Research Article ISSN 3069-8316

Journal of Advances in Artificial Intelligence and Machine Learning

Validation of the MAL using Neural Networks

Éric ZELTZ*

Independent Researcher, La Motte-en-Champsaur, France.

***Correspondence:**
Éric ZELTZ, Independent Researcher, La Motte-en-Champsaur, France.

Received: 03 Jul 2026; **Accepted:** 10 Aug 2026; **Published:** 20 Aug 2026

Citation: Éric ZELTZ. Validation of the MAL using Neural Networks. J Adv Artif Intell Mach Learn. 2026; 2(1): 1-9.

ABSTRACT

This document presents a validation of the results obtained using the Method of Average Lengths (MAL) through neural networks, specifically MLPRegressor and Random Forest models. Developed by the author starting in 2021, the MAL method enables the detection of non-linear or time-lagged interactions in climate time series, which often remain undetected by classical statistical methods (e.g., correlation, R^2, Granger-Newbold cointegration).

The study focuses on three key interactions:
1. Between heat of Upper Ocean Stratum (UOS) and atmospheric temperature (T): Neural networks confirm a strong relationship, with a lagged effect of T on UOS, reflecting oceanic thermal inertia.
2. Between ocean stratification (strat) and UOS: Models validate strong thermal coupling but fail to capture the structural influence of ENSO on stratification, which is detected by MAL.
3. Between oceanic cloud cover (OC) and UOS: A feedback loop is confirmed, where UOS influences OC and vice versa, supporting the hypothesis of a "natural oceanic thermostat."

The results demonstrate that neural networks complement the MAL method by quantifying relationships, while MAL detects dynamic signals inaccessible to traditional methods. This hybrid approach opens new perspectives for modeling complex climate interactions.

Keywords

Method of Average Lengths (MAL), Neural networks, Multilayer Perceptron (MLP), Random Forest, Climate time series, Nonlinear interactions, Time-lagged relationships, Upper Ocean Stratum (UOS), Ocean stratification, ENSO.

Introduction

Climate time series are often characterized by non-linear interactions and time lags, which are difficult to identify using classical statistical tools. The Method of Average Lengths (MAL), introduced by Zeltz [1-3], offers an innovative approach to detect these hidden signals by relying on Markov chains. However, to validate and refine these results, it is essential to compare them with predictive models capable of assessing their ability to explain observed variations.

This is where neural networks (*MLPRegressor* and *Random Forest*) come into play as a tool for quantitative validation. By directly modeling the relationships between climate variables identified by MAL, they allow us to:

- Confirm detected interactions by quantifying their strength (using metrics such as R^2 or RMSE),
- Refine the understanding of underlying mechanisms, particularly lagged effects or feedback loops,
- Complement MAL by providing a numerical dimension to the detected signals, where MAL is limited to qualitative detection.

For example, where MAL reveals alternating Markov-1 signals between atmospheric temperature (T) and heat content in Upper Ocean Stratum (UOS), neural networks demonstrate that including time lags (T(t-1), T(t-2), T(t-3)) significantly improves UOS prediction, thus validating the hypothesis of oceanic thermal inertia.

Similarly, for ocean stratification (strat) and UOS, models confirm strong thermal coupling, while highlighting that the structural influence of ENSO on stratification—detected by MAL—eludes classical metrics like R^2. This synergy between MAL and neural networks enables a more robust analysis of climate mechanisms, while highlighting the strengths and limitations of each approach. Neural networks, as universal approximators, provide essential quantitative validation, while MAL detects non-linear dynamics that traditional models cannot capture.

To achieve these objectives, this work is divided into four sections and a general conclusion, summarized here:

1. Theoretical Foundations of Neural Networks for Validating and Refining MAL Results. Presentation of the principles of MAL and neural networks, as well as their complementarity in detecting and modeling non-linear interactions in climate time series.
2. Validation of the Relationship Between UOS and Atmospheric Temperature (T)
   Analysis of the interactions between these two variables using predictive models, incorporating a time-lag effect to quantify the thermal inertia of the oceans.
3. Validation of the Relationship Between Ocean Stratification (strat) and UOS
   Study of the influence of UOS on stratification and the structural role of ENSO, detected by MAL but difficult to capture with classical linear models.
4. Validation of the Relationship Between Oceanic Cloud Cover (OC) and UOS
   Exploration of the link between these variables, confirming the hypothesis of a feedback loop acting as a "natural thermostat" in the climate system.
5. General Conclusion
   Synthesis of the results, highlighting the complementarity between MAL and neural networks for studying complex climate dynamics.

