SSRG International Journal of Civil Engineering
ISSN: 2348-8352/ https://doi.org/10.14445/23488352/IJCE-V12I11P106

Volume 12 Issue 11, 68-88, November 2025

Original Article

Thermo-Mechanical Pavement Deformation Prediction Using Monte Carlo Simulation and Regularized Neural Networks: A Comparative Study of FFNN and LSTM Models

Oumaima EL ABIDI[1], Mouna EL MKHALET[2], Nouzha LAMDOUAR[3]

[1,2,3]*Civil Engineering and Construction structure GCC laboratory, Mohammadia School of Engineers, Mohammed V University, Rabat, Morocco.*

[1]*Corresponding Author : oumaima_elabidi2@um5.ac.ma*

Received: 04 September 2025 Revised: 05 October 2025 Accepted: 03 November 2025 Published: 29 November 2025

Abstract *- This study addresses the challenge of predicting pavement performance under the combined influence of traffic-induced mechanical loads and daily thermal variations, a critical issue for road infrastructure in Morocco, where harsh climatic conditions and increasing traffic intensities exacerbate pavement deterioration. Traditional monitoring methods, such as visual inspections or simplified mechanical models, remain limited in their ability to capture the complexity and uncertainty inherent to thermo-mechanical interactions. In contrast, artificial intelligence methods, particularly neural networks, have shown strong potential for modeling nonlinear phenomena and improving predictive accuracy in pavement engineering. Building on this perspective, the present research develops a predictive framework that integrates two constitutive equations reflecting thermo-mechanical interactions, solved through deep learning architectures including feed-forward neural networks and long short-term memory networks, with and without dropout regularization. The study pursues a dual objective: to compare the predictive performance and robustness of these models, and to assess the reliability of their associated uncertainties, ultimately aiming to provide actionable insights for predictive pavement management and maintenance planning.*

Keywords *- Artificial Intelligence, Artificial Neural Network, FFNN, LSTM, Pavement Deformation, Pavement Performance Prevision.*

1. Introduction

1.1. Explanation

Reliable assessment of pavement condition is a strategic issue for the sustainability and quality of the road network. In Morocco, this need is accentuated by extreme climatic conditions and increasing road traffic, leading to complex thermomechanical stresses. Earlier studies have explored the use of neural networks to model pavement condition, particularly through indices such as PCI based on visual inspection data.

However, the majority of approaches remain deterministic, based on classical structural models or empirical correlations, without simultaneously considering mechanical and thermal effects, nor quantifying uncertainty critical for predictive planning in variable operational contexts [1]. In this context, some studies have used physical models such as the Boussinesq model coupled with Monte Carlo simulations to assess the deformation progression of flexible pavements, thus providing a reliable mechanistic basis. This approach provides a natural transition to our objective: to combine physical modeling (thermomechanical effects) and probabilistic approaches (predictive uncertainty) within an artificial intelligence framework.

Our research supports this approach. We propose two explicit formulations of Moroccan pavement deformation: one classical, the other incorporating a memory term ε(t–1), and model them using FFNN and LSTM networks, with and without Dropout regularization. We evaluate not only the predictive accuracy of these models, but also the reliability of the generated uncertainties, through statistical analyses, correlations with errors, and calibration of confidence intervals.

Preliminary results have shown that Dropout-regularized FFNNs offer better correlations between prediction errors and estimated uncertainties, while LSTMs enriched with the memory term partially capture the temporal dynamics but remain less reliable in terms of calibration of confidence intervals. This approach aims to identify a technologically reliable, scientifically based, and interpretable predictive approach to anticipate road deformation while integrating a level of uncertainty useful for proactive maintenance.

1.2. Novelty

The present study introduces a hybrid predictive framework combining Monte Carlo simulation and regularized neural networks (FFNN and LSTM) for thermo-mechanical pavement deformation prediction. Unlike previous works that separately addressed mechanical or thermal effects

Using either analytical or statistical methods, this study integrates both domains within a probabilistic-AI hybrid approach. The Monte Carlo simulation is used to capture the stochastic variability of key parameters (temperature, viscosity, load, and elasticity modulus), while the neural networks learn the nonlinear relationships governing the deformation evolution. This combination enables higher robustness and generalization capability compared to traditional regression-based or deterministic finite element approaches.

1.3. Originality

A comparative overview of related studies is presented in Table 1, emphasizing the methodological differences and the novelty introduced by the proposed hybrid Monte Carlo–neural network framework.

Table 1. Comparative overview of related studies

Study	Methodology	Region / Dataset	Model Type	Key Metric(s)	Added Value of the Present Study
Abd-elfattah et al. [2]	Crude Monte Carlo simulation of input uncertainties (material, temperature)	Egypt, 4-layer pavement section	Probabilistic	Reliability indices	Focuses on reliability only, not AI-driven prediction of deformation
Shatnawi et al. [3]	ANN vs regression for rutting prediction	Jordan, 33 highway segments	ANN	$R^2 \approx 0.82$	Uses ANN but without explicit thermo-mechanical coupling or Monte Carlo uncertainty modelling.
Cheng et al. [4]	ANN for rutting development in overlays, including climate, traffic	Canada LTPP database	ANN	Sensitivity/ML model	Good AI application, but focused on overlay pavement, not full thermo-mechanical + uncertainty coupling.
Present study	Monte Carlo simulation + Regularized FFNN and LSTM combining thermo-mechanical loads (temperature, viscosity, load, modulus)	Morocco, 1,000 simulations (synthetic dataset)	Hybrid AI-Probabilistic (FFNN & LSTM)	$R^2 = 0.95$ (FFNN) / 0.97 (LSTM)	Integrates thermo-mechanical variability + uncertainty via Monte Carlo + modern neural nets for deformation prediction over pavement lifetime

2. Methodology

To provide a clear overview of the adopted approach, a methodological framework has been structured into three main stages:

- The simulation of thermo-mechanical pavement behavior through Monte Carlo runs.
- The modeling of the generated dataset using different neural network architectures.
- The evaluation and comparison of the predictive performances.

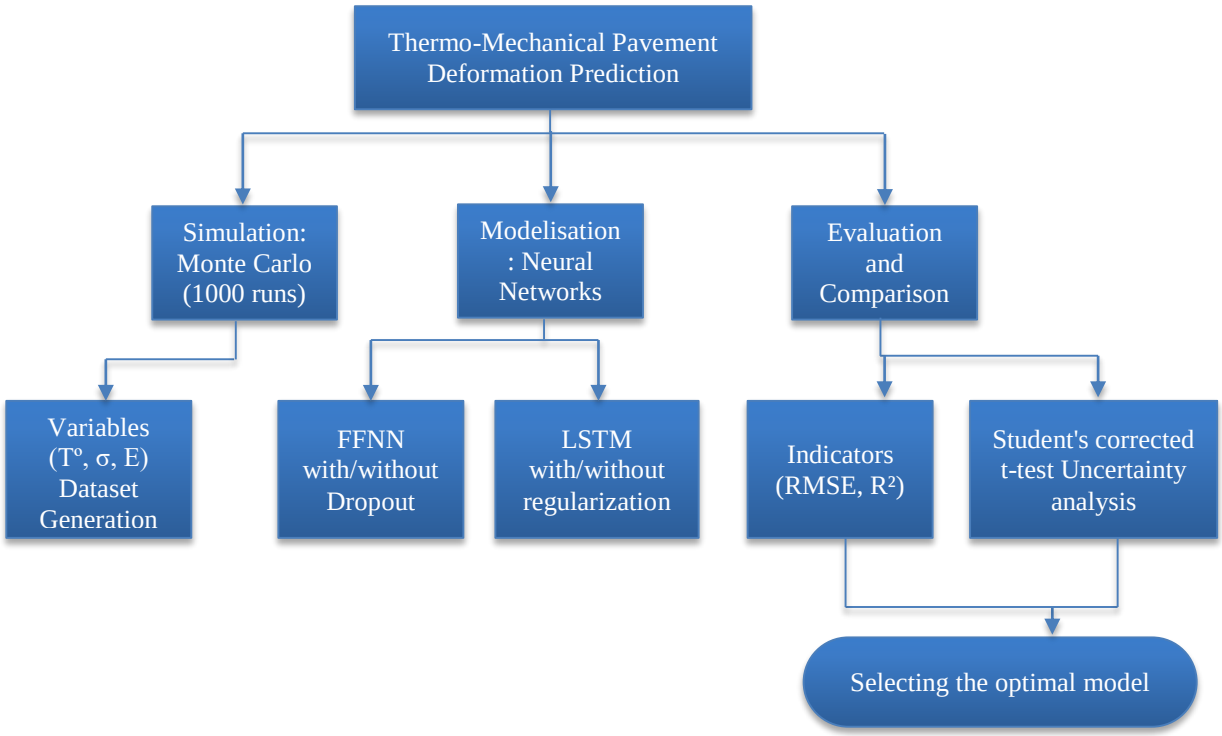

Fig. 1 Mind-map of the proposed methodology

The above figure presents a concept map summarizing these stages and their sub-components, highlighting the logical flow from data generation to model selection.

The prediction of Moroccan pavement deformation was approached by explicitly integrating the combined effects of mechanical traffic loads and thermal variations. Two constitutive formulations were adopted to capture these thermo-mechanical interactions. The First Equation (1) represents a direct thermo-mechanical relationship, while the Second Equation (2) introduces an alternative formulation in which the mechanical parameters-namely the Elastic Modulus (E) and the Viscosity (η)-are expressed as temperature-dependent variables.

To provide a robust training and testing basis for the models, a dataset of 1,000 samples was generated through Monte Carlo simulations. This process ensured that the inherent uncertainties associated with environmental conditions and traffic loading were adequately represented. Each Monte Carlo iteration randomly sampled the key thermo-mechanical variables within physically realistic bounds:

- Temperature (T) uniformly distributed between 5°C and 60°C to reproduce daily thermal cycles;
- Viscosity (η) log-normally distributed between 10^7 and 10^9 Pa·s, temperature-dependent through an Arrhenius-type relation;
- Elastic modulus (E) is usually distributed around 3 GPa ± 0.5 GPa;
- Traffic load (σ_0) is usually distributed around 800 MPa ± 10 %.

A total of 1,000 Monte Carlo runs were conducted, generating a diverse synthetic dataset that integrates the stochastic variability of environmental and loading conditions. This ensures that the resulting deformation values realistically reflect the thermo-mechanical Behavior of Moroccan flexible pavements under uncertainty.

The modeling process was carried out in two main phases. In the first phase, Equation (1) was addressed using two families of deep learning models: Feed-Forward Neural Networks and Long Short-Term Memory networks. Each architecture was tested with and without dropout regularization. This design allowed for the assessment of the dropout mechanism in mitigating overfitting and improving generalization. It also facilitated a comparison between FFNNs, which are well-suited for direct input–output mappings, and LSTMs, which are better equipped to capture sequential dependencies.

The generated dataset was randomly split into three subsets: 70 % training, 15 % validation, and 15 % testing subsets. All input variables were normalized to the [0, 1] interval to ensure uniform feature scaling, while deformation outputs were standardized based on their absolute maxima.

To enhance the robustness of the comparison, a 5-fold cross-validation scheme was implemented, allowing each subset to serve once as a test set. Model performance metrics were averaged across folds to reduce sampling bias.

In the second phase, the same modeling strategy was applied to Equation (2), again employing both FFNN and LSTM architectures. This additional formulation enabled the evaluation of how temperature-dependent mechanical parameters influence predictive performance. Preliminary

experiments revealed that FFNNs remained effective for straightforward mappings but exhibited limitations in capturing temporal memory effects. Conversely, LSTMs proved more adapted to such dynamics, especially when the strain term ε(t−1) was introduced as an exogenous input, thereby reflecting the viscoelastic memory of pavement materials.

