Meta-modelling paths of Simple Climate Models using Neural Networks and Dirichlet polynomials: An application to DICE

based on joint work with

E. Gobet[1] Y. Liu[1] G. Vermandel[1]

[1]École Polytechnique

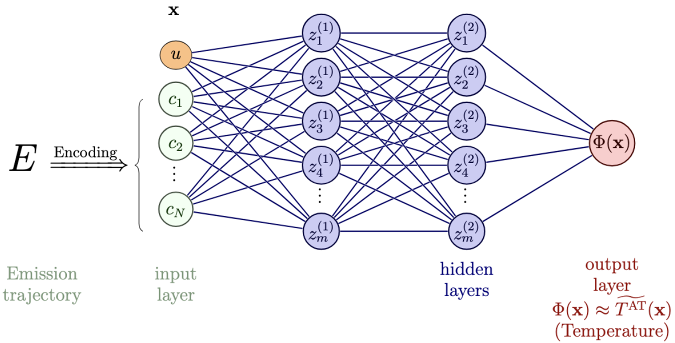
https://hal.archives-ouvertes.fr/hal-04990321

Insurance Data Science Conference - London - 19 June 2025

- **Theoretical justification of using neural surrogates for ODEs system**
  - How to infer temperature from carbon emission?
  - **Simple Climate Models (eg. DICE)**: **computationally demanding** for century-scale scenario (ODE solvers)
  - Semi-infinite time interval : ODEs system defined on a non-compact time domain
- **Scenarios that reflect exponential change...**
  - **CO_2 Dynamics**: Multi-timescale effects Pierrehumbert
  - Requires **exponential solutions**: Green tech surges, Renewables...

Pierrehumber: CO_2 acts like a mixture of decadel-, centennial-, millennial-, and infinite-lifetime gases [Pierrehumbert, 2014]

From Emissions to Instant Climate Forecasts: Neural Surrogate



Exponential decoded emission trajectories & AI

- **Break Emission trajectories into Exponentials**:
  Generalized Dirichlet polynomials
  $E(t) \approx E_{GD}(t) := \sum_{l=1}^{N_{GD}} \left(c_{2l-1} e^{-\lambda_l t} + c_{2l} e^{-\lambda_l t} t \right)$
- **Neural Surrogate**: 100× faster than DICE

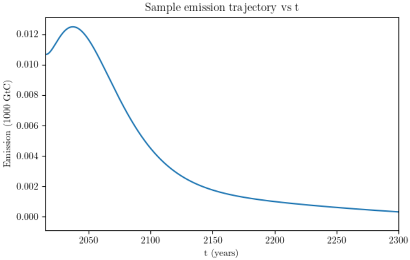
Theoretical Backbone

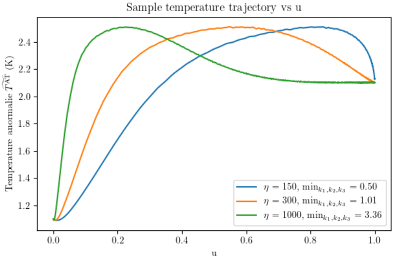
The ReLU Neural network can approximate **smooth** functions well[Yarotsky, 2017]. Our strategy

- Time change: $(0, \infty) \to [0, 1]$ using parameter η,

$$u = 1 - e^{-t/\eta}. \tag{1}$$

- Smoothness condition modulated by η:





Main theoretical results

Theorem 1

Assume the parameter $\eta > 0$ *in the time-change* (1) *satisfies:*

$$\begin{cases} \lambda_l \eta \notin \mathbb{N} \text{ and } \lambda_l \eta > 1, & \forall \lambda_l \in \{\lambda_1, \ldots, \lambda_{N_{GD}}\}, \\ \lambda_{M_i} \eta \notin \mathbb{N} \text{ and } \lambda_{M_i} \eta > 1, & \forall \lambda_{M_i} \in \{\lambda_{M_1}, \lambda_{M_2}\}, \\ \lambda_{T_i} \eta \notin \mathbb{N} \text{ and } \lambda_{T_i} \eta > 1, & \forall \lambda_{T_i} \in \{\lambda_{T_1}, \lambda_{T_2}\}. \end{cases} \tag{2}$$

Define:

$$k_1 := \min_{\lambda_l \in \{\lambda_1, \ldots, \lambda_{N_{GD}}\}} \lfloor \lambda_l \eta \rfloor \geq 1, \quad k_2 := \min_{i=1,2} \lfloor \lambda_{M_i} \eta \rfloor \geq 1, \quad k_3 := \min_{i=1,2} \lfloor \lambda_{T_i} \eta \rfloor \geq 1.$$

