



HAL
open science

Assessment of Deep Learning models for syndromic surveillance

Lavoine Joseph, Jean-Luc Vinard, Jean-Philippe Amat, Viviane Henaux, Eric Morignat

► **To cite this version:**

Lavoine Joseph, Jean-Luc Vinard, Jean-Philippe Amat, Viviane Henaux, Eric Morignat. Assessment of Deep Learning models for syndromic surveillance. Society for Veterinary Epidemiology and Preventive Medicine (SVEPM), Mar 2023, Toulouse, France. anses-04517355

HAL Id: anses-04517355

<https://anses.hal.science/anses-04517355>

Submitted on 22 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

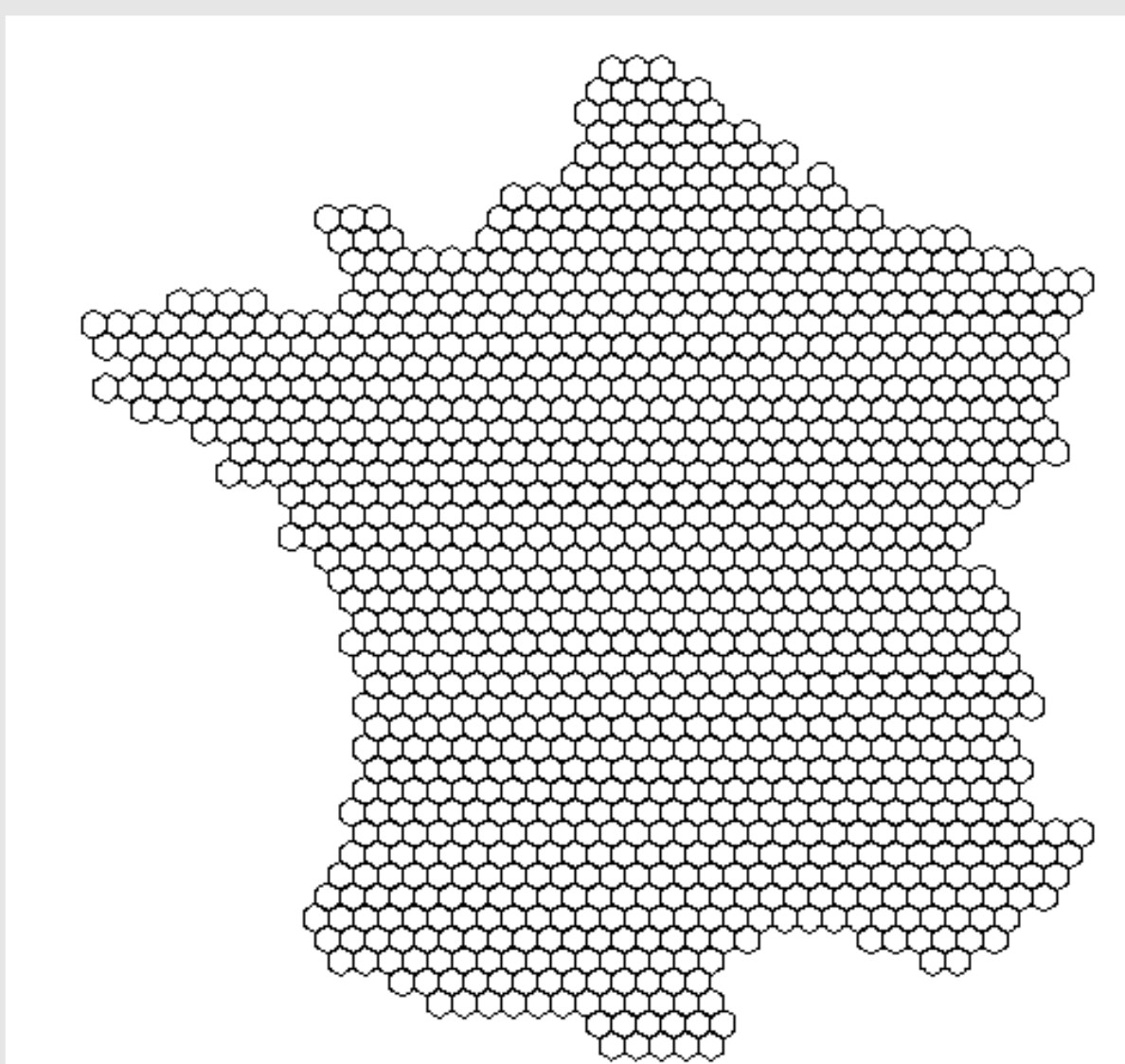
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CONTEXT