Model accuracy and generalization were evaluated using standard performance metrics, including the Root Mean Square Error (RMSE) and the Coefficient of Determination (R^2)**:**

$$\begin{cases} RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \\ R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y}_i)^2} \end{cases}$$

These metrics quantify both the average prediction error and the goodness-of-fit between observed and predicted deformation values. The combination of statistical tests and performance indicators ensures an objective and reproducible comparison between FFNN and LSTM models.

Finally, the two constitutive formulations were systematically compared using rigorous statistical and probabilistic methods. Corrected Student's *t*-tests [5] were applied to evaluate the significance of differences in predictive performance. At the same time, uncertainty analysis was conducted to quantify the robustness and reliability of the models under varying conditions. This comprehensive methodological framework was designed to ensure a fair comparison of the proposed modeling strategies and to provide a solid foundation for subsequent predictive applications in pavement engineering.

Compared with previous approaches that treated thermal and mechanical effects separately or ignored parameter uncertainty, the proposed methodology provides a hybrid and probabilistic framework that integrates stochastic simulation with advanced neural architectures. This design enables both physical interpretability and data-driven adaptability, improving the realism and predictive power of deformation modeling for Moroccan pavement systems.

3. Results

3.1. State of the Art

This section reviews recent developments related to thermo-mechanical pavement deformation modelling, Monte Carlo simulation, and Neural Network Applications (FFNN and LSTM), as well as the main approaches to regularization in predictive models.

3.1.1. Thermo-Mechanical Deformation Studies

The deformations experienced by flexible pavements under combined thermal and mechanical loads have long been a critical issue in pavement engineering. Temperature variations induce changes in material properties (elastic modulus, viscosity) and generate thermal gradients, which, in combination with moving vehicle loads, lead to complex stress-strain responses and permanent deformations (rutting, warping). Early mechanistic-empirical models treated thermal and mechanical effects separately or in a simplified additive manner.

In more recent years, advanced numerical approaches - especially Finite Element Methods (FEM) and multiscale simulations - have emerged to capture the coupled thermo-mechanical (T–M) behaviour more accurately. For example, a 2024 3D multiscale model by Gong et al. analysed seasonal temperature variations and their influence on long-term deformation of bridge-deck pavements [6].

Similarly, Li et al. [7] investigated thermal–mechanical coupling on long longitudinal slopes of asphalt pavements and demonstrated that temperature gradients significantly amplify deformation rates.

Review studies underline that while many models address either the thermal or mechanical aspect, relatively few fully integrate both fields, plus stochastic variability of inputs. For instance, Joumblat et al. [8] highlight that most permanent deformation models still treat thermal effects as boundary conditions rather than as dynamically interacting fields.

These advancements establish a strong foundation, yet two main gaps remain:

- The need to consider probabilistic uncertainty in thermal-mechanical parameters,
- The integration of data-driven predictive models that can exploit large simulation or field datasets.

The present study addresses these by combining Monte Carlo simulation of thermo-mechanical input variability with regularised neural networks (FFNN and LSTM) to predict pavement deformation under realistic T–M coupling.

3.1.2. Artificial Neural Networks

Used to imitate the human brain, Artificial Neural Networks (ANN) are computer-based models designed to mimic the functioning of biological neurons. To grasp complex patterns or extract rules from large and complicated datasets, many machine learning researchers are turning to ANN.

Artificial neural networks are computer systems made up of numerous simple and densely connected processing units, which handle information by updating their internal state in reaction to external inputs [9]. Each neuron making up the network is connected to other neurons by directed links, and each directed link has a weight associated with it. The weights acquired during the training process represent information extracted from the dataset, which is used by the network to solve a particular problem [10].

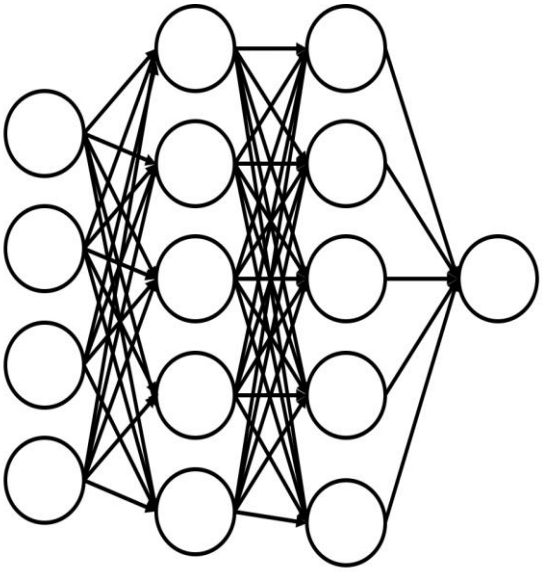

Fig. 2 Regular Neural Network

However, due to their high learning capacity, neural networks are particularly susceptible to overfitting, a phenomenon where the model excessively memorizes training data at the expense of its ability to generalize to new data. Overfitting is a key issue in supervised learning, arising from noisy data, limited training samples, and the use of highly complex classifiers [11]. Regularization is a key element of machine learning [12], as it allows good generalization to unseen data, even when training is done on a finite dataset or with an insufficient number of iterations. Good regularization is necessary for the successful application of neural networks, and as an example of this technique, we have dropout.

This Dropout technique refers to the process of dropping out units of neural networks. This means removing it from the network, along with all its incoming and outgoing edges [13].

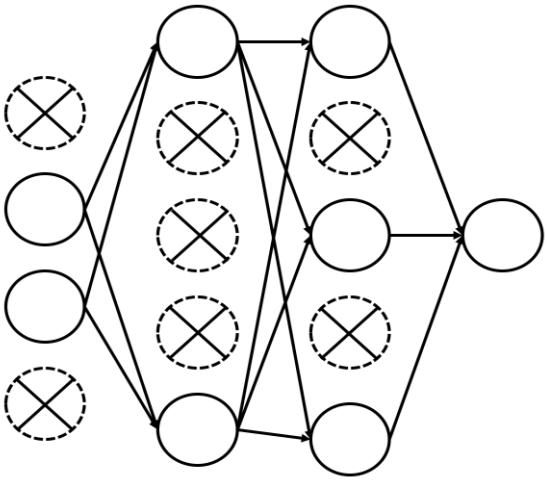

Fig. 3 After applying dropout

3.1.3. Long Short-Term Memory

LSTMs, or Long Short-Term Memory, are a type of neural network that was proposed in 1997 by Hochreiter and Schmidhuber [14]. Their primary purpose is to handle better information that changes over time, such as in time series. Unlike classic neural networks, LSTMs can "remember" important information over a long period of time thanks to a system of gates (such as input, output, and forget gates). This makes them very effective in fields such as speech recognition, traffic forecasting, and hydrology. Several researchers, such as Greff et al. [15], have compared different versions of LSTMs and confirmed that certain parts of their structure are indeed essential for good performance. Even though newer, more powerful models, such as Transformers, have recently emerged, LSTMs remain a reliable and widely used method for processing time-dependent data. Zhang et al. [16] have designed an LSTM model fused with multi-head attention (LSTM+MA) to predict the International Roughness Index (IRI), which measures road roughness. By combining traffic, climate, and maintenance history, this model achieves a correlation of 0.965 with real data, much higher than traditional methods.

In 2022, Mers et al. [17] conducted an extensive analysis of 31 years of pavement data from the Florida Department of Transportation (1989–2020), covering 7,615 segments and over 42,000 miles of roads. They compared several Methods: Linear Regression (MLR), Fully Connected Networks (FCNN), RNN, GRU, LSTM, and a hybrid LSTM-FCNN model. They concluded that LSTM, due to its temporal gates, better captures the progression of road deterioration. Combining it with an FCNN network benefits from both temporal memory and static feature processing capability, resulting in the most accurate predictions for proactive infrastructure management.

3.1.4. Monte Carlo Simulation

Monte Carlo simulation is a technique that utilizes repeated random sampling and statistical analysis to derive results. It closely relates to random experiments, where the specific outcomes are not predetermined. In this approach, we first identify a statistical distribution for each input parameter, which serves as the basis for generating random samples. These samples represent the values of the input variables. For each combination of input parameters, a corresponding set of output parameters is produced. Each output parameter reflects one possible outcome of the simulation run. By conducting multiple simulation runs, we collect a range of output values. Ultimately, we conduct statistical analyses on these output values to inform decision-making regarding the next steps. The sampling statistics derived from the output parameters allow us to characterize the variation in the results [18].

3.1.5. Regularization Techniques

To address the overfitting problem, various regularization techniques have been developed. These methods aim to control the model's complexity and promote its generalization capacity by introducing constraints or modifications to the training. They have become essential in modern deep neural network architectures, particularly in fields such as computer

vision, natural language processing, and recommendation systems [19]. The different techniques can be summarized in the following table with their advantages and disadvantages:

Table 2. Advantages and disadvantages of regularization techniques

Name of the technique	Description	Advantages	Disadvantages
Regularization L1[19]	Adds a penalty to the sum of the absolute values of the weights. Encourages sparsity (variable selection).	Encourages sparsity, useful for trait selection.	May cause loss of information if too aggressive; not differentiable at 0.
Regularization L2 [20]	Penalizes the sum of squares of the weights. Reduces the magnitude of the weights without canceling them out.	Reduces weight amplitude, stabilizes learning.	Does not favor variable selection.
Dropout [21]	Randomly removes neurons during training, preventing co-adaptation of units.	Reduces overfitting; improves robustness.	Extends training time, potentially slowing convergence.
Monte Carlo dropout [22]	Bayesian version of dropout, also used in inference to quantify uncertainty.	Allows probabilistic predictions; good Bayesian compromise.	Sometimes underestimates uncertainty for out-of-distribution inputs; costly in inference.
Early Stopping [23]	Stop training as soon as performance on the validation set stops increasing.	Simple, effective, and prevents overfitting.	Requires a validation game; it depends on the criterion being followed correctly.
Data Augmentation [24]	Artificially generates more data (transformations, noise, rotations, etc.) to generalize better.	Improves generalization, artificially increases the dataset.	It can introduce irrelevant noise if misused.
Batch Normalization [25]	Normalizes intermediate activations, which also has a regulating effect.	Stabilizes and accelerates training, regulating effect.	Unstable behavior with small batches or RNN.
Weight Constraint [19]	Limit the weight norm (eg, max-norm regularization).	Controls model capacity; simple to implement.	Less common, it requires manual tuning.
Label Smoothing [26]	Replace “hard” labels (0 or 1) with softer values (e.g., 0.9 / 0.1) to avoid overconfidence.	Reduces model overconfidence; improves calibration.	May interfere with the interpretation of outputs.
Noise Injection [27]	Adds noise to input data, weights, or activations to improve robustness.	Promotes robustness and regularization.	It can disrupt learning if poorly calibrated.
Mixup / CutMix [28]	Recent data augmentation techniques consist of mixing images and labels to generalize better.	Improves robustness, interpolation, and reduces overfitting.	Less intuitive; may hinder interpretability.
Structural regularization (pruning, distillation) [29]	Removes neurons or compresses the model while maintaining its performance (secondary regularizing effect).	Reduction in complexity without significant loss of accuracy.	More complex implementation; may require retraining.

