Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Describing 2008 PPR epizootic in Morocco through the lens of mathematical modeling: Uncovering its multiscale dynamics

Manal Nhili

manal-nhili@um5r.ac.ma

ENSIAS, Mohammed V University in Rabat

Stéphanie Bougeard

Ploufragan-Plouzané-Niort Laboratory, the French Agency for Food, Environmental and Occupational Health & Safety (ANSES)

Mounir Khayli

Faculty of Sciences, Mohammed V University in Rabat

Noha El Khattabi

Faculty of Sciences, Mohammed V University in Rabat

Mounia Abik

ENSIAS, Mohammed V University in Rabat

Andrea Apolloni

Ploufragan-Plouzané-Niort Laboratory, the French Agency for Food, Environmental and Occupational Health & Safety (ANSES)

Research Article

Keywords:

Posted Date: November 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11172387/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

RESEARCH

Describing 2008 PPR epizootic in Morocco through the lens of mathematical modeling: Uncovering its multiscale dynamics

Manal Nhili [1,2,3*], Stéphanie Bougeard [4], Mounir Khayli [5], Noha El Khattabi [6], Mounia Abik [1] and Andrea Apolloni [2,3,4]

Correspondence:
nanal-nhili@um5r.ac.ma
ENSIAS, Mohammed V University in Rabat, Rabat, Morocco
Full list of author information is available at the end of the article

Abstract

Background: Peste des petits ruminants (PPR), a devastating viral disease of sheep and goats, poses a continuous threat to agricultural economies and food security. Effective control of PPR virus (PPRV) relies on an understanding of its spatial transmission dynamics. However, the lack of mobility data remains a major obstacle to accurately modeling and predicting its spatial spread. Here, we developed a stochastic metapopulation model and fitted to outbreak data from the 2008 epizootic in Morocco to reproduce spatio-temporal pattern and unravel the multi-scale nature of disease spread. In the absence of mobility records, we used the ST-DBSCAN spatio-temporal clustering algorithm to identify epidemiological clusters and Wallinga–Teunis approach to establish connectivity among outbreaks. Transmission parameters were inferred via Truncated Sequential Neural Posterior Estimation (T-SNPE).

Results: By applying a hierarchical framework across the nine affected administrative regions, we found significant geographical heterogeneity: depending on the region, local daily transmission rates ranged from 0.44 to 0.70, and intra-regional rates from 0.13 to 0.31, driving regional basic reproductive numbers (R_0) ranging from 3.09 to 4.91. We show that a clear spatial hierarchy governed the epizootic: while relatively rare weekly inter-regional seeding events (rate $= 0.03$) acted as the primary driver introducing the pathogen to distant areas, the epizootic was sustained by daily local and intra-regional spread.

Conclusions: The added value of this work is both scientific and practical. It provides new quantitative evidence on a major PPR epizootic and presents a transferable framework for studying disease spread under data-limited conditions. Although our case study focuses on Morocco, the framework can be adapted to other countries facing similar limitations in animal-mobility data. Furthermore, the estimated transmission parameters provide empirical benchmarks for researchers developing PPR models and may serve as informative prior information for model calibration and comparative analyses when direct estimates are unavailable.

Keywords: Peste des petits ruminants; Metapopulation model; Truncated Sequential Posterior Estimation; Simulation-based inference; Livestock disease epidemiology; Missing data reconstruction; Spatial transmission dynamics; Outbreak modeling; Morocco; Transboundary animal disease

Background

Transboundary animal diseases (TADs) are defined by FAO as those diseases that are of significant economic, trade and/or food security importance for a considerable number of countries; which can easily spread to other countries and reach epidemic proportions; and where control/management, including exclusion, requires cooperation between several countries . From this point of view, PPR is a classic example of a TAD. PPR is a contagious viral disease that affects primarly sheep and goats and some wild ruminants. It is caused by a morbillivirus in the family *Paramyxoviridae*, closely related to the viruses that cause rinderpest in cattle and measles in humans [1]. Infected animals develop fever, ocular and nasal discharge, erosive lesions in the mouth, and severe diarrhoea. Mortality can reach 90% in fully susceptible goat herds and around 50% in sheep [2]. PPR transmits mainly through direct contact, and movements of herds between pastures, markets, or during the seasonal transhumance, is the principal driver of spatial spread. The disease is endemic across large parts of Africa, the Middle East, Central and South Asia, and parts of East Asia (China). Between 2018 and 2025, outbreaks were confirmed in Bulgaria, Greece, Romania, Albania, Kosovo, and Croatia. Consequently, the European Parliamentary Research Service has classified PPR as an emerging transboundary risk, explicitly citing seasonal transhumance across the Mediterranean basin and susceptible wildlife as vectors for northward spread [3]. The Food and Agriculture Organization of the United Nations (FAO) and the World Organization for Animal Health (WOAH) estimate that PPR threatens the livelihoods of over 300 million pastoral and agro-pastoral households worldwide, causing annual losses of between USD 1.4 and 2.1 billion [4]. An efficient vaccine against PPRV exists since the the early 1990s [4, 5]. In 2015, FAO and WOAH launched a global campaign to eradicate PPR by 2030. Progress has been made in some regions but the disease remains endemic in many countries, and new outbreaks continue to be reported each year [6].

Four PPR lineages exist, with the same serotype, and until the first decade of the XXI century their distribution was geographically characterized: lineage I was predominant in West Africa, lineage II in East Africa, lineage III Middle East till India; lineage IV appeared in China. PPR was first reported in North Africa in Morocco in July 2008, with outbreaks subsequently confirmed in Algeria and Tunisia

[7], the strains collected belonging to lineage IV. The introduction of lineage IV marked a distinct shift in the region's epidemiology; unlike Mauritania, where PPR caused by lineages I and II has been endemic for decades [8], Maghreb countries had no prior exposure to this highly virulent strain. The rapid geographic expansion of the 2008 Moroccan outbreak pointed directly to animal movement as the primary early driver [9]. Furthermore, the geographical proximity of the Maghreb to Europe has since elevated PPR from a regional endemic to a continental concern.

