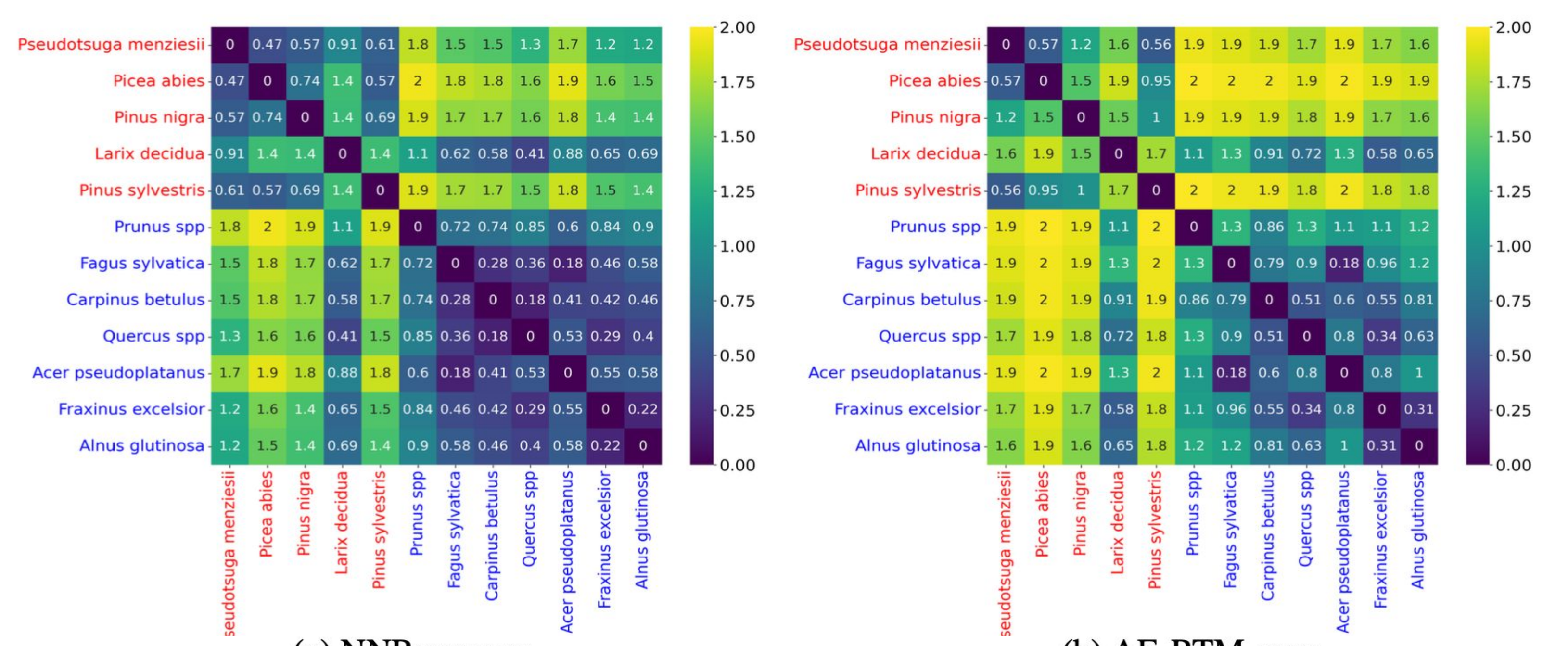
Baseline application to real spectra yields implausible distributions of variables, while AE_RTM_corr learns **plausible and distinguishable distributions** across forest types.



Pairwise co-distributions of variables shows that AE_RTM_corr can learn **distinct physical patterns**. Red: coniferous forest. Blue: deciduous forest.



Our model effectively captures **distinct, temporally smooth and plausible variations** of inferred physical parameters for different forest types.



(a) NNRegressor (b) AE_RTM_corr

Pairwise Jeffreys-Matusita (JM) distance between species based on the learned variables. Our model can learn **more disentangled variables** within the latent space.

4. Conclusion

- End-to-end learning with a complex RTM.
- Plausible recovery of variables with bias correction.
- Implications for inverting biased physical models.

Acknowledgement: This work was supported by the UKRI Centre for Doctoral Training in Application of Artificial Intelligence to the study of Environmental Risks (reference EP/S022961/1) and Cambridge Centre for Carbon Credits. We would also like to thank Markus Immitzer from Mantle Labs for sharing the Sentinel-2 data with us.

Contact: ys611@cam.ac.uk