Syndromic surveillance is based on the **monitoring of non-specific indicators** to enable the **early identification of the impact of potential human or animal health threats** that require public and/or animal health action. Temporal fluctuations of these indicators are usually modeled with regression models on reference periods, i.e. free of health events. The monitoring of these indicators allow to draw statistical alarms when they significantly differ from model predictions.

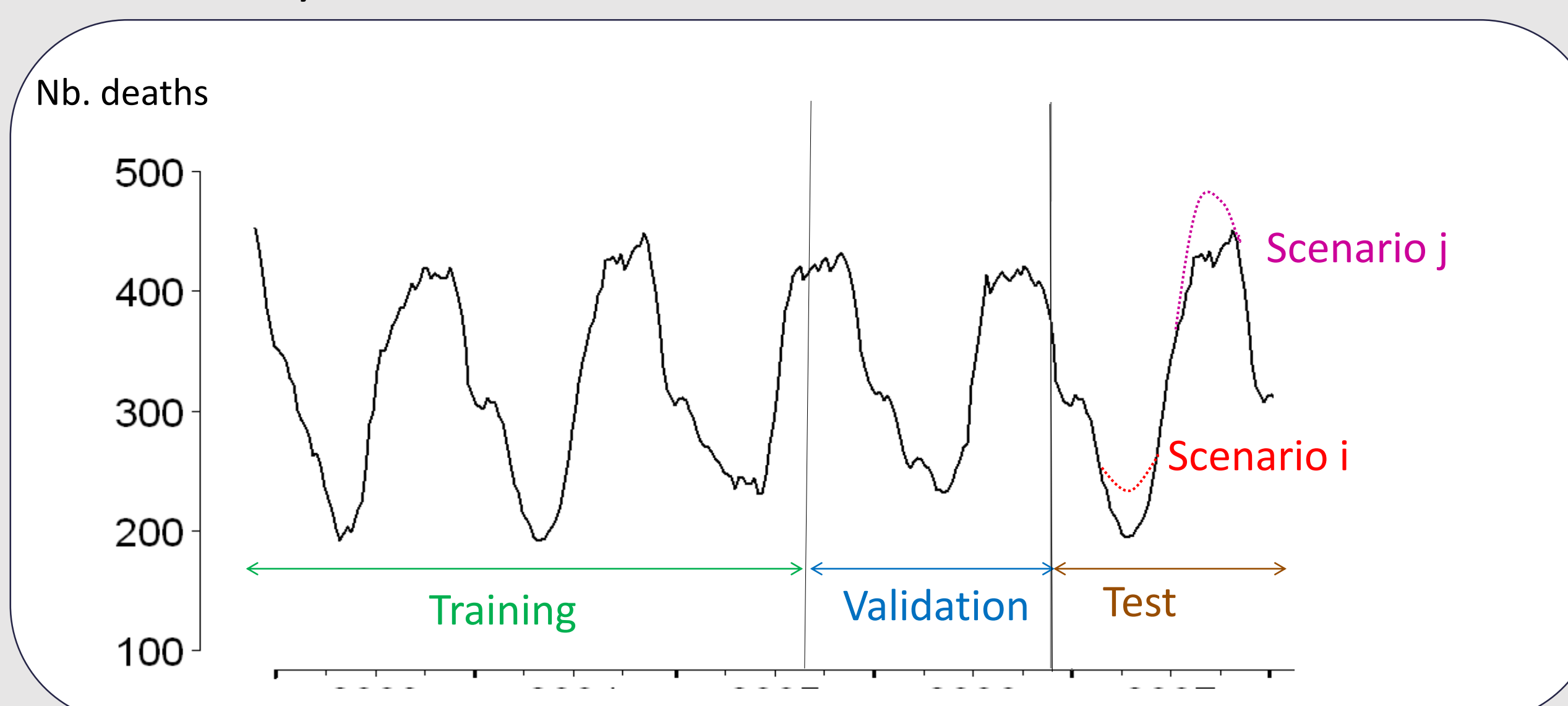
Recently, **machine learning** and **deep learning** have shown promising results in epidemiology¹. In order to investigate their usefulness in the field of syndromic surveillance, **we assessed the performances of different structures of neural networks in detecting anomalies in time series**, using cattle mortality data in France as a study case. Their performances were compared with results from a previous study², combining a regression model with a cluster detection method (= reference model).