Because of the naive status of national herds and the prominent role of livestock mobility, Morocco faced the challenge of controlling a disease whose expansion was driven by this mobility. Mass vaccination campaigns following the 2008 epidemic yielded high serological protection, and the country was considered PPR-free for several years. However, ten new outbreaks were reported across four provinces in 2015 [10], suggesting either undetected local persistence or, more likely, reintroduction via regional livestock movements. Survey-based network studies in Morocco have shown that sheep movements between breeder areas and fattening centers form a highly connected, scale-free network, with a few trading hubs capable of spreading a pathogen across the entire system [9, 11]. These information were used in previous studies to identify high-risk municipalities and periods of intense movement, such as the weeks preceding "Eid al-Adha". Yet, they describe movement patterns from survey data and do not model the transmission dynamics of a specific epidemic that could be useful to improve surveillance and study possible evolution scenario.

This work represents the first attempt to describe the 2008 diffusion of PPR in Morocco using a spatial metapopulation model to reproduce the spatio-temporal pattern of the PPR epizootic. Despite animal mobility is recognized as the principal driver of PPR spread, yet the data needed to model it are seldom publicly available in the very regions where the disease is most prevalent. In the absence of mobility data, we reversed the problem by leveraging information on outbreak locations and times, to unravel plausible connection patterns among regions in Morocco. Through preliminary analysis, we identified possible epidemiological clusters and the relations among locations within them. In a second step, we inferred epidemiological parameters of interest by calibrating the stochastic model to epidemiological data collected during the 2008 epidemics. Cluster analysis and calibration revealed the multi-scale nature of propagation and the role of local and inter-regional movements

in the diffusion of epidemics. Furthermore, the analysis revealed the unique role of the different clusters in controlling the propagation.

Methods

Study area and data sources

The present study focused on the 2008 epizootic of PPR in Morocco, corresponding to the first introduction of the disease in the country and involving the circulation of PPR virus lineage IV. At the moment of the epidemics, Morocco had an estimated population of 21.3 million sheep and 6.2 million goats, nationally distributed across a range of production systems: pastoral grazing in the south-western plains and mountainous regions, agro-pastoral systems in irrigated areas, oasis-based stall-feeding in the pre-Saharan zones, and semi-intensive systems near urban centers.

A total of 257 outbreaks were reported in 2008 between 12 June and 22 October, with a peak in week 31 (28 July-3 August) affecting 5,609 small ruminants and causing the death of 2,595. The National Veterinary authority started the mass vaccination campaign on 25 August 2008, reducing the number of outbreaks per week from 22.3 to 14.9. The epizootic lasted approximately 5 months. The period of the epizootic coincided with a period far from the peak of natality (late December-January) and Eid al-Adha (December 2008) and areas where transhumance was the main production systems, South central part of Morocco [12], were not touched.

An ensemble of statistical methods and dynamical models was used to describe the epizootic and estimate epidemiological parameters. Fig. 1 summarizes the workflow used to develop and calibrate the model. Cleaned geo-referenced data from WAHIS dataset, together with information of small ruminant distribution and market locations, were used to identify epidemiological clusters, areas that were connected to the same introduction event, and their characteristics (sub-section: spatio-temporal clustering). At the same time, transmission trees were reconstructed using the Wallinga–Teunis approach (sub-section: identification of transmission patterns). The analysis of the network resilience identified possible nodes to prioritize for control By aggregating at different spatial scale, we reconstructed the distribution of contacts between different localities in Morocco to inform the metapopulation model in form of "Who Acquire the Infection From Whom" (WAIFW) matrices. The latter,

were used to inform the stochastic metapopulation model to be calibrated against outbreak declarations (sub-section: spatio-temporal propagation).

Spatio-temporal clustering

For this analysis, we used the Spatio-Temporal Density-Based Spatial Clustering of Applications with Noise (ST-DBSCAN) approach, which was developed by Birant and Kut [13]. This approach creates clusters of spatio-temporal data according to non-spatial, spatial and temporal attributes by employing spatio-temporal distance functions. The algorithm can create neighboring clusters and distinguish between *core* points (those located in a dense region with enough nearby points in space and time) and *outliers* (those that do not belong to any cluster). While core points could indicate active areas of transmission, outliers could be related to sporadic events. More details about the parameters and choice of partitions are provided in Supplementary Section S1.

As pointed out by Lezaar et al. [9, 11] most of the outbreaks occurred in the area where livestock markets and fattening centers were concentrated. We associated to each outbreak location a livestock market (either rural or urban) based on spatial proximity, to assess if closeness to market could be a predictor to outbreak occurrence. Livestock markets geo-located data were provided by the National Office of Food Safety (ONSSA).