3.1.6. Monte Carlo Dropout

Context

Monte Carlo Dropout (MCD) is an advanced method that measures uncertainty in neural network predictions, based on the principles of Bayesian inference [30]. It employs dropout not only as a regularization technique, but also as a purposeful approach to approximate the posterior distribution of network weights. Throughout the training stage, dropout randomly inactivates portions of neurons, progressively refining the model across iterations. This injected randomness helps limit overfitting by preventing neurons from co-adapting [31]. By integrating MCD, the model gains an awareness of uncertainty, enabling more reliable and informative predictions that reflect the inherent uncertainties in the data. This uncertainty-aware approach is essential for enhancing the robustness and credibility of predictive models [32].

Dropout is used to prevent overfitting and offers an efficient means of approximately blending a vast number of different neural network architectures. The term "Dropout" pertains to the temporary removal of units (both hidden and visible) from a neural network. When we say a unit is dropped out, we mean that it, along with all its associated incoming and outgoing connections, is temporarily excluded from the network [21]. The units that are dropped are chosen randomly. In the most basic scenario, as shown in Figure 4 below, every unit is kept with a set probability p that is independent of other units.

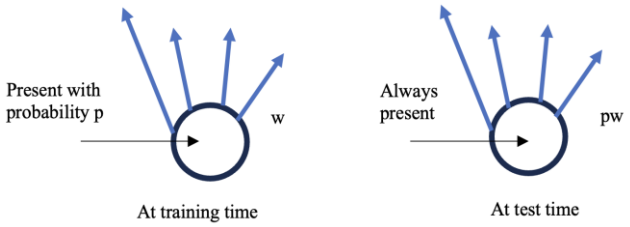

Fig. 4 Dropping out a unit [21]

During training, a unit is present with a probability of p and connects to the next layer with weights w. During testing, the unit is always present, and the weights are adjusted by multiplying by p. Consequently, the test output matches the expected output from training.

Application

Concretely, the process proceeds as follows: once the model is trained with the dropout mechanism active, N stochastic passes (typically 20 to 100) are carried out on the same test example. In each pass, specific units of the network are randomly deactivated according to the defined dropout probability. This generates a series of different predictions. These predictions are then used to calculate a mean (Central Prediction) as well as a Standard Deviation, which is used to quantify the uncertainty of the model for this example [22]. In the following figure, a flowchart explaining the process is presented:

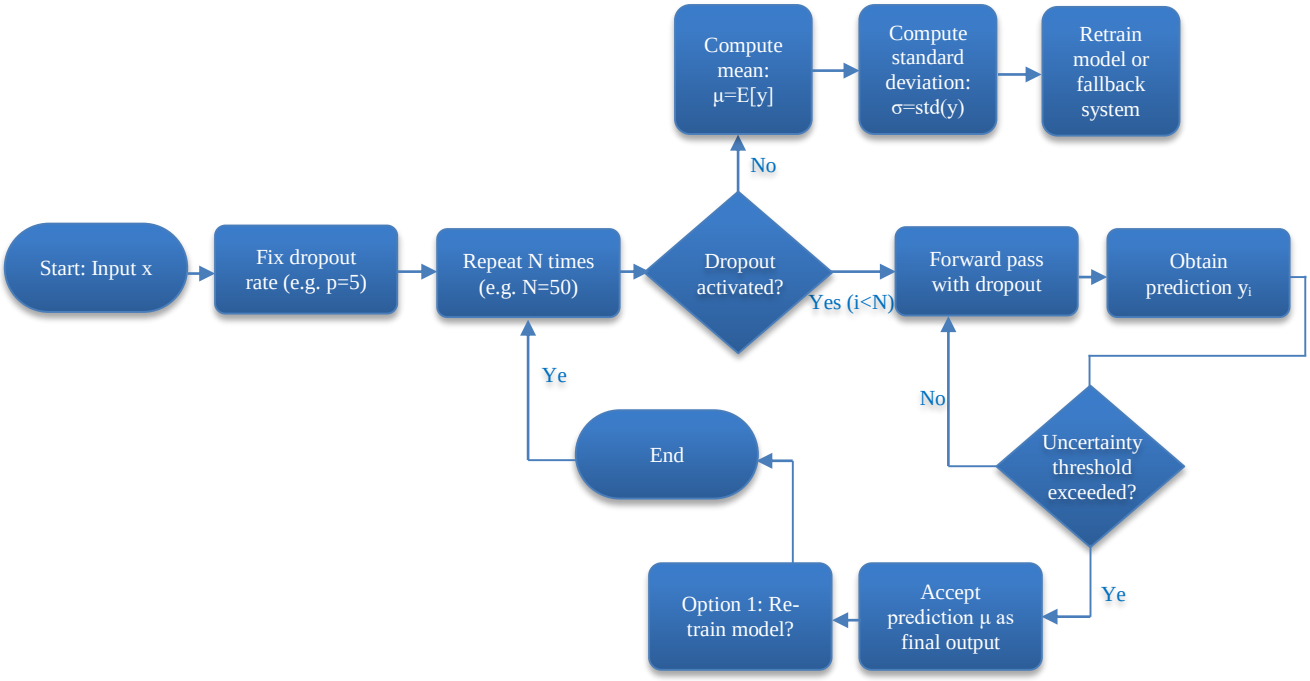

Fig. 5 Flowchart of Monte Carlo dropout

The flowchart in Figure 5 describes the operation of Monte Carlo Dropout (MC Dropout). The process begins with an input x, for example, a test sample.

Next, a dropout rate is set (for example, p=0.5), i.e., the probability with which some neurons will be deactivated during inference.

Then, the decision is made to repeat the operation N times (for example, N=50). This means that 50 passes are made on the same sample, each time with a different dropout mask.

At each iteration:

- We check whether the dropout is enabled.
- If so, we perform a forward pass with the dropout enabled.
- This produces a prediction yi, which will be stored.
- Once the N passes are complete:
- We calculate the mean of the predictions (denoted μ=E[y]): this is the model's final prediction.
- We also calculate the standard deviation of the predictions (denoted σ=std(y)) to measure the uncertainty.

We then check whether the uncertainty exceeds a predefined threshold:

- If yes, this means that the model is not sufficiently confident. Two possibilities then arise: either we retrain the model or we use a backup system.

- If no, then the mean μ is accepted as the final output.

In all cases, the process ends at this point.

Srivastava et al. [21] applied dropout to classical neural networks to examine its impact on various fields: object recognition, handwritten digit recognition, speech recognition, document classification, and computational biology. In the past few years, theories and methods related to structural reliability have evolved considerably, and they now represent a valuable approach for rationally assessing the safety of complex or unconventional structural systems. Furthermore, recent advances suggest that their use will continue to expand, even for typical structural configurations. By extending dropout to graphical models such as Restricted Boltzmann Machines (RBMs), Dropout RBMs were developed, and they have empirically shown promising Behavior. Hinton et al. [22] explored a theoretical interpretation of dropout by comparing it to L2 regularization in the case of linear regression. They concluded:

- a significant improvement in the performance of neural networks;
- a reduction in overfitting by disrupting the learning of fragile co-adaptations between neurons;
- an increase in training time: models take 2 to 3 times longer to train because the model changes randomly at each iteration, making gradients less stable;
- Proposal of an alternative to avoid this slowness: replace the stochastic dropout with an equivalent deterministic regularizer, at least in simple cases such as linear regression.

Another example from Thaler et al. [31] can be cited; they used Dropout Monte Carlo (DMC) as a method for estimating Uncertainty (UQ) in GNN predictions. This approach involves enabling dropout both in training and inference, thus generating a prediction distribution and estimating uncertainty. From this research, they concluded that there is good generalization in the same domain (large MOFs of the same type) and that even when the structure of MOFs is different (e.g., ARC-MOF, IZA-Zeolite data), DMC generally succeeds in reporting high uncertainties, which is useful for alerting about the limits of prediction reliability. In addition, Liu et al [32] explored the Monte Carlo dropout by integrating it into the Multi-Fidelity Deep Neural Network (MFDNN) model, which increased the accuracy of predictions of the risk level of Retrograde Erosion of Dikes (BEP). The model studied outperforms four advanced machine learning models, particularly in the context of limited data. In this respect, the MC Dropout allows a probabilistic evaluation of predictions, i.e., it quantifies the uncertainty associated with each model prediction, which is particularly important in a critical area such as flood risk management, as it helps distinguish between situations with high uncertainty (requiring more attention) and those with low uncertainty. While exploring the application examples of Monte Carlo Dropout, it is necessary to present its advantages and disadvantages in the following table:

Table 3. Monte Carlo dropout advantages and disadvantages

Aspect	Advantages	Disadvantages
Uncertainty quantification [22]	Allows estimation of epistemic uncertainty (model uncertainty) using a standard network with dropout.	Less effective for random uncertainty, especially on noisy data
Simplicity of implementation [33]	Easy to integrate into already trained models with dropout; no re-architecture required.	Requires keeping dropout enabled during inference, which is non-standard.
Bayesian approximation [34]	Allows approximation of the Bayesian process without training expensive true Bayesian models.	This is only an approximation, sometimes not very precise on extreme or out-of-distribution releases.
Computing efficiency [35]	Less expensive than classical Bayesian approaches like Gaussian Processes or Bayesian NNs.	Each prediction requires N forward passes (e.g., 20 to 100), which increases inference time.
Multi-domain adaptability [21]	Used effectively in vision (ImageNet), GNN, NLP, biomedicine, etc.	May underestimate uncertainty if activations are insensitive to dropout.
Calibration [36]	Allows better calibration of predictions than a network without regularization.	Sometimes overconfident in areas with little training data (out of distribution).
Versatility [37]	Works with CNNs, LSTMs, GNNs, and Transformers.	The effect of dropout depends on the architecture (e.g., less effective in RNNs without specific adaptation).

3.2. Case Study

To continue the case study and follow the work of the author's previous work in the case of sizing the road network in the Moroccan kingdom. This study focuses on the combination of thermosensitivity and viscoelasticity of the road material, and in particular, bituminous coatings. The proposed formula is as follows [38]:

$$\epsilon(t) = \frac{\sigma_0}{E} + \frac{\sigma_0 t}{\eta} + \alpha(T(t) - T_0 \qquad (1)$$

- ϵ(t): total strain at time
- σ0: applied mechanical stress (constant in this model)
- E: modulus of elasticity (elastic stiffness of the material)
- η: viscosity (resistance to strain over time)
- α: coefficient of linear thermal expansion
- T(t): temperature at time
- T0: reference temperature (usually initial or ambient temperature)

The equation used is a combination of the Maxwell Model (elastic + viscous in series) and a linear thermal correction, which is used in modeling the behavior of asphalt pavements to:

- Simulate creep under constant load (stationary or slow-moving trucks);
- Incorporate the effect of daily or seasonal temperature variations.
- Calibrate numerical or analytical models used in design or analysis tools.

This equation combines three effects:

- Instantaneous elastic Behavior: The term $\sigma 0/E$ represents the immediate response of the material to the applied load. Typical of a Hooke spring.
- Viscous Behavior (Creep): The term $\sigma 0t/\eta$ reflects an increasing deformation over time, typical of viscous creep (like a dashpot). The longer the load lasts, the greater the deformation.
- Thermal effect: The term α(T(t)−T0) reflects the thermal expansion or contraction due to the temperature change. If the temperature increases, the pavement expands.