Section 1: Theoretical Foundations of Neural Networks for Validating and Refining MAL Method Results

Introduction

The Method of Average Lengths (MAL) method, introduced by Zeltz [1-3] and detailed in Zeltz [4], is an innovative statistical approach based on Markov chains [5]. It enables the detection of non-linear or lagged interactions between climate time series—interactions that remain undetected by conventional methods (correlation, R^2, Granger-Newbold cointegration).

The work presented here aims to validate and refine the results obtained via the MAL method using multi-layer neural networks (Multi-Layer Perceptron, MLP). After briefly outlining the principles and stages of the MAL method, this section presents the theoretical foundations of MLPs, their advantages and limitations, and their role in validating the MAL method.

MAL Principle

Briefly summarized, the MAL method relies on the following steps:

1. **Time series transformation:** Successive differences are calculated from a data series (e.g., monthly temperature).It is verified that this series of differences can be modeled as white noise [5].
2. **Binarization of differences:** A value of 0 is assigned to each decrease and 1 to each increase. The approximate equiprobability of 0s and 1s is verified. The resulting chain exhibits all the characteristics of a Markov chain; determining its order is the key to the method.
3. **Finalization:** This is achieved by calculating the average lengths of sequences of 0s or 1s: if they deviate significantly from the theoretical value expected under a binomial distribution (≈1.94 for 100 terms), the Markov chain in question is likely not of order 0 (the binomial case) but of a higher order. This reveals a probable interaction with another climatic entity that explains this acceleration—or, conversely, this retardation—in the alternation of 0s and 1s.

MAL thus enables the detection of interactions that are not captured by classical statistical methods for time series analysis.

Role of Neural Networks in Validating MAL

Multi-Layer Perceptrons (MLPs) are used to validate the MAL results by directly modeling the relationships between the relevant climatic variables. By capturing the complex non-linearities identified by MAL, they will enable a numerical assessment of the impacts of these interactions.

According to the Universal Approximation Theorem [6], MLPs can approximate any continuous function with arbitrary precision, provided they have:

- A sufficient number of neurons.
- A non-linear activation function (e.g., ReLU, sigmoid).

MLPs can therefore model non-linear relationships between climate variables, confirming or refuting the interactions detected by MAL. However, they have several limitations:

- Extrapolation: MLPs may perform poorly when predicting outside the training data range.
- Noise: MLPs do not distinguish noise from signal and may overfit noisy data.
- Convergence: Gradient descent can converge to local minima.

MLPs must therefore be rigorously validated to avoid confusing noise with signal.
Their use should complement MAL, not replace it.

Validation and refinement of results obtained via MAL

MAL has enabled the detection of potential interactions between time series. This work is presented in Zeltz [1-3]. For a summary, see Zeltz [4].

To validate and refine these results using MLPs, we will proceed as follows:

1. Use MLPs to directly model the relationships between the variables identified by MAL.
2. Compare the results:
    - If MLPs confirm the interactions detected by MAL, this reinforces the validity of the results.
    - If MLPs refute or refine the MAL results, this helps clarify the underlying mechanisms.

This approach offers several advantages:

- Complementarity: MAL detects non-linear interactions, while MLPs model these interactions.
- Robustness: MLPs allow for the validation of MAL results on independent datasets.
- Precision: MLPs can refine MAL results by quantifying the relationships between variables.

Neural networks (MLPs) thus offer a complementary framework for validating and refining these results by directly modeling the relationships between variables.

Section 2: Validation of the relationship between UOS and atmospheric temperature (T)

Overview

The results obtained using multilayer neural networks (implemented in Python using the MLPRegressor model, unless otherwise specified) confirm and reinforce the conclusions drawn from the MAL method regarding the interaction between upper-ocean heat content (UOS) and global atmospheric temperature (T). While the MAL method revealed alternating Markov-1 signals for T and UOS—suggesting non-linear interaction dynamics between these two climate parameters—the neural network approach allows for the quantification and validation of this relationship by explicitly incorporating a time-lag effect.

Recap of MAL results regarding interactions between T and UOS

In Zeltz [1], the MAL method demonstrated that:

- The time series for T and UOS exhibited alternating Markov-1 signals, indicating a tendency for rapid reversals in monthly variations for T and quarterly variations for UOS.
- These signals were interpreted as the result of physical interactions between the atmosphere and the ocean, particularly through feedback mechanisms (e.g., increased evaporation → increased cloud cover → atmospheric cooling).
- However, while the MAL method is powerful for detecting interaction signals, it does not allow for the direct quantification of the temporal impact of T on UOS.