Identification of transmission patterns

Previous works has shown the important role of livestock mobility in the propagation of 2008 PPR epidemics. Not having information accessible for mobility pattern that could inform the metapopulation model, we used statistical tools to infer possible transmission trees and the relationships between the outbreak locations. For this, we used the Wallinga–Teunis approach [14] to reconstruct the probabilistic transmission tree, starting from the dates of outbreak declaration or symptomatic onset. The benefit of this approach is that it is data-driven and does not require any prior knowledge of contact patterns between locations. Based on previous work by Wallinga and Teunis and by Jombart et al. [14, 15]. Given a general interval distribution for each node, the likelihood that location i has infected location j is

given by the following formula:

$$p_{ij} = \frac{w(t_j - t_i)}{\sum_{k:t-k<t_j} w(t_j - t_k)} \quad t_i < t_j \tag{1}$$

where $w(t)$ is the kernel of the generation interval. The probabilities, p_{ij}, defined all the plausible transmission links at a given stage of the epidemic. The results are a probabilistic tree in which each link between outbreak locations corresponds to the probability that location i has infected location j. To account for the fact that most of the data were outbreak declarations per week, and to consider the delay between declaration and intervention (culling or mobility restriction) of up to one week, the generation interval distribution, needed to inform the model, was empirically estimated by the lag time between 2 consecutive declarations. Realizations of transmission trees were obtained by sampling a unique infector for each case according to p_{ij} and calibrating using MCMC. All analyses were conducted using the R package *outbreaker2* [16], which was developed explicitly for this type of analysis. Transmission trees provide a probabilistic link between the infecting and infected locations that should respect the temporal evolution of the pathogen. Based on the results of the calibration, we estimate the frequency with which a link respects the temporal evolution (i.e. the time of infection for the infected node should be later than that for the infector node) over all simulated trees, and use this as a proxy for the probability of a link connecting the infector and infected node existing. Transmission trees were then aggregated at the cluster and regional/commune levels to describe the patterns of interaction at the different spatial scales. In particular, WAIFW matrices (Who Acquires Infections From Whom) at cluster levels were used to understand the role of the different groups of nodes, while matrices at regional and communal levels were used to identify possible administrative hierarchies and calibrate the mechanistic model. Finally, to assess the role that the locations could have in PPR propagation we estimated the network's resilience when nodes are removed. In particular, we checked whether the network would remain connected after the removal of the nodes and all the satisfying certain properties (being of a specific cluster or region). This corresponds to the fact that the removal could stop the diffusion of the pathogens and create sub-networks that are isolated. The indicator is the greatest-connected components, that is the number of the nodes in

the largest sub-network a low value could give the maximum extent a disease could reach after removal.

Spatio-temporal propagation

The 2008 PPR epizootic in Morocco was modeled using a spatially SEIRD metapopulation framework in which each commune constituted a patch. We categorized transmission into three scales: local, intra-regional, and inter-regional (Fig. 2). This captured both short-range epidemiological coupling between neighboring villages through shared grazing (local transmission), and widespread circulation via strategic livestock markets that aggregate small ruminants from multiple regions before redistributing them (regional and inter-regional). The spatial unit of the model was defined as the commune, which is the smallest administrative unit in Morocco. Since outbreak locations were reported as Villages, a geocoding procedure was performed to assign each reported locality to it's corresponding commune and to the current administrative province and region. The transmission probabilities within and between populations were inferred from the calibration of the model. To this end we used outbreaks records extracted from the World Organization for Animal Health (WOAH)-World Animal Health Information System (WAHIS) historical reports.

We modeled local transmission dynamics using a continuous-time SEIRD model. While we modeled between-commune spread using a stochastic importation process with different probabilities depending if the communes were in the same region or not. More specifically, communes linked in the binary contact matrix from our clustering procedure had potential exposure once per week, matching the typical frequency of livestock market exchanges.

Local transmission model

For each commune i, animals were partitioned into susceptible (S_i), exposed (E_i), infectious (I_i), recovered (R_i), and deceased (D_i) compartments. The within-commune force of infection was defined as $\lambda_i^{\text{within}} = \beta_{\text{local},g(i)} I_i / N_i^{\text{live}}$, giving a within-commune incidence rate of $\lambda_i^{\text{within}} S_i = \beta_{\text{local},g(i)} S_i I_i / N_i^{\text{live}}$. Here, $g(i)$ denotes the epidemic-trajectory group, $N_i^{\text{live}} = S_i + E_i + I_i + R_i$ is the number of live animals in commune i, and $N_i^{\text{tot}} = S_i + E_i + I_i + R_i + D_i$ is the total population, including deceased animals. Deceased animals are excluded from the transmission term because they no longer contribute to infection dynamics. Exposed animals pro-

gressed to the infectious state at rate σ, infectious animals left that compartment at rate γ, and a proportion d of removals resulted in death, while the remainder entered the recovered compartment. Demographic turnover was omitted because the epidemic lasted less than five months, did not coincide with the main lambing or reproductive season, and unfolded over a time scale on which PPR-related mortality was assumed to dominate background births and natural deaths.

The within-commune dynamics were governed by:

$$\frac{dS_i}{dt} = -\frac{\beta_{local,g(i)} S_i I_i}{N_i^{\text{live}}}, \tag{2}$$

$$\frac{dE_i}{dt} = \frac{\beta_{local,g(i)} S_i I_i}{N_i^{\text{live}}} - \sigma E_i, \tag{3}$$

$$\frac{dI_i}{dt} = \sigma E_i - \gamma I_i, \tag{4}$$

$$\frac{dR_i}{dt} = (1-d)\gamma I_i, \tag{5}$$

$$\frac{dD_i}{dt} = d\gamma I_i, \tag{6}$$

Small ruminant population sizes at the commune level were obtained from the Gridded Livestock of the World version 3 (GLW3) dataset. Goats and sheep were modeled combined, as only limited number of goat-specific outbreaks were identified, and in Morocco, small-ruminant herds are commonly mixed.

The progression rate from exposed to infectious was fixed from the incubation period of 5 days as:

$$\sigma = \frac{1}{5}\ \text{day}^{-1} = 0.2\ \text{day}^{-1}, \tag{7}$$

and the removal rate from the infectious compartment was fixed from the infectious period of 7 days as

$$\gamma = \frac{1}{7}\ \text{day}^{-1} \approx 0.143\ \text{day}^{-1}. \tag{8}$$

The parameter $d \in [0,1]$ denotes the PPR case fatality proportion, so that a fraction d of infectious animals enter D_i and the remaining fraction $(1-d)$ enter R_i.

