AI-Based Pathogenicity Classification of Flea and Tick-Borne Diseases and Seasonal Distribution of Some Species in Algeria

Noureddine Rabah Sidhoum, Mehdi Boucheikhchoukh and
Noureddine Mechouk

December 19, 2023

AI-based pathogenicity classification of flea and tick-borne diseases and seasonal distribution of some species in Algeria

Noureddine Rabah sidhoum[1], Mehdi Boucheikhchoukh[1], Noureddine Mechouk[2,3]

1. Biodiversity and Ecosystems Pollution Laboratory, Faculty of Life and Nature Sciences, Chadli Bendjedid University, El Tarf 36000, Algeria.
2. Ecology of Terrestrial and Aquatics Systems Laboratory (EcoSTAq), Department of Biology, Faculty of Science, Badji Mokhtar University, Annaba 23200, Algeria.
3. Department of Parasitology and Parasitic Diseases, Faculty of Veterinary Medicine, University of Agricultural Sciences and Veterinary Medicine of Cluj-Napoca, Calea Mănăștur 3-5, Cluj-Napoca 400372, Romania.

***Abstract*—Vector-borne diseases and their classification are paramount epidemiological importance since they are an important cause of death or debilitating situations, especially in African countries. In this study, a database of vector-borne diseases was generated out of Algerian data to train two different classifiers and compare the fitness of each one. In addition, 1037 fleas and ticks were sampled in different regions of the central north of the country to understand the seasonal activity of these vectors and their implication in transmitting diseases to humans and animals.**

***Keywords*—ANN; epidemiology; Arthropods; north-Africa; zoonoses**

I. INTRODUCTION

In North Africa, the coexistence of various arthropod-borne pathogens poses a substantial public health challenge [1], [2]. Ticks and fleas transmit these pathogens, encompassing a spectrum of diseases with diverse clinical manifestations. Among these are TG group *Rickettsia felis* and *R. typhi*, bartonelloses, *Dipylidium caninum*, and plague, predominantly transmitted by fleas [3], [4]. In contrast, ticks are vectors cable of transmitting SFG group rickettsials, certain bartonelloses, and filarial helminths [5], [6]. Understanding the dynamics of these pathogens is critical for developing effective preventive measures and interventions.

The prevalence and distribution of these arthropod-borne pathogens vary across the region, and several factors influence their transmission [7]. These factors encompass a complex interplay between humans, animals, and arthropod vectors. Thus, it is imperative to employ advanced tools and methodologies to categorize these pathogens, shedding light on their spatial distribution, host specificity, and the competence of arthropod vectors.

This study seeks to employ artificial neural network models to classify the pathogens mentioned above, focusing on four fundamental aspects: reproduction rate, prevalence, vector/host presence, and vectorial competence. Prevalence is a critical indicator of the risk posed by a particular pathogen in a given region. By discerning the prevalence detected in humans, animals, and arthropods, we can gain insights into the potential burden of these diseases on public health and animal populations. Vector presence, which pertains to the geographic distribution of arthropod vectors and their vectorial competence, plays a pivotal role in transmitting these pathogens. In summary, this research contributes to our knowledge of tick and flea-borne pathogens in North Africa, enhancing our capacity to mitigate the impact of these diseases on human and animal populations. Through the innovative application of artificial neural network models, we aspire to offer a more nuanced classification that integrates factors critical to disease prevalence, vector presence, and vectorial competence, ultimately advancing our understanding of the transmission of these pathogens in the region.

II. MATERIALS AND METHODS

A. Study area and ticks and fleas sampling

Between June 2022 and July 2023, a sampling was conducted in four provinces of the central north of Algeria: Algiers, Boumerdes, Blida, and Tizi Ouzou.

The sampling targeted domestic and peridomestic, farms and stray and wild animals. The localization of visited counties during the fieldwork is illustrated in Figure 1.