Validation using neural networks

To validate and quantify this interaction, we trained a multilayer neural network (MLPRegressor) to predict UOS based on past values of T. Two models were tested:

1. Model without time lag: UOS = f(T), where UOS is predicted solely from the instantaneous atmospheric temperature T. To achieve this, the NOAA-provided UOS data[1] —available on a quarterly basis (values for March, June, September, and December of each year from 1955 to 2026)—were aligned with the monthly temperature (T) data provided by NOAA[2] as follows:
2. each monthly UOS value for a given quarter is set to one-third of the quarterly UOS value for that quarter. This approach yields monthly UOS data while preserving the original quarterly information. Figure 1a shows the curve of the actual data (solid blue line) and the curve of the model's predictions (dotted pink line).

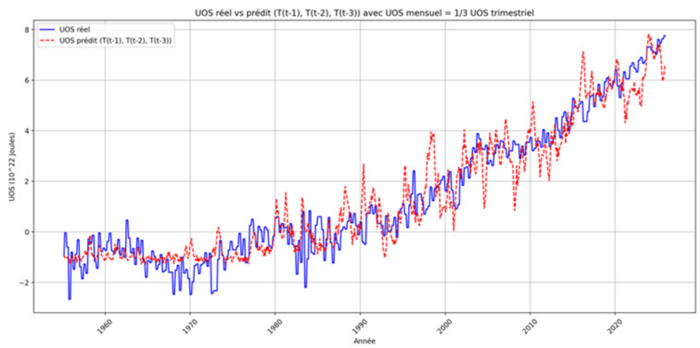

Figure 1a: UOS = f (T) and actual UOS data.

2. Model with time lag: UOS = f(T(t-1), T(t-2), T(t-3)), where UOS is predicted based on temperatures from the preceding three months. Figure 1b shows the curve of the actual data (solid blue line) and the curve of the model's predictions (dotted red line).

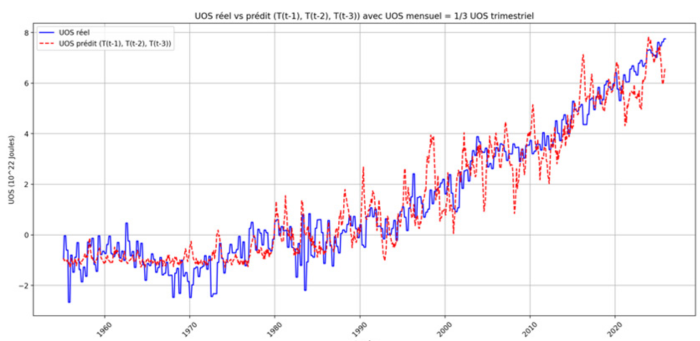

Figure 1b: UOS = f(T(t-1), T(t-2), T(t-3)) and actual UOS data.

The metrics used to evaluate the models are:

- R^2 (coefficient of determination), which measures the proportion of UOS variance explained by the model.
- RMSE (Root Mean Square Error), which measures the average error between predicted and actual values. Table 1 presents these metrics as well as the number of iterations required for the predictions to stabilize. The final column contains the interpretations, which we elaborate on subsequently.

1 https://www.climate.gov/media/13603
2 https://www.ncei.noaa.gov/access/monitoring/climate-at-a-glance/global/time-series/globe/land_ocean/

Table 1: Comparison of the two models.

Model	R^2	RMSE (× 10^{22} Joules)	Iterations	Interpretation
UOS = f(T)	0.87	2 .81	50	Good prediction, but does not capture the temporal inertia of UOS.
UOS = f(T(t-1), T(t-2), T(t-3))	0.90	0.84 × 3 = 2.52	94	Better prediction: The model captures the lagged effect of T on UOS.

The significant improvement in metrics (R^2 +3%, RMSE -10%) demonstrates that UOS indeed depends on the recent history of T, rather than solely on its instantaneous value. Ocean thermal inertia is a key factor: variations in T take several months to impact UOS, which is consistent with known physical mechanisms (see, for example, Durack et al., [7], or Levitus et al., [8]).

The prediction plots (see Figures 1a and 1b) clearly illustrate this improvement:

- In the UOS = f(T) model, predictions (dashed curve) generally follow the trend of the actual data (solid curve) but show notable discrepancies, particularly during rapid UOS fluctuations. This reflects the model's inability to capture the time lags between T and UOS.
- In the UOS = f(T(t-1), T(t-2), T(t-3)) model, predictions align more closely with the actual data, including during peaks and troughs. The overlap between the two curves is much tighter, confirming that including time lags allows for a better reproduction of UOS temporal dynamics.