0.1 Regional and inter-regional spread model

To account for spatial spread in the absence of observed mobility data, a weekly stochastic importation step was applied prior deterministic integration.

Spatial spread between communes was modeled through a weekly stochastic seeding step applied before integrating the local deterministic dynamics. Let $T = (T_{ij})_{i,j=1}^{n}$ denote the $n \times n$ matrix of between-commune transmission rates, where n is the number of communes and T_{ij} quantifies transmission from commune j to commune i. At each simulation step, the between-commune force of infection acting on commune i was defined as:

$$\lambda_i^{\text{between}} = \sum_{\substack{j=1 \\ j \neq i}}^{n} T_{ij} \frac{I_j}{N_j}. \tag{9}$$

For a commune that had not yet been infected, the number of imported exposed animals during a given weekly step was drawn from a negative binomial distribution,

$$\Delta_i \sim \text{NegBin}(k, p_i), \tag{10}$$

with mean

$$m_i = \rho \, S_i \, \lambda_i^{\text{between}}, \tag{11}$$

and success probability

$$p_i = \frac{k}{k + m_i}, \tag{12}$$

so that

$$\mathbb{E}[\Delta_i] = m_i. \tag{13}$$

Whenever $\Delta_i > 0$, the state variables were updated as

$$S_i \leftarrow S_i - \Delta_i, \tag{14}$$

$$E_i \leftarrow E_i + \Delta_i, \tag{15}$$

The transmission matrix distinguished intra-regional, and inter-regional spread:

$$T_{ij} = \beta_{\text{local},\, g(i)}, \tag{16}$$

where $g(i)$ denotes the epidemic-trajectory group assigned to commune i.

For distinct communes in the same region,

$$T_{ij} = \beta_{\text{intra},\, g(i)} \quad \text{if } i \neq j,\ \text{region}(i) = \text{region}(j),\ \text{and } g(i) = g(j), \tag{17}$$

whereas for communes in different regions,

$$T_{ij} = \beta_{\text{inter}} \quad \text{if } i \neq j,\ \text{region}(i) \neq \text{region}(j). \tag{18}$$

Regions have different commercial patterns and raising systems that could impact the transmission probabilities. To reduce the number of free parameters, regions exhibiting similar epidemic curves were grouped together and shared coefficients were estimated for local and intra-regional transmission within each group, while a single coefficient was used for inter-regional transmission.

Calibration and Summary statistic

We estimated model parameters in two stages. First, we inferred the fraction of removals resulting in disease-related death (d) separately and fixed it across all regions and communes thereafter. This reduced calibration dimensionality and separated a parameter estimable directly from case and death counts from parameters informed only indirectly by the full model.

To estimate d, we specified a Bayesian binomial model in which observed deaths were modeled among reported cases with probability d. We assigned a $Beta(2, 2)$ weakly informative prior centered at 0.5 and sampled the posterior using the NO-U-Turn Sampler (NUTS) implemented with BlackJAX, with four chains, 1000 tuning steps, and 5000 draws per chain. We adopted this formulation because the likelihood was tractable and required no likelihood-free methods.

The resulting estimate was fixed in the transmission model.

The free parameter vector was denoted by:

$$\theta = \left(\{\beta_{\text{local},g}\}_{g=1}^{G}, \{\beta_{\text{intra},g}\}_{g=1}^{G}, \beta_{\text{inter}}, k, \rho\right), \tag{19}$$

where G is the number of regional epidemic-trajectory groups, $\beta_{\text{local},g}$ and $\beta_{\text{intra},g}$ are the group-specific local and intra-regional transmission rates, β_{inter} is the between-region transmission rate, ρ is the scaling parameter governing stochastic importation, k is the negative binomial dispersion parameter.

Because the model is stochastic and the observed data were compared with simulations through a derived summary statistic rather than through a tractable observation model, the likelihood

$$p(\mathbf{s}_{\text{obs}} \mid \theta) \tag{20}$$

could not be evaluated in closed form. Parameter estimation was performed using simulation-based inference (SBI) [17] from the SBI toolkit [18], which aims to approximate the posterior

We used likelihood-free inference for this second step because the parameters were linked to data by a complex metapopulation simulator making the likelihood intractable. Calibration used outbreak timing across communes rather than raw case counts. Estimating stochastic epidemic-model parameters is difficult. The likelihood is often intractable because it requires summing over possible sequences of unobserved events that could have produced the data. Bayesian spatiotemporal models also help map disease risk and changing spatial patterns, particularly when surveillance is incomplete [19]. Approximate Bayesian Computation (ABC) has been the standard workaround for this problem in epidemiology for over a decade [20, 21] because it avoids likelihood computation at the cost of approximation. ABC is conceptually simple and widely used, but it can become computationally expensive when the parameter space is large. Recent advances in machine learning and neural networks have led to new calibration techniques at the interface between Bayesian statistics and deep learning. A more recent family of methods, collectively known as simulation-based inference (SBI), replaces the rejection step of ABC with a neural network that learns the mapping from data to parameters directly from simula-

tions [17]. We applied neural posterior estimation (NPE), a variant of SBI, with hand-crafted summary statistics derived from the epidemic data to estimate the transmission parameters of our metapopulation model. NPE estimates the posterior distribution using one or more neural networks trained on simulated data. NPE learns a conditional approximation to the posterior distribution, denoted ($\hat{p}_\phi(\theta \mid \mathbf{s})$), where θ represents model parameters and $\mathbf{s}$ represents simulated summary statistics. Although we use hand-crafted summary statistic here, NPE reduces reliance on ad-hoc distance functions [22, 23]. Benchmark studies across several domains have confirmed that NPE can produce well-calibrated posteriors with fewer simulations than ABC [24, 25] and has been applied to parameter estimation problems in cosmology [26], developmental neuroscience [27], and epidemiology [28, 29].