Therefore, to apply the above-mentioned equation, we take the case of a road section subjected to mechanical and thermal effects over time as modeled in the following figure:

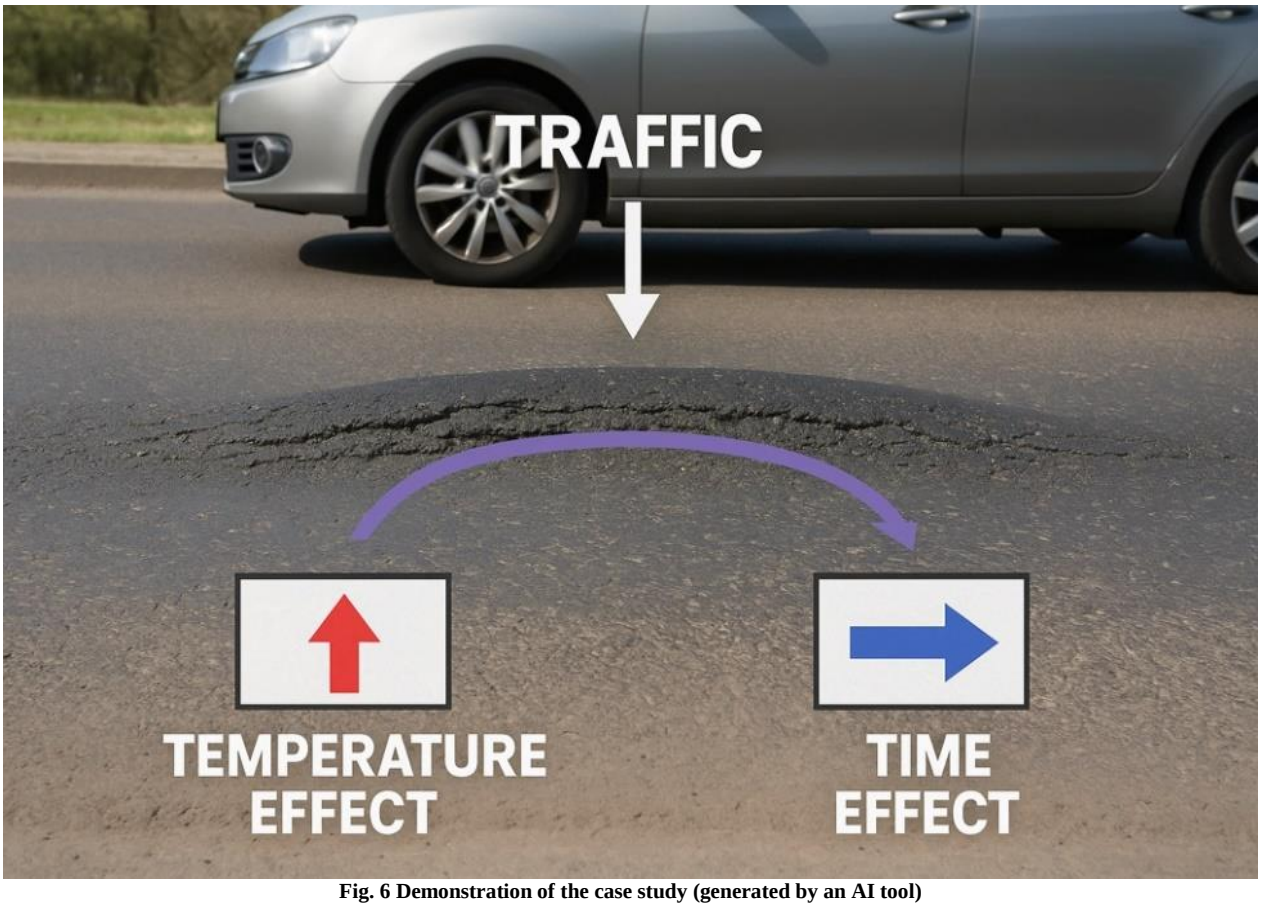

Fig. 6 Demonstration of the case study (generated by an AI tool)

The example studied is a road section subjected to a mechanical load ranging from 655MPa to 906MPa. For a period of one month, a summer month to be precise, the temperature varies by 20°C (night) and 40°C (day) over 24 hours; with this variation, the Young's modulus also varies.

Using a MATLAB script, we can model a time series that shows how the viscoelastic deformation ϵ(t) of the asphalt mix evolves over 30 days, under the modeled variable loads and temperatures.

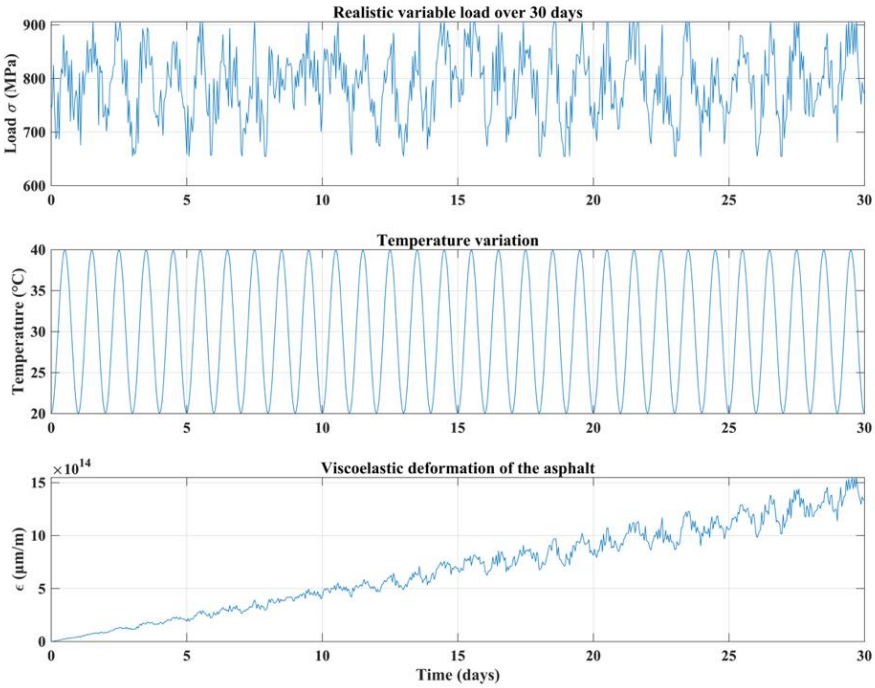

Fig. 7 A time series of the viscoelastic deformation ϵ(t) of the Asphalt mix

Given the complexity of property management, modeling is done with neural networks that have the ability to model these complex relationships better than simple linear models. Once trained, it can provide deformation predictions in real time or for variable conditions very quickly, which is useful for optimization, control, or predictive maintenance applications.

By applying an FFNN on MATLAB, with an architecture of 3 hidden layers with (30, 20, and 10 neurons respectively), 3 inputs (load, temperature, and time of day), and one output (deformation), the comparison of actual and predicted values can be visualized in the following figure:

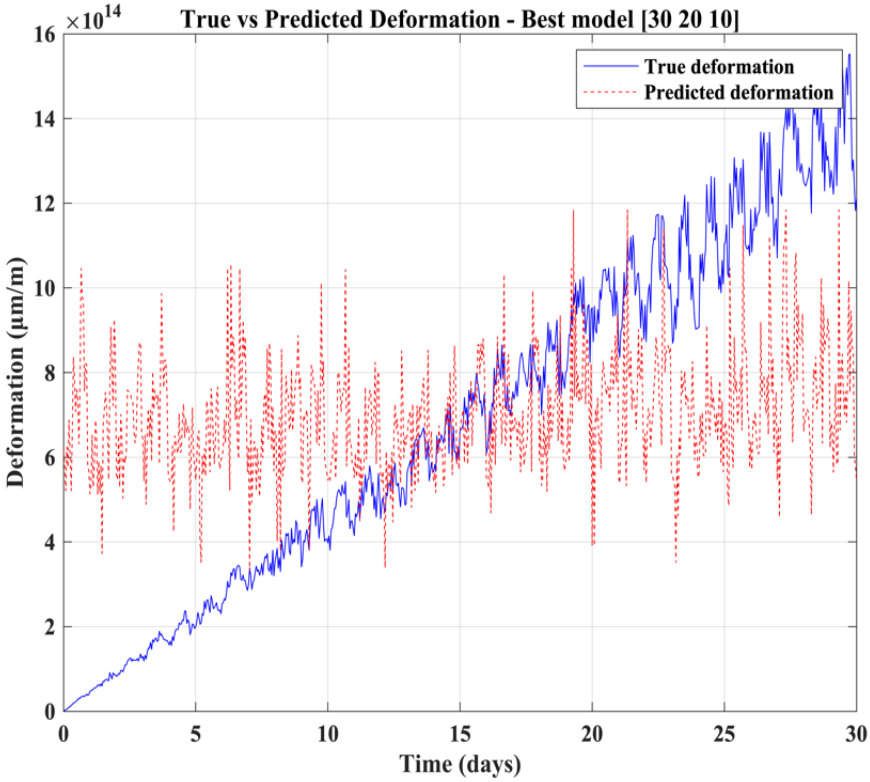

Fig. 8 True vs Predicted values of deformation using an FFNN

However, using this type of model, we obtained an RMSE error value of 0.52274, which is relatively unacceptable. This can be justified by the need to take into account the entire recent history of these conditions (e.g., accumulated load, thermal evolution over time). The FFNN model only "sees" a given instant t; it has no intrinsic memory of previous instants.

Therefore, it is proposed to use an LSTM model that captures these temporal dependencies, thus better modeling the cumulative and delayed Behavior of the deformation.

Unlike FFNNs, which process each piece of data independently, LSTMs are able to take into account the order and relationship between data over time. This temporal memory allows them better to model progressive phenomena, such as road deterioration.

Following the application of the LSTM, the error obtained is of the order of 0.3499, which is practically less than the value of the FFNN. In the following figure, we can model the comparison of the actual and predicted values:

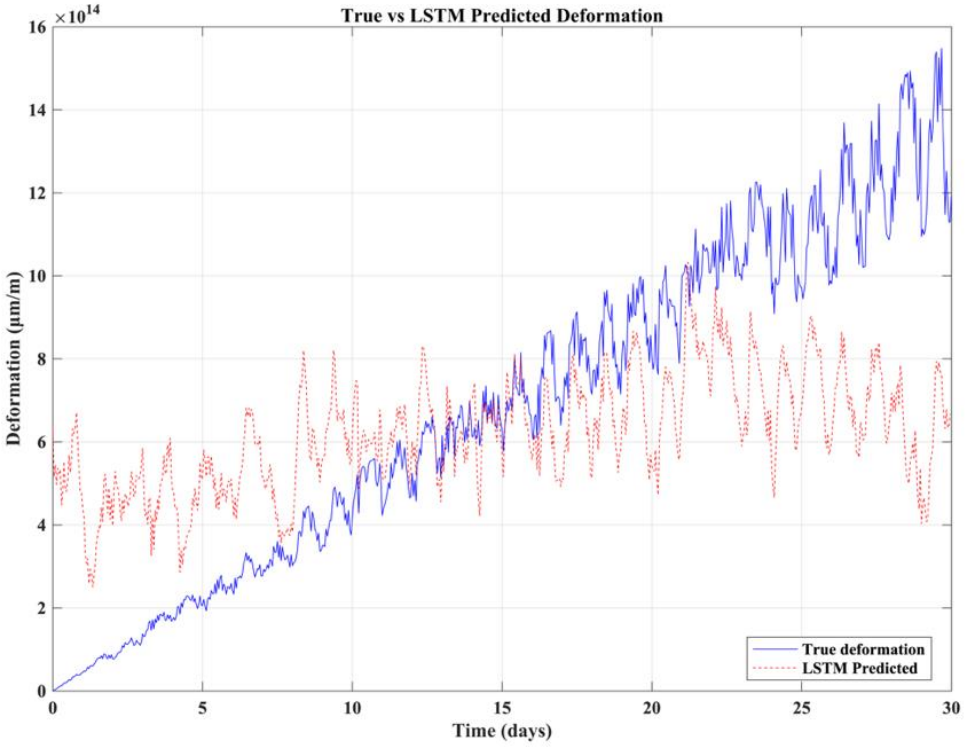

Fig. 9 True vs Predicted values of deformation using an LSTM

3.3. Results

Following the case study, the modeling is done on a road section whose parameters defining its performance are summarized in the following table:

Table 4. Modeling parameters

Input	Variation
Applied Mechanical Stress σ_0	655Mpa to 906Mpa
Temperature T	20°C (night) and 40°C (day)
Young Modulus E	1 to 5 GPa
Time t	0 seconds to 24 hours
Approximate viscosity η	10^7 to 10^9 Pa.s

Following the Monte Carlo simulation on MATLAB, here is a box plot of the deformation at different times.