For any $\varepsilon \in (0,1)$ *there exists a standard fully-connected (dense) feedforward Neural Network* Φ *with ReLU activations such that*

$$\sup_{x=(u,\omega) \in [0,1] \times \mathcal{B}_\omega} |\widetilde{T^{\mathrm{AT}}}(x) - \Phi(x)| \leq \varepsilon, \tag{3}$$

with $D = 2N_{GD} + 6$, *we have complexity bounds:*

$$L(\Phi) \leq c_1 \left(\log(\frac{1}{\varepsilon}) + 1 \right), \qquad N(\Phi) \leq c_2 \varepsilon^{-\frac{D}{\min(k_1, k_2, k_3)}} \left(\log(\frac{1}{\varepsilon}) + 1 \right),$$

with some constants c_1 *and* c_2 *depending on* D, $\min(k_1, k_2, k_3)$ *and the function* $\widetilde{T^{\mathrm{AT}}}$.

High precision NN approximation with simple architecture

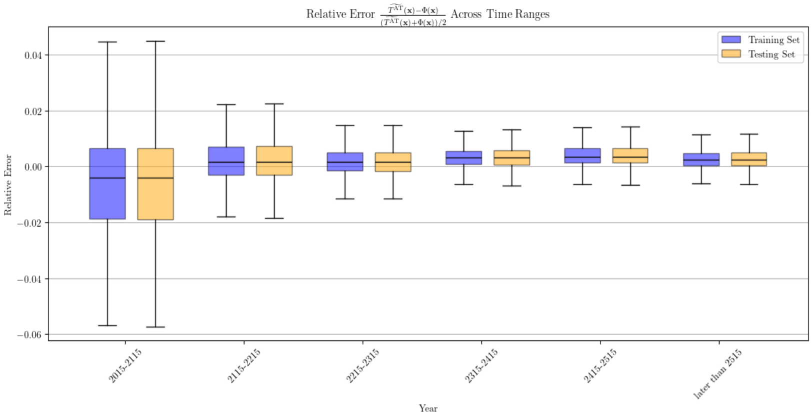


Figure: Boxplot comparison of relative error across time ranges. The neural network has been trained using the augmented dataset, where $n_{\text{traj}} = 1000 \times 8$ (1000 trajectories for each SSP emission trajectory) and $n_{\text{time}} = 500$.

Faster mapping

Method	$n_{\text{time}} = 100$	$n_{\text{time}} = 500$	$n_{\text{time}} = 1000$
ODE Solver	75.4654	158.5378	338.36
Neural Network (NN)	0.6414	2.5254	4.20

Table: Computation time (in seconds) for ODE solver and neural network for 500 emission trajectories with different values of n_{time}.

Conclusion and Perspectives

Achievements:

- **Traditional model**: ODEs defined on non-compact time domain;
- A suitable time change to retrieve boundedness;
- Emission trajectory encoding using exponentials (Dirichlet Polynomials);
- The transformed ODEs satisfy suitable smoothness properties with respect to input parameters;
- A fast-to-evaluate meta-model from emission trajectories to temperature ones.

Future-Proofing:

- **Beyond CO_2**: Expand to Methane, other aerosols, encode multi-gas exponentials for holistic policy assessments.
- **Beyond DICE**: Extend to more complex climate systems.

More info: hal-04990321

References I

Pierrehumbert, R. T. (2014).
Short-lived climate pollution.
Annual review of earth and planetary sciences, 42(1):341–379.

Yarotsky, D. (2017).
Error bounds for approximations with deep ReLU networks.
Neural Networks, 94:103–114.

Emission Trajectories approximated Generalized Dirichlet Polynomials

Theorem 2

Assume that E is k-times continuously differentiable on $[0,\infty)$ and that the derivatives of E converge η-exponentially fast to 0. Then, there is a finite constant c that depends on $k, \kappa_1, \cdots, \kappa_k, \eta$ (but not on N_{GD}) such that

$$\inf_{\lambda_1,\cdots,\lambda_{N_{GD}}} \inf_{c_1,\cdots,c_{2N_{GD}}} \sup_{t\geq 0} \left| E(t) - \sum_{l=1}^{N_{GD}} \left(c_{2l-1} e^{-\lambda_l t} + c_{2l} e^{-\lambda_l t} t \right) \right| \leq c \frac{1}{{N_{GD}}^{k-1}},$$

for any $N_{GD} \geq k$.