In this work, we used the truncated sequential neural posterior estimation (T-SNPE), which restricts later simulation rounds to regions with higher posterior support and controls posterior leakage outside the supported parameter space [30].

The posterior density estimator $\hat{p}_\phi(\theta \mid \mathbf{s})$) was parameterized using normalizing flows. As for the choice of the flow architecture, we opted for the neural spline flows (NSF) [31]. NSF uses monotonic rational-quadratic spline transformations instead of affine transformations, which increases flexibility for representing posterior distributions with non-Gaussian or irregular geometry.

We focused on first-outbreak weeks to preserve the spatial-temporal invasion chronology of the 2008 epizootic while remaining robust to under-reporting and uncertainty in local epidemic sizes. This summary statistic was chosen because the available surveillance data were more informative about the order and timing of first outbreak appearance across communes than about the full underlying transmission process.

For each infected commune, $Z_i(t)$ denotes the simulated cumulative incidence, including infections generated locally and exposed animals imported into the commune. Let $M_i(w)$ denote the number of exposed animals imported into commune i during week w. We define

$$Z_i(t) = \int_0^t \frac{\beta_{\text{local},g(i)} S_i(u) I_i(u)}{N_i^{\text{live}}(u)} \, \mathrm{d}u + \sum_{w \le t} M_i(w), \tag{21}$$

where the integral represents local transmission and the summation represents imported infections. Thus, $Z_i(0) = 0$.

The summary statistic was defined as the first ISO week at which the simulated cumulative incidence $Z_i(t)$ exceeded 90% of the observed cumulative number of cases recorded at the first reported outbreak week in that commune. Denoting by C_i^{obs} the observed cumulative number of cases at first notification in commune i, the simulated first-outbreak week was identified as

$$\tau_i^{\mathrm{sim}} = \min\left\{\tau \in \mathcal{W} : Z_i(\tau) > 0.9\, C_i^{\mathrm{obs}}\right\}, \tag{22}$$

where $\mathcal{W}$ denotes the set of simulated ISO weeks. The observed summary vector was therefore

$$\mathbf{s}_{\mathrm{obs}} = \left(\tau_1^{\mathrm{obs}}, \tau_2^{\mathrm{obs}}, \ldots, \tau_m^{\mathrm{obs}}\right), \tag{23}$$

for the set of m infected communes...

$$p(\theta \mid \mathbf{s}_{\mathrm{obs}}) \propto p(\mathbf{s}_{\mathrm{obs}} \mid \theta)\, p(\theta), \tag{24}$$

using only simulated pairs

$$\left\{\left(\theta^{(r)}, \mathbf{s}^{(r)}\right)\right\}_{r=1}^{R}, \qquad \theta^{(r)} \sim p(\theta), \quad \mathbf{s}^{(r)} \sim p(\mathbf{s} \mid \theta^{(r)}). \tag{25}$$

We chose this approach because it provides a direct approximation to the posterior distribution from simulated parameter summary statistic pairs, unlike approximate Bayesian computation (ABC), NPE does not require the specification of a distance threshold in summary-statistic space.

Calibration diagnostics

We assessed posterior calibration using Tests of Accuracy with Random Points (TARP) [32]. This diagnostic evaluates whether, parameter vectors were sampled

from the prior and used to generated synthetic datasets with the same simulator used in the inference pipeline. For the simulated dataset, 2000 posterior samples were then drawn from the trained posterior, and calibration was evaluated by checking whether the true parameter values were recovered at the expected nominal coverage levels. Repeating this procedure across 1000 prior-predictive replicates yielded the expected coverage probability ECP(α), which was compared with nominal coverage α.

Under good calibration, empirical and nominal coverage should agree closely over the full range of α. These results were summarized with the TARP curve, which plots empirical against nominal coverage, so that departures from the diagonal indicate miscalibration. Overall performance was further quantified by the area under the trust curve (ATC), where values near zero indicate good calibration, negative values indicate overconfidence, and positive values indicate underconfidence. A global Kolmogorov-Smirnov test was also used to compare posterior quantiles with the uniform distribution.

We performed a posterior predictive check (PPC) using the summary statistic defined above. Posterior predictive simulations were used to generate its distribution under the fitted model, which was then compared with the observed value for each commune. This evaluation was intended to determine whether the model captured the temporal ordering and pace of epidemic emergence represented in the data. Because the available observations did not allow a reliable fine-scale comparison of complete epidemic trajectories, this posterior predictive check was based on a reduced statistic. This choice was considered sufficient for the present analysis because the primary objective was to assess temporal dynamics of spread.

Evaluation metrics

Predictive and inferential performance were assessed using the discrepancy between observed and simulated first-outbreak weeks across infected communes. Pointwise error was quantified by the mean absolute error (MAE):

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}\left|\tau_i^{\mathrm{sim}} - \tau_i^{\mathrm{obs}}\right|, \tag{26}$$

and the median absolute error (MedAE),

$$\text{MedAE} = \text{median}\left(\left|\tau_i^{\text{sim}} - \tau_i^{\text{obs}}\right|\right), \tag{27}$$

which is less sensitive to outliers.