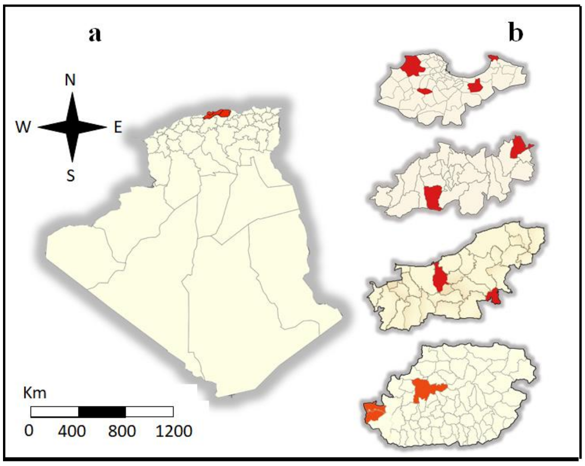

Figure 1: (a) Study area (Algiers, Blida, Boumerdes and Tizi-Ouzou) (b) Visited counties (Aïn Benian, Cheraga, Bab Ezzouar, Bab Hassen, Oued Smar, El Marsa, Bouarfa, Meftah, Zemouri, Timezirt, M'kira, Tizi Ghennif, and Tizi Ouzou).

The arthropods were conserved in 70° ethanol micro-tubes at T=-12C°.

The identification of ectoparasites was established according to taxonomic keys [8], [9] the fieldwork was completed by data analysis using artificial neural networks based on published data from the country.

B. *Artificial neural network classifiers*

In order to generate the dataset for artificial neural network analysis a systematic review was carried out following the guidelines of Preferred Reporting Elements for Systematic Reviews and Meta-analysis (PRISMA). The focus was on references related to flea and tick-borne pathogens in Algeria The search, conducted up to November 2023, targeted materials published online in English and French, , including peer-reviewed publications, theses, DVM and M's dissertations, local authorities reports, and published communications.

In addition to data regarding the reported prevalence in humans, animals, and arthropods, the presence or absence of the arthropod vector (fleas and ticks), the geographical location (urban, rural, forest, or desert), the competence of the arthropod in transmitting the disease and the type of animal species (domestic, peridomestic, feral, synanthropic, or wild) were noted.

Seven covariates were considered, with the pathogenicity as a dependent variable.

Each covariate was attributed a score based on an appropriate scale.

The pathogenicity score was determined using classifications from NIH, APSA, ePATHogen, and German scales, ranging from (1) for non-pathogenic microorganisms to (4) for highly pathogenic and communicable ones. Scores for pathogenic and highly pathogenic organisms were (2) and (3), respectively. A binary score (0/1) denoted the absence/presence of host generalist (or host-specific) fleas. Host behavior was assessed based on lifestyle and contact with humans, with scores ranging from (1) for wild animals to (5) for pets. Locations were individually examined using coordinates from consulted studies and reports, employing Google Earth (Version 7.3.6) and the United Nations DESA database (2018 revision) to estimate scores related to urbanization and population density. The lowest scores (1) and (2) were assigned to desertic and forest lands, while suburban and rural areas received a score of (3), and urban areas were assigned the highest score (4). Prevalences of detected pathogens in humans, animals, and fleas were extracted or calculated based on available data and incorporated into the database (Supplementary Material).

Statistical analyses, including multilayer perceptron and radial basis function neural networks, were conducted using IBM-SPSS version 26.0.0 (IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp).

Information regarding data analysis can be found in the **supplementary material 1**.

III. RESULTS

A. *Sampling outcomes*

During the sampling, 162 animals were examined. Among them, 126 (77.78%) were infested by at least one ectoparasite species. A total number of 1037 ticks and fleas was recovered from infested animals. Details about ectoparasite species with their relative count are depicted in Figure 2.