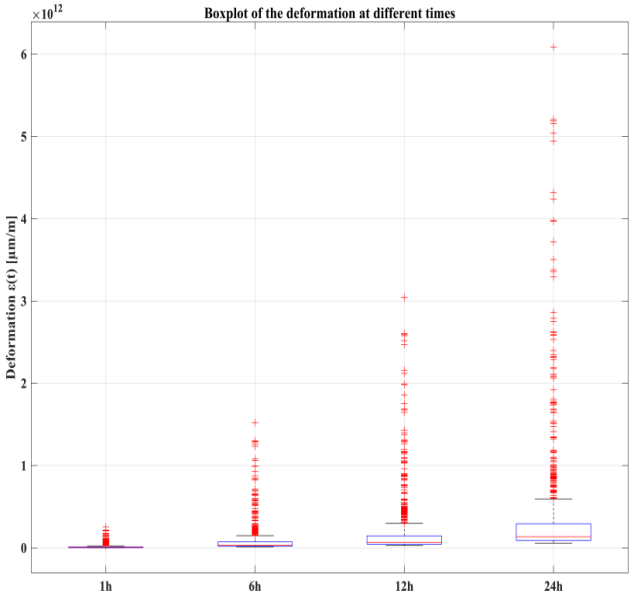

Fig. 10 Boxplot of the deformation at different times

This graph shows how the strain evolves over time, with a general increase in strain as time passes. For each time point, the following can be observed:

- 1h: At 1h, the strain is very low, the box is close to zero, and outliers are rare.
- 6h: At 6h, the strain increases, the box widens, and there are a few outliers.
- 12h: At 12h, the strain continues to increase, with an even greater spread of values, indicating greater variability in the strain measurements.
- 24h: At 24h, there is a sharp increase in strain, and outliers are numerous, suggesting that there are specific phenomena or conditions affecting the strain at that time.

Modeling by Artificial Neural Networks:

Following the modeling by artificial neural networks, to choose the best architecture, four architectures were tested, and then, the comparison of RMSE errors can be visualized in the following figure:

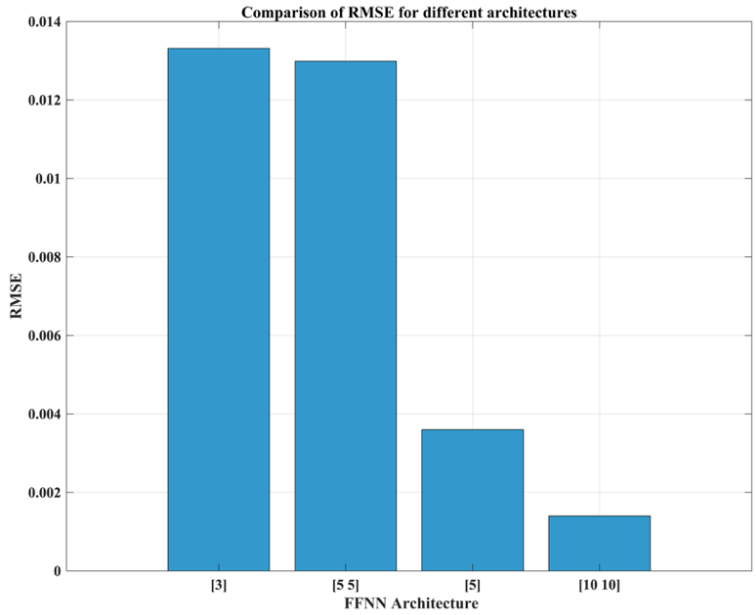

Fig. 11 Comparison of RMSE for different architectures

The figure above shows that simple architectures like [3] and [5 5] give a relatively high RMSE (~0.013). While architecture [10 10] is clearly the best performing with an RMSE < 0.002. This confirms that deeper and wider architectures significantly improve the model accuracy. Following these results, the architecture [10 10] can be chosen, and a visualization of the predicted and actual values is shown in the following figure:

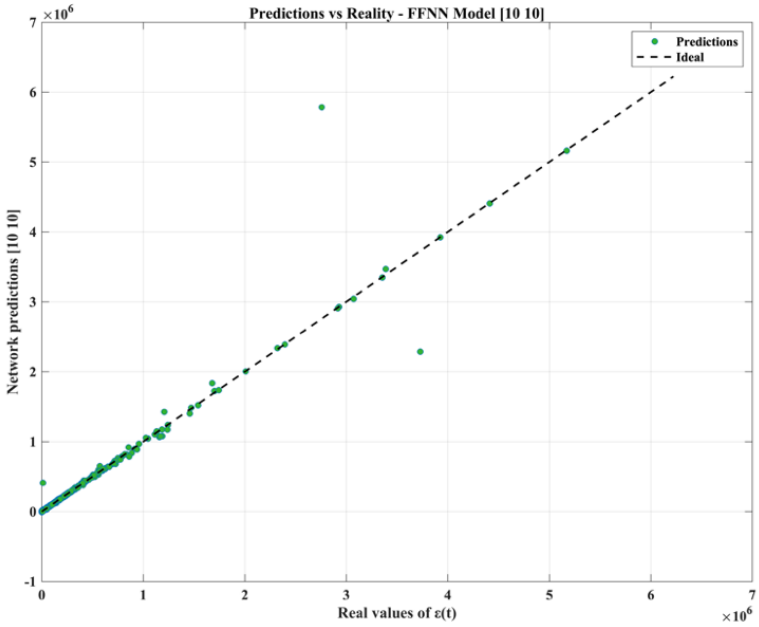

Fig. 12 Comparison of prediction vs Reality values

By applying the Monte Carlo Dropout on our FFNN model with the architecture [10,10], the RMSE errors obtained are as follows:

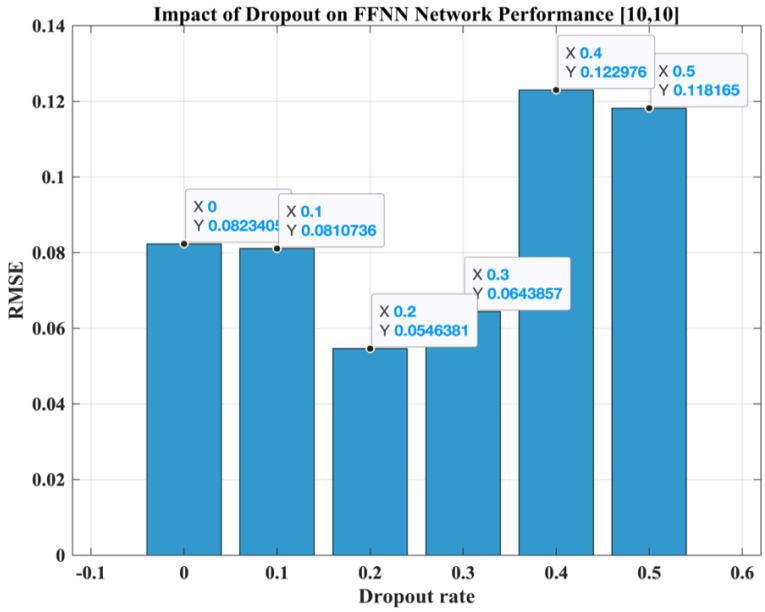

Fig. 13 Impact of dropout on FFNN model

According to the figure above, we can observe that a dropout of 0.2 gives the best performance (RMSE ≈ 0.055). Beyond 0.3, the RMSE increases significantly (>0.12 for 0.4). On this, we can say that the moderate dropout (0.2) improves the generalization, but too high rates degrade the performance.

Given the nature of the problem, and as previously explained, LSTM modeling is necessary due to the temporal dependence of the equation. Four architectures are tested, each with the following properties:

Table 5. LSTM used architectures

Architecture	Description
A1_LSTM_50	1 LSTM layer (50 units)
A2_LSTM_100	1 LSTM layer (100 units)
A3_LSTM_50x2	2 LSTM layers (50 + 50)
A4_LSTM_100x50_dropout	2 layers (100 + 50)

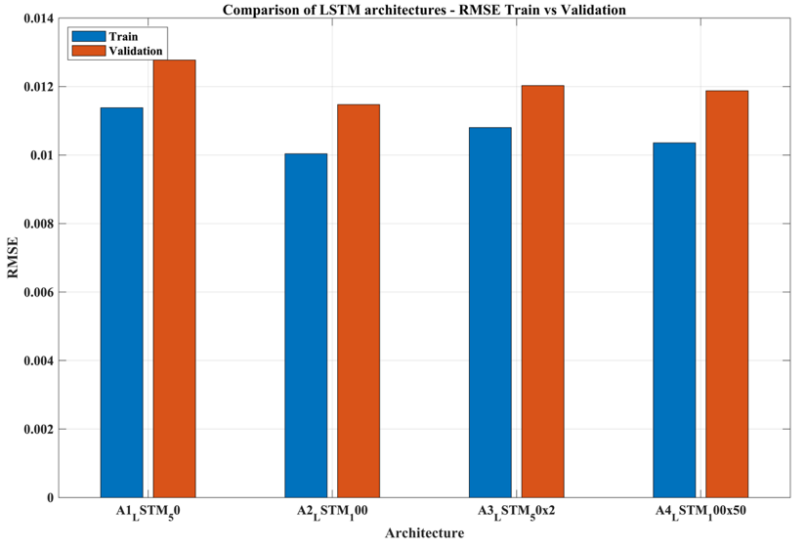

Fig. 14 Comparison of LSTM architectures

Based on these results, the following interpretations can be made:

- A2 (LSTM 100) overestimated the problem complexity, which likely caused overfitting or poor generalization.
- A3 (two LSTM layers without regularization) performed better than A1 and A2 → depth seems beneficial.
- A4 (two layers with dropout) yielded the best RMSE, likely due to better regularization.