Linear agreement between observed and simulated first-outbreak weeks was assessed using Pearson's correlation coefficient, whereas motonic agreement was assessed using Spearman's rank correlation coefficient. Goodness of fit was additionally summarized using the coefficient of determination,

$$R^2 = 1 - \frac{\sum_{i=1}^{m}\left(\tau_i^{\text{obs}} - \tau_i^{\text{sim}}\right)^2}{\sum_{i=1}^{m}\left(\tau_i^{\text{obs}} - \bar{\tau}^{\text{obs}}\right)^2}, \tag{28}$$

Uncertainty quantification was evaluated through 95% credible interval derived from posterior predictive distribution. In addition, posterior calibration was assessed by coverage, i.e the proportion of observed values lying within nominal posterior intervals.

Results

Descriptive Analysis

A descriptive analysis of the observed data set revealed that epidemic curve exhibited a rapid increase in the number of cases during the early stages of the epidemic, followed by a gradual decline as control measures were implemented. The peak of cases occurred on July 21, 2008, almost 40 days after the first reported outbreak (Fig.3 (A)).

A total of 257 Outbreaks occurred in 127 municipalities spread across 9 administrative regions of Morocco, with the highest number of outbreaks recorded in the region of "Béni Mellal-Khénifra" (n = 71), spread across 21 municipalities, followed by the region of "Fès-Meknès" (n = 68), spread across 37 municipalities, and the Oriental region (n = 26), spread across 14 municipalities. The remaining outbreaks were distributed across various regions, some of which reported only a few cases (Fig. 3 (B)). The chronology of outbreak onset is shown in Fig. 3 (C), where Fès-Meknès region reported the first outbreak, followed by Casablanca-Settat region. The total number of cases was 5,609, with a total of 2,595 deaths; the population at

risk recorded in the WAHIS report was 47,080 goats and sheep. Casablanca-Settat region reported 1,585 cases and 1,073 deaths, with a population at risk of 11,930. It was followed by Fès-Meknès region, with 1,219 cases and 328 deaths, affecting a population at risk of 10,345 (Supplementary Fig. S1).

Most of the cases were concentrated close to markets (either rural or urban) between 10 and 15 km (Supplementary Fig. S2).

Cluster Identification

The ST-dbscan algorithm identified 6 clusters (id from 1 to 6) and a group of outliers (id 0) of different sizes and geographical extension (Table 1). In particular, we notice that with the exception of cluster 3, all the clusters are Multi-regions. Clusters 1 and 2 are the largest ones both in terms of locations and cases occurred. Furthermore, clusters 3 to 6 include outbreaks occurring in a short time window (around one week, while in clusters 1 and 2 outbreaks occurred in one month. Outlier outbreaks, instead, occurred all along the epidemics.

The distance of each outbreak location from the closest market has been evaluated per cluster (Fig. 4). Clusters 1, 2 and 4 contain several urban markets. The distance from rural markets in these clusters is approximately the same (around 10 km), whilst the distance from urban markets is generally larger for clusters 1 and 2. Clusters 3, 5 and 6 have a single urban market. For clusters 5 and 6, outbreak locations are closer to rural markets than for the other clusters, while for cluster 3, the opposite is true. For cluster 0, the outliers, the distance from rural markets is shortest in most of the cases.

The spatial distribution of outbreaks varies by cluster (Supplementary Fig. S3). Clusters 3 and 4 have closer outbreaks and a more compact distribution. For the other clusters, the average distance between consecutive outbreaks is around 200 km; however, for clusters 1 and 2, consecutive outbreaks could also be more than 800 km apart. The distribution of outbreaks in the outliers group, however, is quite broad and, on average, larger than in all the other clusters. In terms of elapsed time, we observe that the time between two outbreaks is generally less than one week for all clusters, while for outliers, the distribution is broader, with some outbreaks occurring more than one month apart. This could be a sign that outliers act as

bridges between different clusters. After PPR diffuses in a cluster, the outliers are affected and spread the infection to another cluster.

We assessed the degree of spatial overlap between clusters by checking if the outbreak locations of each pair of clusters were located in the same commune, province or region (Supplementary Table S1). At the regional level, there is high overlap between clusters, i.e. outbreaks of different clusters occur in the same regions. However, cluster 3 does not share locations with clusters 4, 5 and 6. This could be due to the later appearance of clusters 5 and 6, and the earlier appearance of cluster 4. At a finer scale, we observe that the degree of overlap between clusters 1 and 2 is the highest, with outbreaks occurring in the same 17 communes. For the other possible couples, the degree of interaction is null or very low. This could indicate that, for low overlap, interaction between clusters could be due to commercial movements, whereas interaction between clusters 1 and 2 could be due to local interaction between herds sharing the same resources.

We notice that from a temporal point of view, outbreaks in Clusters 1 to 6, with the exception of cluster 3, they don't overlap (Supplementary Fig. S4). Cluster 3 could be thought as a spin-off of cluster one and indicate a local outbreak following introduction, whilst locations of cluster 3 can be thought as backbone locations, acting as bridges between the different clusters. Finally, we notice that in clusters 1, 3 and 4, outbreaks occurred before the vaccination campaign started, while outbreaks in clusters 5 and 6 occurred after the beginning of the vaccination campaign, whilst the others, include outbreaks both before and after implementation of the vaccination campaign.

Based on the characteristics of the clusters, we could argue that outliers act as a sort of bridge between the different clusters and could be responsible for the introduction of the disease to new areas through random exchange. Cluster 4 was where the disease first appeared in relation to the introduction of the pathogen in Morocco. Clusters 1 and 2 represent two sub-communities of the production chain and each was involved in a specific wave of the PPR epidemic. Transmission occurred not only via nodes of Cluster 0, but also through local transmission in 17 communities. Cluster 3 is a spin-off of Cluster 1, related to movement towards the Oriental region. Most of the connections were among locations in a sub-region.

Finally, Clusters 5 and 6 originated from Cluster 2, but their extent was strongly affected by the vaccination campaign.