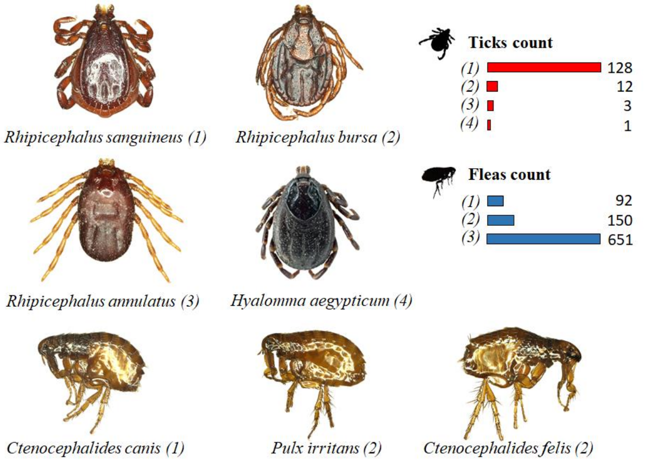

Figure 2: Identified flea and tick species with the number of collected specimens.

Figure 3 shows the bimestrial chart of the ectoparasites in terms of their numbers.

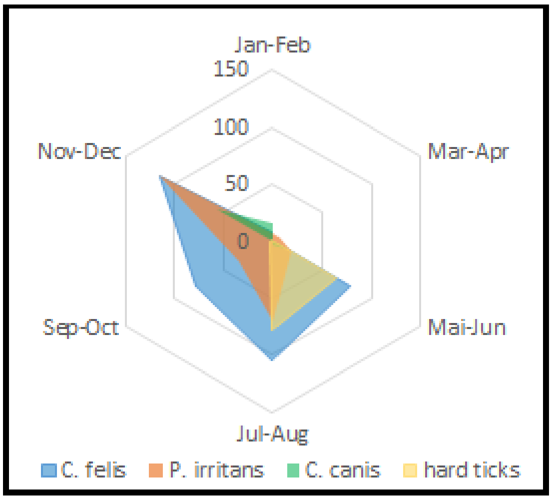

Figure 3: Radar chart of recovered fleas and ticks seasonal activity

B. *Pathogens classifying models*

Tick-borne pathogens that were previously detected in Algeria and considered in this study are the following: *Anaplasma bovis, A. centrale, A. marginale, A. ovis, A. phagocytophilum, A. platys, Babesia bigemina, B. bovis, B. ovis, B. vogeli, Borrelia bugdorferi, B. hispanica, B. theileri, B. turicantae, Bartonella bovis, B. dromedarii, B. tamiae, Coxiella brunetti, Ehrlichia canis, Hepatozoon canis, Mycoplasma wenyoni, Rickettsia aeschlimannii, R. africae, R. conorii, R. helvetica, R. massilae, R. monacensis, R. raoulti, R.*

sibirica, R. slovaca, Theileria annulata, T. equi, T. orientalis, and *T. ovis.*

Several flea-borne pathogens have also been reported in Algeria, including *Bartonella claridgeiae, B. elizabethae, B. henslae, B. tribocorum, B.v. berkhoffi, Dipylidium caninum, Rickettsia felis, R. typhi,* and *Yersinia pestis.*

The training and testing results of each models are as follows:

Table 1 represents the cross entropy error of training and testing as well as percentage of incorrect predictions.

TABLE I. MULTILAYER PERCEPTRON SUMMARY

Training	Cross Entropy Error	83,274
	Percent Incorrect Predictions	28,4%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error
	Training Time	0:00:00,03
Testing	Cross Entropy Error	28,098
	Percent Incorrect Predictions	21,3%

Figure 4 represents the ROC (Receiving Operating Characteristic) curve of the multilayer perceptron model.

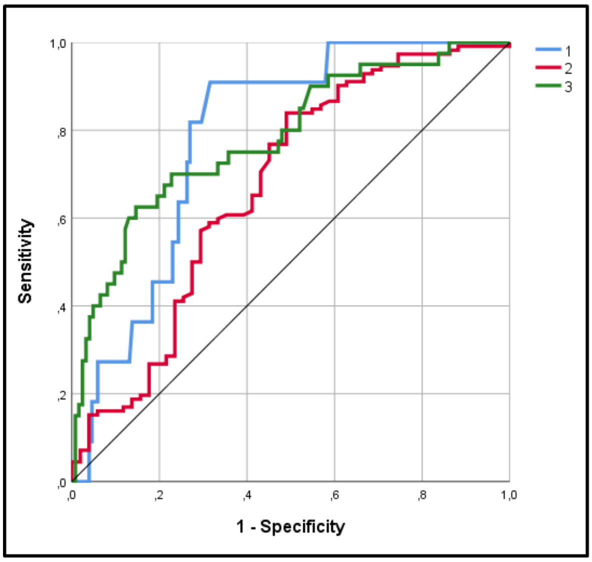

Figure 4: performance of the multilayer perceptron model in classifying the pathogenicity of tick and flea-borne pathogens in Algeria according to published data.