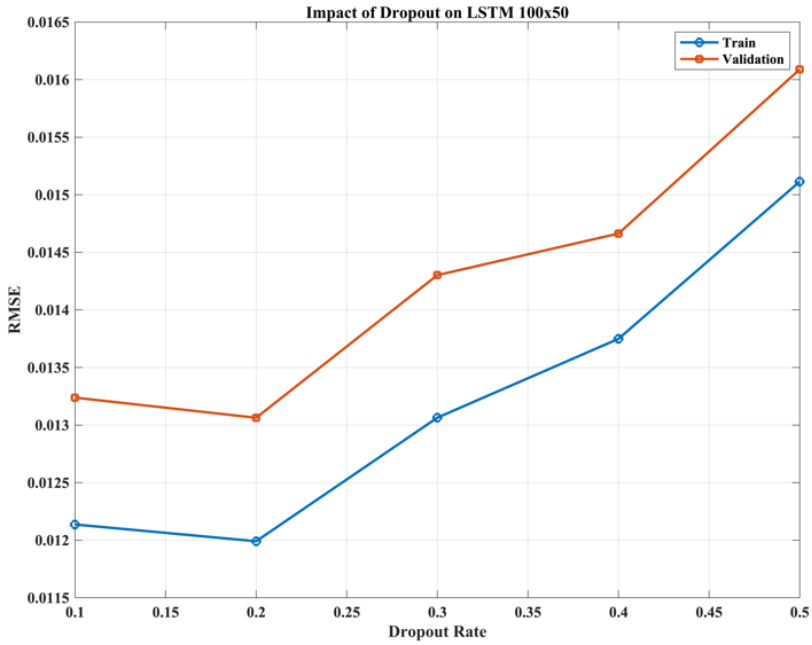

Fig. 15 Impact of dropout on LSTM

Unlike the FFNN, here, according to the figure above, the dropout degrades the performance: the RMSE increases from 0.012 to 0.016 when the rate increases from 0.1 to 0.5. This suggests that the LSTM architecture studied is already sufficiently regularized and that the dropout is not beneficial.

When we trained the LSTM model with a dropout rate of 0.2, we obtained an RMSE of approximately 0.013 on the training data and 0.012 on the validation data. The fact that these two values are close shows that the model generalizes well, without over-adapting to the specific details of the training data.

Initially, the FFNN and LSTM models were trained without any specific regularization. These tests yielded very low validation errors (≈0.01 in normalized values), but this apparent performance actually reflected overfitting: the models perfectly reproduced the training data but lost generalization ability, with unstable differences between the test and validation sets.

To address this limitation, we adopted an option that consists of retaining the current physical equation of linear creep with thermal effect, while improving the training framework with regularization techniques. These include data normalization (log + z-score), L2 penalization (weight decay), batch normalization, moderate dropout (0.10–0.25), and early stopping. The goal was not to artificially reduce error, but to obtain more stable and generalizable models.

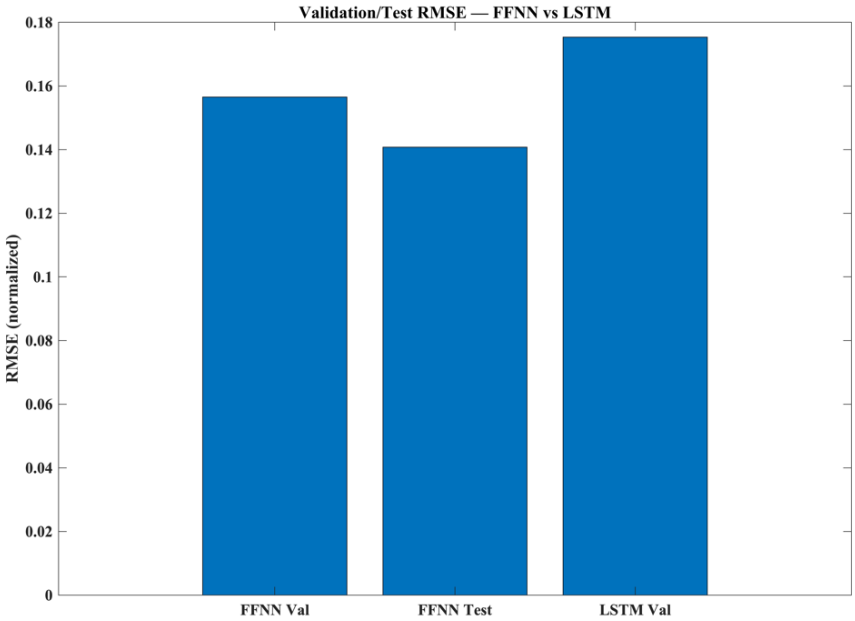

Fig. 16 RMSE FFNN vs LSTM

According to the figure above, the results obtained actually show higher RMSE values (≈0.15–0.17 in normalized, or 50,000–65,000 in original scale) compared to the first experiments, but this time the errors are consistent between validation and test, which reflects better robustness of the model. In other words, this option makes it possible to limit overfitting and to strengthen the reliability of predictions, even if the raw precision remains limited by the simplicity of the physical equation used.

In summary, this option ensures reliability better than raw performance. If the goal is to reduce absolute error significantly, we will then need to consider Option B, which consists of changing the physical equation to a richer model.

4. Discussion

4.1. Test Part

In the initial application, the equation used is:

$$\epsilon(t) = \frac{\sigma_0}{E} + \frac{\sigma_0 t}{\eta} + \alpha(T(t) - T_0) \qquad (1)$$

Based on the data in the equation, the linear approximation of creep does not capture:

- the transient phase (delayed creep),
- the separation between reversible and irreversible deformation,
- nor the true temperature dependence.

Therefore, even with optimal regularization (Option A), the errors remain high, indicating that the physical formulation is too simple.

On this, we propose the most widely used model for bituminous mixes, which combines permanent creep ($\eta 1$) and delayed creep ($E2$,$\eta 2$), the Burgers Model (Maxwell + Kelvin). It easily allows the dependence of E and η on temperature (via Arrhenius or WLF-type laws).

$$\epsilon(t) = \frac{\sigma_0}{E_1} + \frac{\sigma_0 t}{\eta_1} + \frac{\sigma_0}{E_2}(1 - e^{-t/\tau_2} + \alpha(T(t) - T_0) \quad (2)$$

For a step charge $\sigma 0$ and any temperature $T(t)$, we can calculate $\varepsilon(t)$ as shown in the following figure:

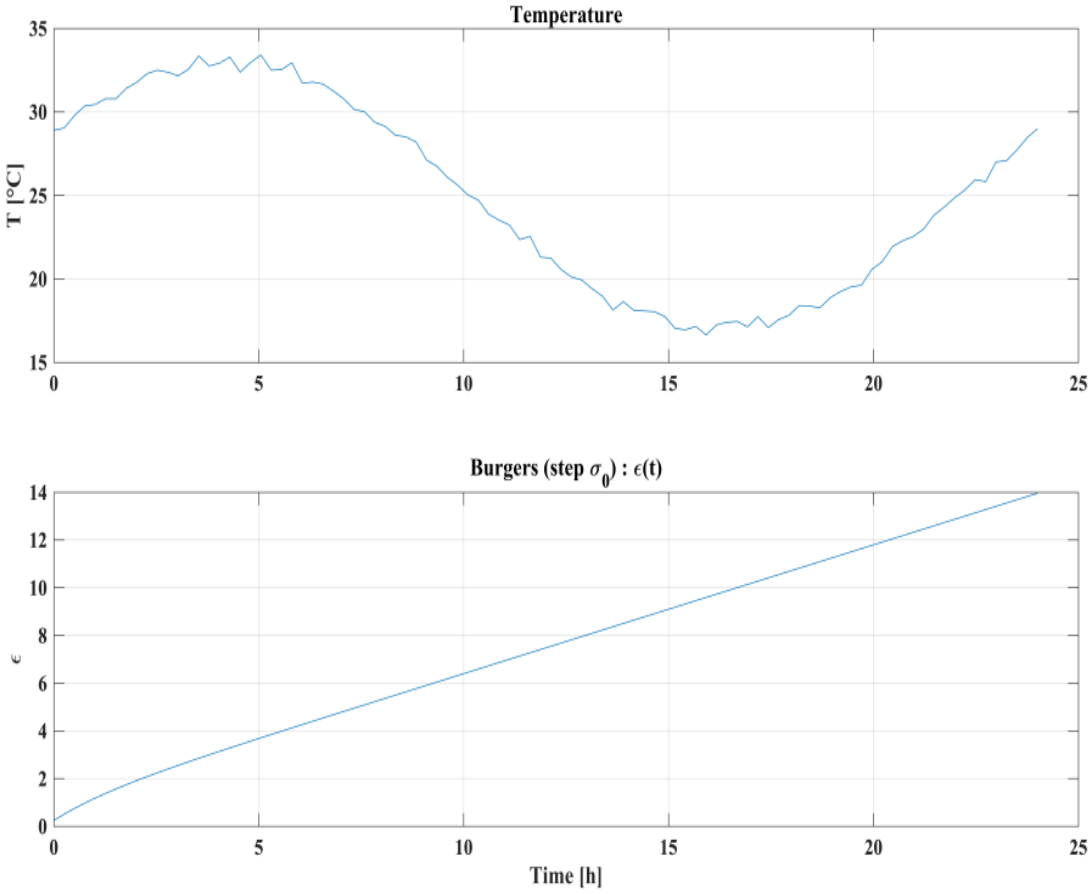

Fig. 17 Creep Burgers under Load-Step

We see a quasi-linear growth of (t): this is the viscous Maxwell term σ_0/η_1. Using Monte Carlo simulation (Simulation of sequences (σ, T, ε) by varying ($E1$, $E2$, $\eta 1$, $\eta 2$, α) and $T(t)$; output: 1,000 simulations × 96 steps (96,000 samples)) and FFNN modeling with a two-layer dense network, Dropout 0.2/0.2), swept L2 and early-stopping on validation, the following results are obtained:

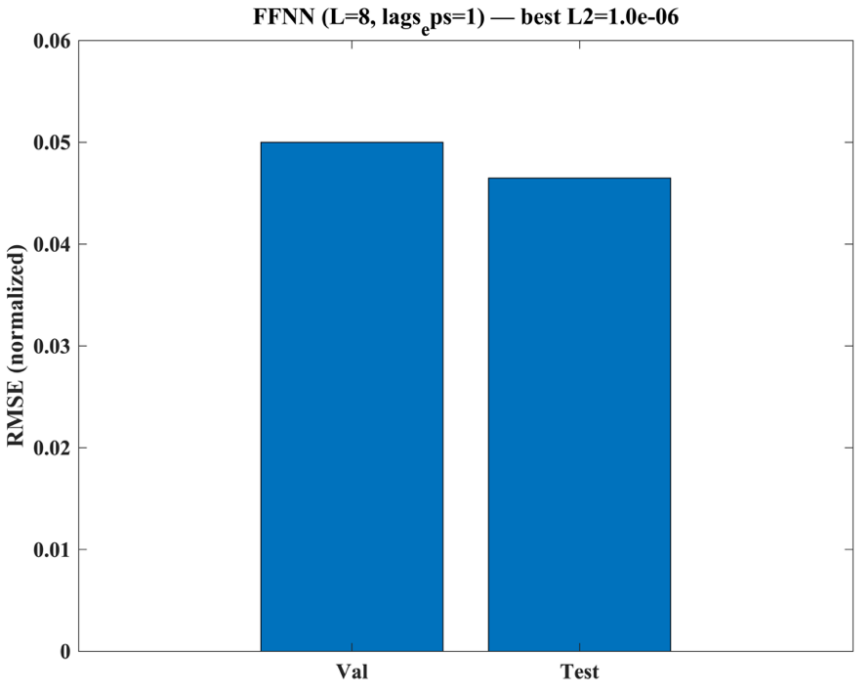

Fig. 18 RMSE FFNN option B

Compared to the "point-by-point" FFNN used before Option B (~0.349 normalized), the window + lags ε divides the error by ~7. We can say that the generalization is good (test < val). Next, we switch to a sequence-to-sequence LSTM that augments the physical inputs $[\sigma, T, \log E1, \log E2, \log\eta1, \log\eta2$ with a teacher-forced $\varepsilon(t-1)$ channel to carry the material memory; we normalize on the train split only, use early-stopping, and produce a bar chart of RMSE (Val/Test):