Transmission Tree and WAIFW matrices

As empirical generation time interval we used the average distribution of elapsed time between outbreaks in cluster: 71 % of the outbreaks occurring the same day, 18% one day after, 7% two days after, 1.3 % day 3 and the remaining between 4 and 14 days. The outbreaker algorithm has been run for 2 10^6 simulations. The results of the procedure is a directed network whose links are weighted by the probability of occurrence. In Supplementary Fig. S5 we show the transmission network after pruning links whose probability was less than 1%, the choice was made to eliminate noisy links, while maintaining the network connected.

WAIFW matrices and the role of clusters

We used the transmission tree as a proxy for contact between different locations, based on the assumption that movement was the main way of introducing disease, with infection indicating a direct or indirect connection between locations. After estimating the frequency of occurrence, we counted the number of links between nodes belonging to clusters and calculated their probability, dividing it by the size of the cluster to create the asymmetric matrix WAIFW, together with the average probability. These matrices encode the strength of the relations between clusters. Fig. 5 shows the WAIFW matrices between clusters. On the left, we show the average number of contacts and on the right, the probability of occurrence. We can see that not all clusters interact with each other directly, but rather through the mediation of clusters 1 and 2. In particular, the number of interactions is quite high between clusters 4 and 1, 2, and 3. This means that, for an infection occurring in cluster 4 to reach a location in cluster 6, it must first infect locations in clusters 1 and 2. However, the probability of establishing contacts is low (between 2 and 6%).

Table 2 summarizes global properties after nodes are removed by cluster. We observe that when the locations of cluster 0 are isolated, for example, the network remains connected and the disease can continue to propagate. Network column identifies the network configuration, with Original denoting the network reconstructed from the transmission trees. The Size and Edges columns give the numbers of nodes and edges remaining after nodes from a cluster or region are removed. Density in-

dicates network redundancy, while Components and GCC give the number of sub-networks and the size of the largest one, respectively. Mean Distance indicates the average number of steps between two locations. We notice that besides the number of nodes and edges the network maintain its characteristics almost invariant after removal. However, only when nodes of cluster 1 are removed, the network breaks down into four sub-networks, the largest of which contains 99 nodes. The removal of other clusters affects the structure of the network less.

WAIFW matrices and the role of administrative Units

Similarly, we estimated the contact matrix between regions (Supplementary Fig. S6). The matrices show that contacts within and between regions are really rare, except in the case of Drâa-Tafilalet, which has strong contacts with locations within the regions of Béni Mellal-Khénifra, Fès-Meknès and Rabat-Salé-Kénitra. We observed that the network remained almost intact when locations were removed, with very few locations remaining protected (Table 2).

The WAIFW matrix at commune level shows different patterns of contact between regions, but appears denser than at regional level (Supplementary Fig. S7). In general, low contact values at regional levels and dense contact network at communal level could indicate a hierarchical network of contacts, with movements being frequent at a lower scale and rarefying as the administrative level increases. Similarly, we assess the role of locations in the region in the diffusion of pathogens by removing locations from the region. The matrix at commune level will be used to inform the mathematical model and estimate the parameters.

Parameter estimation

We estimated the case fatality rate among infected individuals from the observed data using MCMC. On average, $d = 0.463$ with a standard deviation of 0.007 and a 95% HDI between 0.45 and 0.476.

We estimated the remaining parameters using TSNPE. We trained the NSF with 3,500 simulations across three sequential rounds. Supplementary Fig. S8 shows the resulting pairplots. Point estimates are presented in the Table 3. The inter-regional transmission rate (β_{inter}) is much lower compared to other parameters, indicating the stochastic nature of the seeding process. We noticed that at local level the local transmission values β_{local} is in the same range of values for the 3 groups of

regions producing a value of $R_0 > 4.2$, with Casablanca-Settat and Fès-Meknès having similar values, whilst for the other regions, the estimated values of R_0 is much lower. Nevertheless, intra-regional transmission rates vary between group of regions, indicating differences in the connections at regional level. In particular, the Béni Mellal-Khénifra and Rabat-Salé-Kénitra regions had the highest local and intra-regional transmission rates.

Calibration and validation

The TARP diagnostic confirmed the robust statistical calibration of the inferred posterior. The TARP curve closely aligned with the ideal value of zero yielding an Area-To-Curve (ATC) of -0.120, indicating perfect calibration with a slight tendency toward over-confidence (Supplementary Fig. S9). Additionally, the Kolomogorov-Smirnov (KS) test applied to the expected coverage probabilities (ECP) failed to reject perfect calibration (KS p-value = 1.0), demonstrating that the model's credible regions accurately reflect true parameter uncertainty mapping without exhibiting significant overconfidence or underconfidence.

We complemented the TARP diagnostic with classical PPCs. PPC confirmed that the calibrated model accurately captures the temporal onset of the epidemic, as the true first week of the first reported outbreak per commune aligned closely with the simulated summary statistic. Specifically, the observed initial outbreak week fell within the 95% credible interval of the posterior predictive distribution for most communes across all regions (Fig. 6), indicating that our model effectively reproduces the observed temporal patterns of outbreak emergence. The MAE between observed and predicted first outbreak weeks varied across regions, with values ranging from 1.96 weeks in Drâa-Tafilalet to 5.47 weeks in Tanger-Tétouan-Al Hoceima, reflecting differences in predictive performance across regions (Supplementary Table S2). These results suggest that our model is well-calibrated and captures the epidemic's temporal dynamics. However, variability in predictive performance across regions indicates that region-specific factors may influence prediction accuracy, and could be explored to refine the model.