These results validate and complement the conclusions of the MAL method:

- MAL had detected interaction signals between monthly T and quarterly UOS, suggesting coupled dynamics between these two parameters.
- Neural networks confirm this interaction by showing that UOS prediction is significantly improved when past values of T are taken into account. This reinforces the hypothesis that the underlying physical mechanisms (e.g., thermal inertia, atmosphere-ocean feedbacks) are effectively captured by both methods, albeit in a complementary manner:
- MAL detects non-linear signals (e.g., rapid alternation).
- MLP quantifies the temporal impact of these interactions.

These results have significant implications for understanding climate dynamics:

1. Validation of physical mechanisms: Improved predictions when accounting for time lags confirm that ocean thermal inertia plays a major role in the relationship between T and UOS. This supports the view that climate models must incorporate time lags to accurately represent atmosphere-ocean interactions.
2. Complementarity of methods: MAL is a powerful tool for detecting hidden interactions in climate time series. Neural networks enable the validation and quantification of these interactions, particularly those involving lagged effects. Together, these methods offer a robust approach to studying complex climate dynamics.

Conclusion of Section 2

Neural networks confirm the results obtained via MAL:

- Lagged temperatures (T(t-1), T(t-2), T(t-3)) are robust predictors of UOS.
- The relationship is strong, reflecting differences in climate memory between the atmosphere (reactive) and the ocean (inertial).
- These results support the hypothesis of strong coupling between UOS and T, wherein UOS acts as a thermal reservoir modulating atmospheric variations.

This cross-validation between a non-linear statistical method (MAL) and a machine learning approach (MLP) opens up new perspectives for climate system modeling.

Section 3: Validation of the relationship between ocean stratification (strat) and UOS

Overview

In this instance, MLPs are used to verify certain conclusions drawn from the MAL method, as presented in Zeltz [3]. Specifically, they allow for the validation and quantification of the interaction detected between ocean stratification and UOS. However, they are unable to confirm another phenomenon detected by the MAL method: the structural effect of the El Niño-Southern Oscillation (ENSO) on the temporal evolution of stratification.

Recap of results obtained via the MAL method

Three variables are involved:

- Semiannual anomalies in ocean stratification (denoted as strat).
- Semiannual anomalies in UOS heat content (denoted as UOS).
- The semiannual ENSO index.

The main conclusion of the Zeltz [3] study is as follows: the MAL signal associated with strat is of the "lengthening" Markov-1 type, whereas the signal associated with UOS is of the "alternating" Markov-1 type. Yet, these two variables are strongly correlated. The lengthening of the strat MAL signal is attributed to the influence of ENSO; without this influence, it would exhibit the alternating Markov-1 pattern seen in UOS.
To test these results, we employed three MLP models.

Stratification model: strat = f(UOS)[3]

For the period studied (1955–2023), we obtained:
- R^2: 0.7825
- RMSE: 0.9288 (in the N^2 units used to measure stratification)
- Interpretation:

3 In this instance, UOS is the variable corresponding to the semiannual anomalies of the heat content within the UOS.

The model confirms that UOS explains 78.25% of the variance in stratification, validating the hypothesis of strong thermal coupling between these two variables.

Figure 2a shows the curve of actual data (solid blue line) and the curve of model predictions (dotted green line).

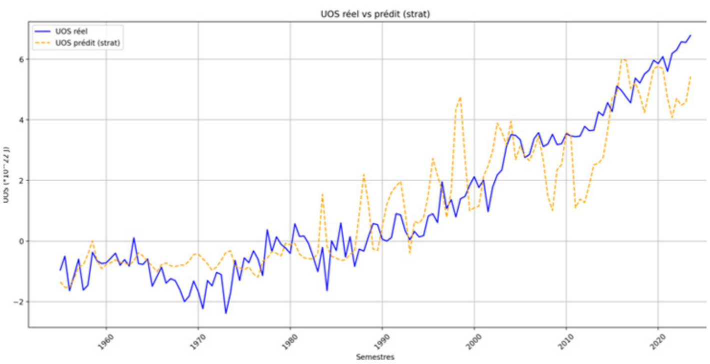

Figure 2a: Strat = f(UOS) and actual stratification data.

Inverse model: UOS = f(strat)

- R^2: 0.8062
- RMSE: 1.0574 ($\times 10^{22}$ Joules)
- Interpretation:

The near-symmetry between strat = f(UOS) (R^2 = 0.7825) and UOS = f(strat) (R^2 = 0.8062) confirms that these two variables are facets of the same climatic phenomenon.