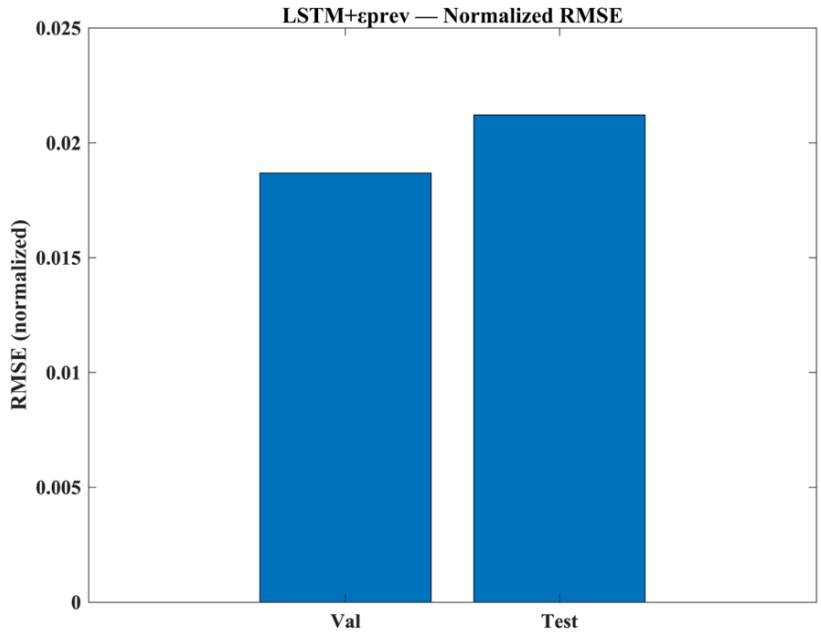

Fig. 19 RMSE LSTM option B

On our dataset, the LSTM + ε(t-1) achieves a normalized RMSE of 0.0187 (val) and 0.0212 (test)≈, 2.2× better than the windowed FFNN (0.0465). In original units: 0.091 (val) and 0.104 (test). Test < val indicates good generalization. The predicted curves follow creep/relaxation well; the remaining deviations appear mainly at the very beginning of the transients.

After comparing the results of the first and second equations, with FFNN and LSTM modeling without and with regularization, we can say that switching to Option B (Burgers + $\alpha\Delta T$) correctly captures the transient viscoelasticity (Kelvin part) and the thermal expansion, which makes the dynamics $\varepsilon(t)$ more predictable, hence the strong drop in RMSE. To conclude, we apply paired tests on the same test points. A paired t-test is applied to compare the means of two populations when the data consist of two samples in which each observation in one sample is matched with a corresponding observation in the other [39].

By adopting the following parameters:

- K=5 (number of folds (partitions))
- R = 1 (number of repetitions (repeated CV))
- $\alpha = 0.05$

After performing the test on MATLAB, we obtain:

- t = 0.047, p = 0.9645
- Average of differences (Equation (1) – Equation (2)) = 2105
- MATLAB Conclusion: No significant difference.

And following the Interpretation Standards:

- $p < \alpha$: Significant difference → one model is better than the other.
- $p > \alpha$: No significant difference detected.

On this note:

- There is no statistical evidence that one of the two models is better than the other.
- Their performances are considered equivalent from a testing perspective.

Even if we adopt a number of repetitions R=10 to compare the two models, we obtain a p=0.8438>>0.05. Hence, the test confirms that there is no statistical proof that one of the models is better and that the performances are statistically equivalent.

Based on these results, we propose comparing the reliability of the two models using the following two tests:

- Correlation test σ / error: where the predicted uncertainty is calculated to show whether the model is aware of its own uncertainties or not.
- Calibration test: where the confidence intervals (e.g., 95%) contain the truth in ~95% of cases, thus testing the reliability of the models.

The results obtained are summarized in the following table:

Table 6. Results of the correlation and calibration tests

Model	Test	Results	Interpretation
Equation (1) (FFNN [10,10] + Dropout)	Correlation test σ / error	Pearson = 0.565 Spearman = 0.941	Uncertainties are very well correlated with errors, which is positive
	Calibration test	coverage95 = 0.4 %	The intervals are too tight; the model greatly underestimates its uncertainty.
Equation (2) (LSTM + ε(t-1))	Correlation test σ / error	Pearson = 0.163 Spearman = 0.710	The model "feels" its errors a little but less well than the FFNN.
	Calibration test	coverage95 = 0 %	No real point is in [μ ± 1.96σ].

From this, we can say that both models generate relatively confident predictions of themselves, but they are poorly calibrated statistically. These results can be explained by the fact that the MC dropout with T=30 does not generate enough dispersion. By adopting an automatic calibration factor k for calibration, we obtain the following results:

Table 7. Results of the correlation and calibration tests after calibration factor

Model	Test	Results	Interpretation
Equation (1) (FFNN [10,10] + Dropout)	Correlation test σ / error	Pearson corr_P = 0.539 Spearman corr_S = 0.898	Uncertainty (σ) correlates well with actual errors, hence a good ability to flag difficult cases.
	Calibration test	The calibration factor k=9.8	Uncertainty is underestimated at the start (interval too narrow).
		coverage95 = 1,2 %	It would be necessary to calibrate even better.
Equation (2) (LSTM + ε(t-1))	Correlation test σ / error	Pearson corr_P = 0.170 Spearman corr_S = 0.179	The uncertainty given by the model has no link with its errors; the model is unreliable.
	Calibration test	The calibration factor k=1.34	No real value fell within the interval, hence poor structural calibration
		coverage95 = 0 %	It is not usable as is

We can conclude that the FFNN with Dropout is much more reliable than the LSTM + ε(t−1) in terms of uncertainty, even if its raw coverage is low; its correlations are high, which is very valuable: it can order the predictions from the most reliable to the least reliable. Whereas the LSTM cannot calibrate its uncertainty and does not provide useful information.

Other researchers can be cited as examples to highlight the applicability of LSTM, Yushun et al. in their research [40], present a new large dataset for the prediction of the International Roughness Index (IRI), comprising 2243 records, ten times more than the databases used previously. From this dataset, they propose a model called LSTM-BPNN, which combines the capabilities of LSTM networks to analyze temporal data and those of BPNN networks to integrate the transverse characteristics of pavements. The two types of information are fused using an attention mechanism, allowing the relative importance of the data to be automatically adjusted. Experimental results show that this model achieves high accuracy (R^2 = 0.867) and outperforms conventional methods, confirming the relevance of this approach to improve IRI prediction.

Another example from Amir et al can be cited, advanced LSTM models were used to represent better the complex Behavior of soil moisture in different compaction layers, which significantly improves the accuracy compared to traditional methods. By combining environmental data with site-specific characteristics, the approach allows for more reliable and rapid predictions, thus facilitating more efficient compaction. The LSTM model was significantly more accurate than traditional water balance models in predicting soil moisture changes in compaction layers, enabling better management of the drying process ("dry back"). It provides reliable real-time estimates of surface and deep moisture, helping to improve pavement quality and durability.

4.2. Limitation

Although LSTM architectures are designed to mitigate vanishing gradients and capture long-term temporal dependencies, several authors have demonstrated that they remain structurally limited. LSTMs require large amounts of trainable parameters due to the use of gated cells (input, forget, output), which increases computation cost and makes them prone to overfitting when the training dataset is not sufficiently large. Recent studies also highlighted that LSTM models still tend to suffer from gradient decay when the temporal horizon becomes very long, especially when the signal has low temporal density or high noise content [15]. Moreover, LSTMs remain purely data-driven black-box models, offering no structural interpretability regarding internal physical mechanisms; this makes their internal state representations difficult to analyze, explain, or validate scientifically [41]. These limitations justify the trend toward hybrid or physics-informed variants instead of standalone LSTMs in engineering sciences.

When applied to pavements, LSTMs predict time series (e.g., surface temperatures, IRI), but struggle to decompose the effect of constitutive laws (viscoelasticity, thermal expansion) and traffic loads because they do not explicitly incorporate the physics of materials [42]. Datasets are often short, heterogeneous, and noisy (weather stations, multiple sites), which weakens generalization between regions and structures. Studies on pavement surface temperature demonstrate the value of LSTM/ConvLSTM architectures but confirm their sensitivity to input choices, measurement quality, and the temporal density of the time series [43]. In other words, good performance at specific points does not imply robust transferability without physical coupling or spatiotemporal attention mechanisms.

In Morocco, these limitations are exacerbated because the equivalent temperature θeq governing the design varies according to the climate and increases with warming, while traffic classes (TPL) and multilayer construction induce highly non-stationary thermo-mechanical behaviors from one section to another. Purely data-driven LSTMs do not explicitly capture the temperature-dependent modulus E and viscosity, nor the spatial heterogeneity of the structures; however, the Moroccan Structures Catalogue and national standards base design and management choices on θeq and on non-trivial mechanical assumptions. In the absence of dense and continuous data series, an LSTM trained on a specific area does not generalize well to other Moroccan climatologies or other types of structures. Hence, the practical interest, for Moroccan networks, of hybrid (physics + AI) or spatio-temporal (ConvLSTM/attention) models constrained by θeq and the catalogue assumptions to guarantee mechanically consistent and transferable predictions [44].

4.3. Contribution

This research provides several contributions that are particularly relevant to the Moroccan context of pavement engineering and road asset management.

First, it introduces a thermo-mechanical modeling framework specifically adapted to the local conditions of Moroccan road networks. By integrating both daily temperature fluctuations and realistic traffic loading scenarios, the study highlights the combined influence of thermal and mechanical factors on pavement deformation, an aspect often overlooked in conventional models.

Second, the work demonstrates the potential of artificial intelligence in Moroccan pavement research. By coupling Monte Carlo simulations with FFNNs and LSTMs, the study shows that deep learning approaches can significantly reduce computational costs while maintaining predictive reliability. This marks one of the first applications of such hybrid methodologies in the field of Moroccan pavement analysis.

Third, the research contributes methodologically by proposing a reproducible framework that bridges classical mechanical formulations (e.g., Hooke, Maxwell, and Boussinesq models) with modern machine learning

techniques. This hybrid approach not only enhances predictive capabilities but also provides a transferable methodology that can be extended to related problems such as fatigue, cracking, and performance index prediction.

Finally, the study offers practical value for Moroccan road authorities and decision-makers. The developed models serve as decision-support tools capable of forecasting pavement deformation and optimizing preventive maintenance strategies. Such tools have the potential to reduce maintenance costs and extend pavement service life, thereby supporting more sustainable infrastructure management in Morocco.

4.4. Comparison with Existing Research Findings

Table 8. Comparison with existing research findings

Appearance	**Neural Network Approach for Fatigue Crack Prediction in Asphalt Pavements Using Falling Weight Deflectometer Data [45]**	**Thermo-Mechanical Pavement Deformation Prediction Using Monte Carlo Simulation and Regularized Neural Networks: A Comparative Study of FFNN and LSTM Models**
Context	Analysis of asphalt pavements in the United States, using LTPP data (FWD, ESAL, climate)	Moroccan road network, local conditions (frequent overloads, strong thermal variations)
Application	Prediction of the occurrence and progression of fatigue cracks	Prediction of thermo-mechanical deformation of flexible pavements
Model type	Classical Artificial Neural Network (ANN)	FFNN with/without dropout and LSTM with sequential memory
Data used	Actual measurements: FWD, traffic, annual average temperature	Data generated by Monte Carlo simulations (variable loads, T°, elastic modulus, viscosity)
Results	Good predictive performance with $R^2 \approx 0.9$	FFNN with dropout: better generalization; LSTM: adapted to temporal dependence (Equation (2))
Contribution	Demonstration of the usefulness of ANNs for anticipating cracking in pavements	First integration in Morocco of a hybrid approach (mechanics + deep learning + Monte Carlo), with potential for preventive maintenance and cost reduction

4.5. Perspectives

In order to ensure the continuity of research, some perspectives can be proposed:

- Test other neural network architectures (CNN, RNN, Transformers) to compare their performance with that of the FFNN and LSTM.
- Increase the size and diversity of the dataset by integrating real-world field data in addition to Monte Carlo simulations.
- Experiment with advanced regularization techniques (dropout, L2, batch normalization) to improve model robustness.
- Use transfer learning to adapt models trained on one type of road surface to other contexts or climatic conditions.
- Integrate several additional variables (e.g., humidity, rainfall) to enrich model inputs.
- Implement a comparative approach between classic deterministic models (Maxwell, Burgers) and hybrid deep learning models to evaluate the contributions of AI better.