Table 2 displays the sum of squares error and the percentage of incorrect predictions in the training/testing phase for the radial basis function model.

TABLE II. RADIAL BASIS FUNCTION SUMMARY

Training	Sum of Squares Error	19,151
	Percent Incorrect Predictions	22,5%
	Training Time	0:00:00,18
Testing	Sum of Squares Error	11,014[a]
	Percent Incorrect Predictions	30,8%

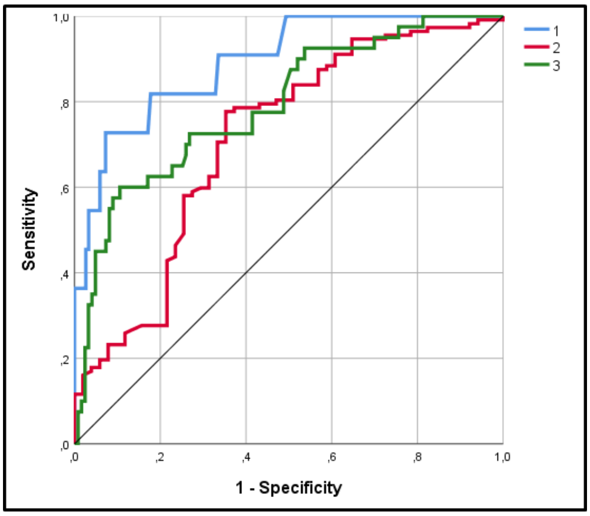

Figure 5: performance of the radial basis function model in classifying the pathogenicity of tick and flea-borne pathogens in Algeria according to published data.

Figure 5 represents the radial basis function model's ROC curve (sensitivity/specificity).

IV. DISCUSSION

Algeria has a warm Mediterranean climate [10] that generally promotes arthropods' burden. Therefore, the results of the small survey that we conducted led to seven different species of ectoparasites. Similar data have been previously reported from different parts of the country [11], [12].

For fleas, *C. felis* and *P. irritans* were the most abundant fleas. Both species had a seasonal activity ranging from May to December, while for ticks *R. sanguineus (s.l.)* was the most abundant tick in the studied hosts with a restricted activity from May to August. These results align with previous studies [13], [14].

Regarding pathogens classifying models, when comparing the summary of both models, it appears that the MLP (multilayer perceptron) model had a better performance than the RBF (radial-based function) model, with reduced training time and error in its prediction, even though it’s known that MLPs are prone to overfitting for small-sized datasets, as the one exploited in this study [15]. However, this case does not prove that MLP is the ideal classifier for this task.

The dataset, consisting of only 163 samples, cannot adequately represent a sufficient number of pathogenicity classes. This necessitates the use of a larger dataset [15]. Both classifiers yielded similar results in classifying pathogens from the second risk group (the second level of pathogenicity), constituting most of the dataset. Furthermore, both models exhibited a similar trend in ROC curve analysis for classifying non-pathogenic microorganisms. However, this cannot be a deciding criterion in choosing one over the other, as only eleven samples were categorized as non-pathogenic. This small sample size makes it insufficient for both models to be thoroughly trained and tested, as described by [16].

Nevertheless, reconsidering the covariates may be appropriate to efficiently determine the pathogenicity of pathogens since some of the covariates, such as vectorial competence and the presence of ticks or fleas, are more epidemiologic than clinical criteria.

Finally, this study brought some insights into some ectoparasites distribution and seasonal activity in the central north of Algeria with a trial on classifying the circulating pathogens in this region, which needs to be completed by further studies to have full insight into flea and tick-borne diseases and their reservoirs in the country.