Figure 2b shows the curve of actual data (solid blue line) and the curve of model predictions (dotted orange line).

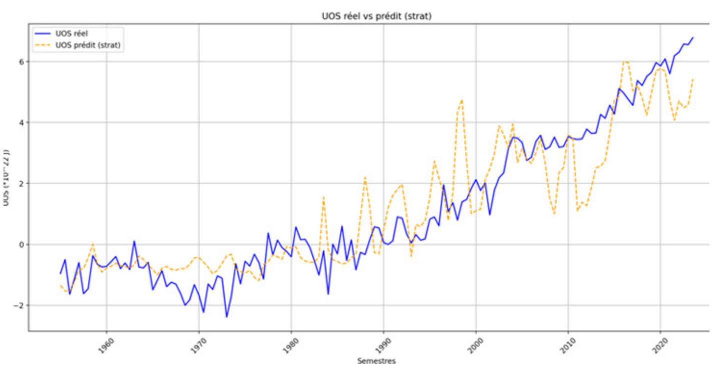

Figure 2b: UOS = f(Strat) and actual UOS data.

Model: strat = f(ENSO)

We obtain:
- R^2: 0.0237
- RMSE: 1.9679%
- Interpretation:

ENSO alone explains only 2.37% of the variance in strat. Figure 2c clearly illustrates ENSO's inadequacy in explaining strat (solid blue line for actual data, dashed red line for predictions).

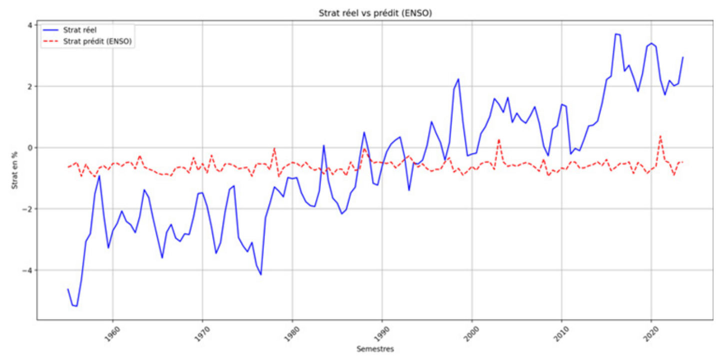

Figure 2c: Strat = f(ENSO) and actual strat data.

This result might appear to contradict the hypothesis proposed in Zeltz [3]—that ENSO alters the temporal structure of stratification (by lengthening the Markov-1 signal). However, there is no contradiction: R^2 measures only the direct linear relationship between ENSO and strat, whereas ENSO's influence on stratification operates primarily through changes in its temporal persistence (lengthening of up/down sequences)—an aspect that R^2 fails to capture. MAL, on the other hand, is capable of detecting this dynamic structure.

Conclusion of Section 3

Neural networks partially validate the MAL signals:
- UOS is the dominant predictor of strat (R^2 = 0.7825), confirming the hypothesis of strong coupling between these two variables.
- ENSO has limited direct influence (R^2 = 0.0237), but its structural role (lengthening of alternation periods) is captured by MAL rather than R^2.
- The quasi-symmetrical relationship between strat and UOS reflects a consistent physical mechanism: stratification acts as an amplifier for UOS thermal anomalies, while UOS serves as the primary driver.

Section 4: Validation of the relationship between oceanic cloud cover (OC) and UOS

Presentation and results obtained via MAL

In this instance, multilayer neural networks are used to verify certain conclusions drawn from the MAL method regarding oceanic cloud cover, as presented in Zeltz [2]. MAL revealed an alternating Markov-1 signal for quarterly OC and UOS anomalies, with a correlation of 0.65 between the two series. The proposed hypothesis involved a feedback loop: UOS warming increases evaporation, thereby raising quarterly OC cloud cover anomalies, which in turn cools the UOS through an enhanced albedo effect. This suggests a "natural thermostat" effect of oceanic cloud cover—a phenomenon frequently proposed in climate literature [9,10].

OC = f(UOS) model

Over the 1955–2008 period, the OC = f(UOS) model yielded an R^2 of 0.63 and an RMSE of 0.59%. Thus, nearly two-thirds of OC variations appear to be explained by UOS variations, a finding consistent with Zeltz [2].

Figure 3 shows the curve for actual data (solid blue line) and the curve for model predictions (dashed orange line).