5. Conclusion

In this work, the problem of pavement deformation in Moroccan road networks was addressed by considering the combined effects of mechanical loading and thermal variations. The study was motivated by the limitations of traditional models in capturing complex behaviors and uncertainties, and sought to explore the potential of modern deep learning approaches.

Two formulations of the deformation problem were proposed. The first directly related deformation to load and temperature, while the second introduced a memory effect through the $\varepsilon(t-1)$ term. Monte Carlo simulations provided the dataset, and several neural architectures were tested, including feed-forward networks of different sizes, with and without dropout regularization, as well as recurrent LSTM networks designed for sequential modeling.

The experiments showed that dropout-regularized FFNNs offered consistent accuracy and reduced overfitting, whereas LSTMs only brought limited improvements, even when applied to the memory-based formulation. Statistical comparisons confirmed that the two best models achieved similar predictive errors. However, the FFNN with dropout stood out by producing informative uncertainty estimates, in contrast with the LSTM, whose uncertainty intervals lacked reliability.

Taken together, these findings indicate that regularized feed-forward networks remain a robust and interpretable option for pavement deformation modeling. Beyond accuracy, their ability to provide meaningful uncertainty measures makes them particularly promising for predictive maintenance applications. Future work should focus on refining uncertainty calibration and extending the framework to broader performance indicators of road infrastructure.

References

[1] T.A. Cruse, *Reliability-Based Mechanical Design*, CRC Press, pp. 1-341, 1997. [Google Scholar] [Publisher Link]

[2] Sameh S. Abd-elfattah, Ahmed I. Abu-Elmaat, and Ibrahim H. Hashim, "Reliability Analysis of Flexible Pavement Using Crude Monte Carlo Simulation," *ERJ. Engineering Research Journal*, vol. 45, no. 3, pp. 447-456, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Nawras Shatnawi, Mohammed Taleb Obaidat, and Amjad Al-Sharideah, "Modeling Road Pavement Rutting Using Artificial Neural Network and Conventional Measurements," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2677, no. 2, pp. 973-984, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Chunru Cheng et al., "Predicting Rutting Development of Pavement with Flexible Overlay Using Artificial Neural Network," *Applied Science*, vol. 13, no. 12, pp. 1-15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Claude Nadeau, and Yoshua Bengio, "Inference for the Generalization Error," *Machine Learning*, vol. 52, pp. 239-281, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[6] Mingyang Gong, Yiren Sun, and Jingyun Chen, "Analysis of Coupled Thermo-Mechanical Response and Damage Behaviour of Curved Ramp Bridge Deck Pavement using a 3D Multiscale Method," *Road Materials and Pavement Design*, vol. 25, no. 11, pp. 2335-2357, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Xu Li et al., "Analysis of Asphalt Pavement Response to Long Longitudinal Slope Considering the Influence of Temperature Fields," *Materials*, vol. 18, no. 15, pp. 1-18, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[8] Rouba Joumblat et al., "State-of-the-Art Review on Permanent Deformation Characterization of Asphalt Concrete Pavements," *Sustainability*, vol. 15, no. 2, pp. 1-34, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[9] Maureen Caudill, "Neural Networks Primer, Part I," *AI Expert*, vol. 2, no. 12, pp. 46-52, 1987. [Google Scholar] [Publisher Link]

[10] Jidong Yang, "*Road Crack Condition Performance Modeling using Recurrent Markov Chains and Artificial Neural Networks*," Digital Commans, University of South Florida, pp. 1-111, 2024. [Google Scholar] [Publisher Link]

[11] Yingjie Tian, and Yuqi Zhang, "A Comprehensive Survey on Regularization Strategies in Machine Learning," *Informatioin Fusion*, vol. 80, pp. 146-166, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep Learning*, MIT Press, pp. 1-775, 2016. [Google Scholar] [Publisher Link]

[13] Nitish Srivastava, "*Improving Neural Networks with Dropout*," Master's Thesis, University of Toronto, pp. 1-26, 2013. [Google Scholar] [Publisher Link]

[14] Sepp Hochreiter, and Jürgen Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997. [CrossRef] [Google Scholar] [Publisher Link]

[15] Klaus Greff et al., "LSTM: A Search Space Odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222-2232, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[16] Tianjie Zhang et al., "LSTM+MA: A Time-Series Model for Predicting Pavement IRI," *Infrastructures*, vol. 10, no. 1, pp. 1-16, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[17] Micah Mers et al., "Recurrent Neural Networks for Pavement Performance Forecasting: Review and Model Performance Comparison," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2677, no 1, pp. 610-624, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Samik Raychaudhuri, "Introduction to Monte Carlo Simulation," *2008 Winter Simulation Conference*, Miami, FL, USA, pp. 91-100, 2008. [CrossRef] [Google Scholar] [Publisher Link]

[19] Deep Learning: Theory, Algorithms, and Implications, dl2025. [Online]. Available: https://dl2025.fbk.eu/

[20] Andrew. Y. Ng, "Feature Selection, L1 vs. L2 Regularization, and Rotational Invariance," *Twenty-First International Conference on Machine Learning - ICML '04*, Banff, Alberta, Canada, pp. 1-78, 2004. [CrossRef] [Google Scholar] [Publisher Link]

[21] Nithish Srivastava et al., "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929-1958, 2014. [Google Scholar] [Publisher Link]

[22] Yarin Gal, and Zoubin Ghahramani, "Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning," *Proceedings of the 33rd International Conference on International Conference on Machine Learning*, vol. 48, pp. 1050-1059, 2016. [Google Scholar] [Publisher Link]

[23] Lutz Prechelt, "Automatic Early Stopping Using Cross Validation: Quantifying the Criteria," *Neural Networks*, vol. 11, no. 4, pp. 761-767, 1998. [CrossRef] [Google Scholar] [Publisher Link]

[24] Connor Shorten, and Taghi M. Khoshgoftaar, "A Survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, pp. 1-48, 2019. [CrossRef] [Google Scholar] [Publisher Link]
[25] Sergey Ioffe, and Christian Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *Arxiv Preprint*, pp. 1-11, 2015. [CrossRef] [Google Scholar] [Publisher Link]
[26] Rafael Müller, Simon Kornblith, and Geoffrey E. Hinton, "When Does Label Smoothing Help?," *Advances in Neural Information Processing Systems*, pp. 1-10, 2019. [Google Scholar] [Publisher Link]
[27] Chris M. Bishop, "Training with Noise is Equivalent to Tikhonov Regularization," *Neural Compution*, vol. 7, no. 1, pp. 108-116, 1995. [CrossRef] [Google Scholar] [Publisher Link]
[28] Hongyi Zhang et al., "Mixup: Beyond Empirical Risk Minimization," *Arxiv Preprint*, pp. 1-13, 2018. [CrossRef] [Google Scholar] [Publisher Link]
[29] Song Han, Huizi Mao, and William J. Dally, "Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding," *Arxiv Preprint*, pp. 1-14, 2016. [CrossRef] [Google Scholar] [Publisher Link]
[30] Pamela Carreno, Dana Kulic, and Michael Burke, "Adapting Neural Models with Sequential Monte Carlo Dropout," *Proceedings of The 6th Conference on Robot Learning, PMLR*, pp. 1542-1552, 2023. [Google Scholar] [Publisher Link]
[31] Stephan Thaler et al., "Active Learning Graph Neural Networks for Partial Charge Prediction of Metal-Organic Frameworks via Dropout Monte Carlo," *NPJ Computational Materials*, vol. 10, pp. 1-10, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[32] Hongchen Liu et al., "Multi-Fidelity Deep Neural Network with Monte Carlo Dropout Technique for Uncertainty-Aware Risk Recognition of Backward Erosion Piping in Dikes," *Applied Soft Computing*, vol. 166, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[33] Alex Kendall, and Yarin Gal, "What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?," *Advances in Neural Information Processing Systems*, pp. 1-11, 2017. [Google Scholar] [Publisher Link]
[34] Yarin Gal, "*Uncertainty in Deep Learning*," University of Cambridge, pp. 1-174, 2016. [Google Scholar] [Publisher Link]
[35] Ian Osband et al., "Deep Exploration via Bootstrapped DQN," *Advances in Neural Information Processing Systems*, pp. 1-9, 2016. [Google Scholar] [Publisher Link]
[36] Yaniv Ovadia et al., "Can you Trust Your Model's Uncertainty? Evaluating Predictive Uncertainty Under Dataset Shift," *Advances in Neural Information Processing Systems*, pp. 1-12, 2019. [Google Scholar] [Publisher Link]
[37] Meire Fortunato, Charles Blundell, and Oriol Vinyals, "Bayesian Recurrent Neural Networks," *Arxiv Preprint*, pp. 1-14, 2019. [CrossRef] [Google Scholar] [Publisher Link]
[38] Menglan Zeng, and Donald H. Shields, "Nonlinear Thermal Expansion and Contraction of Asphalt Concrete," *Canadian Journal of Civil Engineering*, vol. 26, no. 1, pp. 26-34, 1999. [CrossRef] [Google Scholar] [Publisher Link]
[39] Rosie Shier, Statistics: 1.1 Paired T-Tests, mathematics Learing Support Centre, 2004. [Online]. Available: https://www.statstutor.ac.uk/resources/uploaded/paired-t-test.pdf
[40] Yushun Dong et al., "Forecasting Pavement Performance with a Feature Fusion LSTM-BPNN Model," *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, Beijing, China, pp. 1953-1962, 2019. [CrossRef] [Google Scholar] [Publisher Link]
[41] Nima Amini, and Qinqin Zhu, "Fault Detection and Diagnosis with a Novel Source-Aware Autoencoder and Deep Residual Neural Network," *Neurocomputing*, vol. 488, pp. 618-633, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[42] Sepideh Emami Tabrizi et al., "Hourly Road Pavement Surface Temperature Forecasting Using Deep Learning Models," *Journal of Hydrology*, vol. 603, 2021. [CrossRef] [Google Scholar] [Publisher Link]
[43] Meng Zhang et al., "A Deep Learning Approach for Enhanced Real-Time Prediction of Winter Road Surface Temperatures in High-Altitude Mountain Areas," *Promet - Traffic Transportation*, vol. 36, no. 5, pp. 958-972, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[44] Mouncif Sarroukh et al., "Effect of Global Warming and New Equivalent Temperature Zoning Maps for Asphalt Pavement Design in Morocco," *Energy and Buildings*, vol. 303, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[45] Bishal Karki et al., "Neural Network Approach for Fatigue Crack Prediction in Asphalt Pavements Using Falling Weight Deflectometer Data," *Applied Science*, vol. 15, no. 7, pp. 1-17, 2025. [CrossRef] [Google Scholar] [Publisher Link]