ACKNOWLEDGMENT

Dr. Farid Tritri, a veterinary practitioner at M’kira Veterinary Clinic, realized a part of the sampling (Tizi Ouzou).

REFERENCES

[1]. El Hamzaoui, B., Zurita, A., Cutillas, C., & Parola, P. (2020). Fleas and flea-borne diseases of North Africa. Acta tropica, 211, 105627.

[2]. Perveen, N., Muzaffar, S. B., & Al-Deeb, M. A. (2021). Ticks and tick-borne diseases of livestock in the Middle East and North Africa: A review. Insects, 12(1), 83.

[3]. Bitam, I., Parola, P., De La Cruz, K. D., Matsumoto, K., Baziz, B., Rolain, J. M., ... & Raoult, D. (2006). First molecular detection of Rickettsia felis in fleas from Algeria. The American journal of tropical medicine and hygiene, 74(4), 532-535.

[4]. Bitam, I., Baziz, B., Kernif, T., Harrat, Z., Parola, P., & Raoult, D. (2009). Molecular detection of Rickettsia typhi and Rickettsia felis in fleas from Algeria. Clinical microbiology and infection, 15, 255-256.

[5]. Kernif, T., Djerbouh, A., Mediannikov, O., Ayach, B., Rolain, J. M., Raoult, D., & Bitam, I. (2012). Rickettsia africae in Hyalomma dromedarii ticks from sub-Saharan Algeria. Ticks and tick-borne diseases, 3(5-6), 377-379.

[6]. Benredjem, W., Leulmi, H., Bitam, I., Raoult, D., & Parola, P. (2014). Borrelia garinii and Rickettsia monacensis in Ixodes ricinus ticks, Algeria. Emerging infectious diseases, 20(10), 1776.

[7]. Pherez, F. M. (2007). Factors affecting the emergence and prevalence of vector-borne infections (VBI) and the role of vertical transmission (VT). Journal of vector-borne diseases, 44(3), 157-164.

[8]. Beaucournu, J. C., & Launay, H. (1990). Les puces (siphonaptera) de France et du Bassin méditerranéen occidental: par Jean-Claude Beaucournu et Henri Launay. Lechevalier.

[9]. Estrada-Peña, A., Mihalca, A. D., & Petney, T. N. (Eds.). (2018). *Ticks of Europe and North Africa: a guide to species identification*. Springer.

[10]. Ammam, I., Rahal, M., Rahal, K., & Bitam, I. (2022). Epidemiological study and identification of some flea species infesting wild rabbits (Oryctolagus cuniculus) and cape hares (Lepus capensis) in northern Algeria. Iraqi Journal of Veterinary Sciences, 36(1), 159-165.

[11]. Amrani, S., Seaward, M. R., Sipman, H. J., & Feuerer, T. (2018). Lichenological exploration of Algeria II: checklist of lichenized, lichenicolous, and allied fungi. Herzogia, 31(2), 817-892.

[12]. Sidhoum, N. R., Boucheikhchoukh, M., Mechouk, N., & Deak, G. (2023). An overview of fleas (Siphonaptera) in wild and domestic mammals from Algeria with new data from the central north and south of the country. Acta Tropica, 107004.

[13]. Hassissen, L., Kebbi, R., Ayad, A., & Nait-Mouloud, M. (2019). Seasonal activity of ticks infesting domestic dogs in Bejaia province, Northern Algeria. Onderstepoort Journal of Veterinary Research, 86(1), 1-6.

[14]. Li, Y., Pont, M. J., Jones, N. B., & Twiddle, J. A. (2001). Applying MLP and RBF classifiers in embedded condition monitoring and fault diagnosis systems. Transactions of the Institute of Measurement and Control, 23(5), 315-343.

[15]. Li, Y., Pont, M. J., Jones, N. B., & Twiddle, J. A. (2001). Applying MLP and RBF classifiers in embedded condition monitoring and fault diagnosis systems. Transactions of the Institute of Measurement and Control, 23(5), 315-343.

[16]. Sug, H. (2010). The effect of training set size on the performance of neural networks for classification. WSEAS Trans Comput, 9, 1297-1306.