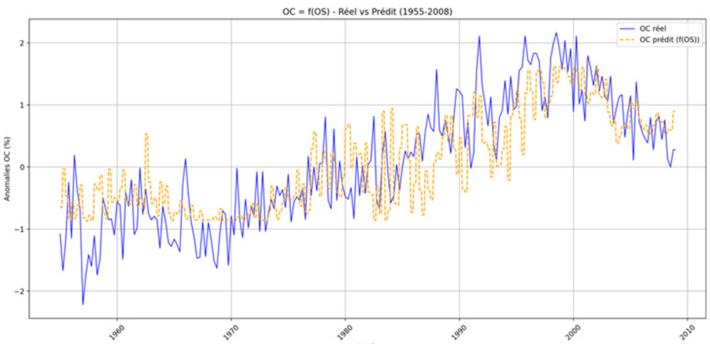

Figure 3: OC = f(UOS) and actual OC data.

OC = f(UOS) model[4]

The OC = f(UOS, T) model was also tested. Table 2 compares the metrics of the two models studied for predicting OC over the 1955–2008 period. We observe that incorporating monthly T data significantly improves the quality of OC predictions.

Table 2: Comparison of the two models.

Model	R^2	RMSE (%)	Iterations
OC = f(UOS)	0.63	0.59	127
OC = f(UOS;T)	0,.72	0.51	214

Figure 4 confirms this, as the predictions (dashed red line) fit the data (blue line) better than those observed in Figure 3 for the OC = f(UOS) model.

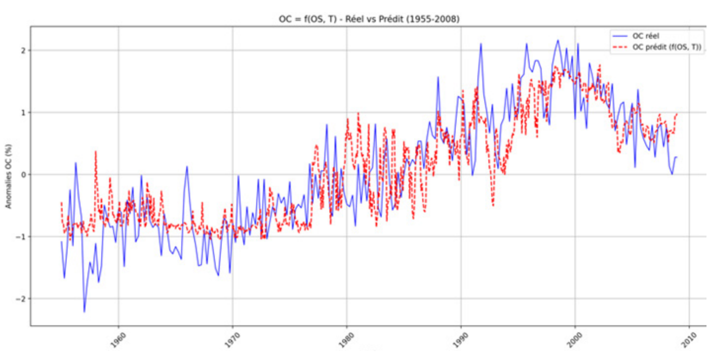

Figure 4: OC = f(UOS, T) and actual OC data.

Inverse model: UOS = f(OC)

To obtain this inverse model, we initially used the standard Python program employing MLPRegressor. However, as clearly shown in Figure 5, this model fails to capture extreme UOS values. The predicted values (orange dashed line) are capped at around 4.9 ($\times 10^{22}$ Joules), whereas the observed data (blue curve) sometimes significantly exceed 10 ($\times 10^{22}$ Joules).

4 UOS is defined on a monthly basis in the same way as in Section 2.

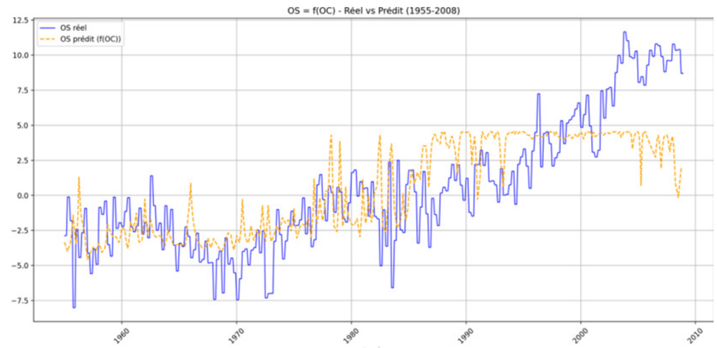

Figure 5: UOS = f(OC) using MLPRegressor and actual UOS data.

This issue arises because, while MLPRegressor is a universal approximator, its design limits its ability to extrapolate beyond the training data.[5]

Consequently, we turned to a neural network model that appeared better suited to handling and utilizing the full dataset, including extreme values: the Random Forest model. Unlike *MLPRegressor*, this model does not require strict data normalization, allowing it to better capture strong non-linearities. It is therefore theoretically capable of providing a satisfactory UOS = f(OC) model.

Figure 6, showing predictions obtained via Random Forest Regressor, confirms this: the issue of capped predicted UOS values has been eliminated.