MATERIAL AND METHOD



France was divided in 1,040 hexagons of 5-km side. In each hexagon, a 9-year historical time series of weekly dead cattle data was computed.

Time series were divided into 3 sets of data: 1) **Training** and 2) **validation** datasets -> used to calibrate the neural networks models. 3) **Test** dataset: -> in which simulated extra deaths mimicking the spread of an infectious fatal disease were injected (1000 scenarios).



7 models evaluated: Deep learning models including recurrent neural networks (e.g Transformers, LSTM), convolutional neural networks, dense neural networks, and combinations of these structures; and a basic neural network (one neuron without activation function) similar to a linear regression model.

- **In each hexagon**, models were used to compute the weekly expected numbers of dead cattle. Based on differences between observed and expected dead cattle numbers, two alternative approaches were used to flag weeks with anomalies for each hexagon : 1/ a time varying cluster approach (DBSCAN algorithm) and 2/ a common threshold to all hexagons.
- **Several performance criteria** in anomaly detection used to compare results from the different models: proportion of success, timeliness, sensitivity, positive predictive value (PPV) and negative predictive value (NPV).

RESULTS

Table 1. Mean performances of the two best deep learning models (Transformers and LSTM-dense) and the basic neural network (LRM) using the threshold method for anomaly detection, and the reference model²

	Transformers	LSTM-dense	LRM	Reference
Success (%) *	65.0	67.1	63.0	73.2
Timeliness (week)**	2.0	1.8	2.4	1.4
Sensitivity (%)	10.9	13.3	12.3	13.0
PPV (%)	43.9	40.0	40.9	12.7
NPV (%)	99.9	99.0	99.0	99.9

* success is defined as the detection of at least one week with simulated death

** Timeliness is defined as the shortest delay of detection among all hexagons

- Better performances of the models with the threshold method (Table 1) than with the DBSCAN algorithm (not shown) for anomaly detection
- Higher PPV with deep learning models (Transformers and LSTM-dense) than reference model but similar sensitivity and lower success
- Surprisingly, LRM has similar performances than deep learning models despite extremely different model sizes (a few tens versus several million parameters)

CONCLUSION AND PERSPECTIVES

- Deep learning models did not show better performances than the basic neural network in spite of the difference in model complexity
- Further steps to improve performances:
 - Investigate other structures of neural networks
 - Increase input information by extending the time series length and integrating new input variables, such as temperature/humidity³ and calving⁴ whose influence on cattle mortality has already been demonstrated

REFERENCES

- ¹ Wiemken, T. L. *et al.* 2020. Annual Review of Public Health – doi: <https://doi.org/10.1146/annurev-publhealth-040119-094437>
- ² Perrin, J. B. *et al.* 2015. PLoS One - doi: <https://doi.org/10.1371/journal.pone.0141273>
- ³ Morignat, E. *et al.* 2018. Journal of Thermal Biology - doi: <https://doi.org/10.1016/j.jtherbio.2018.11.001>
- ⁴ Perrin, J.-B. *et al.* 2010. Epidemics - doi: <https://doi.org/10.1016/j.epidem.2010.10.002>