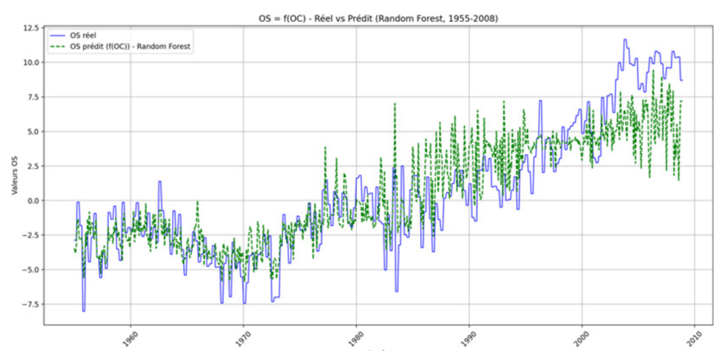

Figure 6: UOS = f(OC) using Random Forest and actual UOS data.

This model yields metrics of $R^2 = 0.72$ and RMSE = 2.42 ($\times 10^{22}$ Joules). Thus, ocean coverage has a significant impact on UOS variations. However, given the differences between the neural systems employed, we cannot conclude that OC is a more significant driver than UOS simply because the inverse model OC = f(UOS) has a lower R^2 value (0.63).

Conclusion of Section 4

Neural networks confirm the results obtained via MAL:
- UOS is a good predictor of OC, and vice versa, validating the hypothesis of a feedback loop; this aligns perfectly with the "natural thermostat" hypothesis regarding OC, as described in Zeltz [2].

5 We specify this in the Appendix.

General Conclusion

The results described above confirm the validity of the MAL method for detecting complex climate interactions. Neural networks have validated the signals identified by MAL by quantifying the direct relationships between variables. For instance, UOS explains 78.25% of the variance in stratification, while lagged temperatures explain 90% of the variance in UOS.

UOS emerges as the primary thermal driver of stratification, whereas oceanic cloud cover is strongly influenced by UOS. Although its direct influence is limited, ENSO alters the temporal structure of stratification—an aspect captured by MAL.

This method is thus validated not only by the consistency of results from models employing it (as seen in Zeltz [1,2] and [3]) but also through the use of a powerful external tool: MLPs.

Acknowledgements

The author wishes to express his sincere gratitude to Professor Jérôme Pousin, who suggested this work.

References

1. Zeltz E. Origins of natural climatic nervousness and its current accentuation. J Mod Green Energy. 2023; 2: 8.
2. Zeltz E. Analysis of the interaction of oceanic cloudiness with the upper oceanic stratum. J Inform Anal. 2024; 2: 4.
3. Zeltz E. Mechanisms driving ocean stratification and mixed layer deepening. Explora Environ Resource. 2024; 1: 4578.
4. Zeltz E. New method for statistical analysis of climate time series. Explora: Environment and Resource. 2025; 2: 6109.
5. Hamilton JD. Time Series Analysis. Princeton University Press. 1994.
6. Cybenko G. Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals, and Systems. 1989; 2: 303-314.
7. Durack P, Gleckler P, Landerer F, et al. Quantification des sous-estimations du réchauffement à long terme des eaux de surface océaniques. Nature Clim Change. 2014; 4: 999-1005.
8. Levitus S, Antonov J, Boyer T, et al. World ocean heat content and thermosteric sea level change (0-2000m), 1955-2010. Geo Res Lett. 2012; 39: L10603.
9. Ceppi P, Hartmann DL, Webb MJ. Mechanisms of the negative shortwave cloud feedback in high latitudes. J Climate. 2016; 29: 139-157.
10. Ceppi P, McCoy DT, Hartmann DL. Observational evidence for a negative shortwave cloud feedback in middle to high latitudes. Geophys Res Lett. 2016; 43: 1331-1339.

Appendix: Comparison between *MLPRegressor* and *Random Forest* for climate data modeling

Overview of *MLPRegressor*

MLPRegressor (Multi-Layer Perceptron Regressor) is an artificial neural network model used for regression tasks. It consists of one or more hidden layers of neurons, followed by an output layer. Each neuron applies a non-linear transformation (via an activation function such as ReLU) to a linear combination of its inputs.

MLPRegressor is based on the universal approximation theorem, which states that a neural network with a single hidden layer can approximate any continuous function with arbitrary precision, provided it has a sufficient number of neurons and a non-linear activation function.

A common issue with *MLPRegressor* is prediction capping—meaning the predicted values do not cover the full range of actual values.

Several reasons explain this phenomenon:

When data is normalized (e.g., using StandardScaler), extreme values are compressed into a reduced interval (typically [-1, 1] or [0, 1]). After prediction, the values are transformed back to the original scale. However, if the model primarily predicts within the normalized interval, the inverse transformation does not correctly restore the extreme values. For example, if the actual values of the target variable range from -2.22 to 2.17, but the model predicts within the [-1, 1] range after normalization, the final predictions will also be limited.

MLPRegressor uses the ReLU (Rectified Linear Unit) activation function, defined as $\mathrm{ReLU}(x) = \max(0, x)$. It has a derivative of zero for negative values and a constant derivative of 1 for positive values. This means that for very large or very small input values, the network may saturate, thereby limiting its ability to extrapolate beyond the range of the training data.

MLPRegressor uses gradient descent to minimize error (typically Mean Squared Error, or MSE). This method tends to prioritize central values—which are more numerous—at the expense of extreme values, which are rarer. Consequently, the model optimizes its weights to minimize overall error, which can lead to a smoothing of predictions for extreme values.

Climate data is often noisy due to natural variability or measurement errors. By minimizing MSE, *MLPRegressor* does not distinguish between signal and noise. Thus, the R^2 (coefficient of determination) is limited by the variance of the noise in the data. Mathematically, R^2 is defined as:

$R^2 = 1 - (\text{Variance of residuals} / \text{Variance of target variable})$

If the target variable y can be decomposed as $y = f(x) + \varepsilon$, where $f(x)$ is the signal and ε is the noise, then the variance of the residuals is at least equal to the variance of the noise. Therefore, even with a perfect model, R^2 cannot reach 1 if noise is present.

For discrete and noisy data, *MLPRegressor* cannot achieve an R^2 close to 1, as it cannot predict the random noise ε.

Introduction to *Random Forest*

Random Forest is a supervised learning model based on an ensemble of decision trees. Each tree is trained on a random subset of the data and features, and the final prediction is the average of the predictions from all the trees.

Unlike *MLPRegressor*, *Random Forest* does not require data normalization and can capture complex non-linear relationships without being constrained by activation functions.

Random Forest can capture extreme data values because it is not limited by activation functions or normalization. Each decision tree can learn local rules that capture specific behaviors, including those associated with extreme values.

By aggregating predictions from multiple trees, *Random Forest* is more robust to noise than *MLPRegressor*. Random noise has less impact on the final prediction because it is averaged out across the ensemble of trees.

Although *Random Forest* is generally robust against overfitting due to aggregation, it can still overfit if the trees are too deep or if the number of trees is excessive. This occurs when the model memorizes noise present in the training data, thereby degrading its performance on unseen data.

And while individual decision trees are interpretable, the ensemble of trees makes the overall model less transparent than simpler models like linear regression.

Practical limitations regarding R^2

R^2 (the coefficient of determination) is a normalized metric that measures the proportion of the dependent variable's variance explained by the model. Since it is defined as R^2 = 1 - (Variance of residuals / Variance of target variable), it is independent of the data scale. However, one cannot, for instance, directly compare the R^2 values of OC = f(UOS) and UOS = f(OC) to draw conclusions about the relative strength of the relationships. Indeed, the models used in this case differ: an R^2 of 0.63 obtained with *MLPRegressor* for OC = f(UOS) and an R^2 of 0.72 obtained with *Random Forest* for UOS = f(OC) do not allow us to conclude that one variable is "more of a driver" than the other. The differences in R^2 values may stem from the models employed (*MLPRegressor* vs. *Random Forest*) rather than the intrinsic strength of the relationships. Consequently, we simply state that the interactions between OC and UOS are significant in both directions.

Furthermore, while the universal approximation theorem guarantees that MLPs can approximate any continuous function with arbitrary precision, several factors limit this capability in practice when dealing with discrete climate data:

- Discrete and noisy data:

Climate data are discrete (sampled at regular intervals) and contain noise. *MLPRegressor* learns from these samples and cannot perfectly reconstruct the underlying continuous function, especially if the data are noisy.

- Limited extrapolation:

Although MLPs are universal approximators, their ability to extrapolate beyond the training data range is limited. If the training data do not cover the full range of possible values, the model will struggle to accurately predict extreme values.

- Suboptimal optimization:

The gradient descent method used to train MLPs can converge to local minima, meaning the model may fail to reach the best possible solution. This limits the model's overall accuracy.

Conclusion

MLPRegressor and *Random Forest* are two powerful regression models, yet they possess different strengths and weaknesses. *MLPRegressor* is limited by its inability to capture extreme values and its sensitivity to noise, whereas Random Forest is more robust but carries a higher risk of overfitting. In the case of discrete, noisy climate data, these limitations explain why MLPs typically achieve R^2 values well below 1.