Discussion

Despite the importance of mobility data for modeling disease spread, most available datasets and models are concentrated in high-income countries [33]. In Africa, de-

spite growing interest in recent decades, empirical livestock mobility data are scarce and rarely available at subnational scales [34]. At the same time, relying on mobility proxies such as gravity models calibrated in other countries could be misleading [34]. In this work, we reverse the approach, trying to infer possible contacts and movements between locations in Morocco from outbreak data, and use this information to calibrate a metapopulation model to extract epidemiological parameters of interest. Although this approach is not detailed or mechanistic, it provides a reliable and effective framework for dealing with missing information.

We developed a first model to describe the spatio-temporal diffusion of PPR in 2008 in Morocco. A distinctive aspect of our study is the use of DBSCAN clustering and Wallinga–Teunis method to define connectivity between communes in the absence of explicit mobility data. The binary connectivity matrix derived from this analysis identifies pairs of communes whose outbreaks were close in both space and time, and therefore likely to have been epistemologically linked. While this approach cannot capture the full complexity of livestock trade and transhumance networks, it offers a pragmatic alternative when movement data are not available. The cluster analysis identified 6 clusters, two (Clusters 1 and 2) correspond to different seeding processes and epidemic waves, while the outlier group (Cluster 0) remained active during the epidemic period. Studies of livestock mobility in Morocco [9] have primarily focused on commercial animal movements and have shown the regional character of mobility communities. These studies underscored the significance of trade networks in disease dynamics. However, the role of local transmission processes was not fully addressed because suitable data were missing. Our analysis provides greater detail on PPR dissemination. Clusters 1 and 2 overlap substantially across all spatial scales. However, few outbreaks occurred within the same communes. Most events also occured towards the end of Cluster 1's epidemic wave. This pattern supports the hypothesis of a two-step transmission process. PPR may have been initially introduced through commercial movements into locations belonging to cluster 1. Subsequently, local transmission (at shared grazing areas or other local interactions between herds) could have seeded infection in nearby areas. These secondary foci may have then spread the disease further through commercial movements, resulting in the geographically distant outbreaks observed in cluster 2. The overlap between regions and communes is weaker for the other clusters.

This is more consistent with introductions and long-distance dissemination driven predominantly by commercial livestock movements. Moreover, the inference of the transmission tree, showed that interaction between and within region, were hierarchical, with , most of the interaction, between locations in the same region, and less with other regions, and highlighted the role of Draâ-Tafilalet as connecting the different regions.

The estimated transmission parameters revealed significant regional heterogeneity. Local transmission rates ranged from 0.44 to 0.70 per day, and intra-regional rates from 0.13 to 0.31 per day. These differences resulted in basic reproduction numbers ranging from 3.09 to 4.91. The group comprising Beni Mellal-Khenifra and Rabat-Salé-Kénitra had the highest R_0 value at 5.1. Next were the regions of Casablanca-Settat and Fès-Meknès, with nearly identical R_0 values of 4.7 and 4.5, respectively, followed by the group comprising Draâ-Tafilalet, Marrakech-Safi, the Oriental region, Souss-Massa, andTangier-Tetouan-Al Hoceima, which had the lowest R_0, around 2.9. Inter-regional transmission contributed little to the overall epidemic. The estimated inter-regional transmission rate was approximately 0.030 per week, roughly an order of magnitude lower than the local rates. We compared our estimates with those reported by Fournié et al. [35] by aligning parameters that share the same epidemiological meaning. Our local intra-patch transmission rate maps onto their within-village parameters (1.56 in lowlands and 6.19 in highlands). Similarly, our intra-regional rate corresponds to their inter-village transmission within the same ecological zone (1.08×10^{-3} for lowland-to-lowland and 1.19×10^{-3} for highland-to-highland), while our inter-regional rate aligns with their lowland-to-highland parameter (0.50×10^{-3}).

Our findings point to a distinct spatial hierarchy. Following the same hypothesis of above, inter-regional jumps likely acted as the primary seeding mechanism, introducing the virus to distant areas, but the epizootic was sustained by local and intra-regional transmission—a dynamic directly reflected in our modeling framework. This pattern is consistent with the findings reported for Ethiopia by Fournié et al. [35], where the authors noted that local transmission was the primary driver of viral persistence, and with the network analysis reported by Lezaar et al. [9], which found that most sheep movements occurred within or between neighboring provinces. Control measures implemented by the Moroccan authorities, particularly

movement restrictions and the rapid deployment of vaccination teams, were effective at limiting inter-regional spread.

Finally our estimates of R_0 values fall within the range previously reported for sustained PPR transmission in [8] and [35]. These comparisons should be interpreted cautiously given structural differences between the models, but also because of the viral lineages involved. Our study focuses entirely on PPRV lineage IV known to be highly contagious compared to earlier lineages. The Ethiopian setting in [35], however, likely reflected co-circulation of both lineage III and lineage IV, and the Mauritanian one was lineage II. Because lineage IV is more transmissible, our higher estimated transmission rates may reflect the dominant viral lineage just as much as differences in spatial connectivity or model assumptions.

The fact that the model produced reasonable estimates of R_0 and reproduced the observed epidemic curve suggests that spatiotemporal outbreak proximity contains useful information about underlying contact patterns. This is consistent with the broader observation that, in many infectious disease systems, spatial proximity is a reasonable proxy for transmission risk [36].

Our results highlight the importance of local and intra-regional transmission driving PPR epidemics in Morocco. This has direct implications for control policy. Surveillance should be strengthened in high-risk regions, and rapid detection and response should be prioritized in areas with dense livestock populations and active trade networks. The identification of critical nodes and cut points in the animal movement network, as described by Lezaar et al. [9, 11], provides a concrete basis for targeting interventions. The estimated R_0 values have direct implications for vaccination strategy. The herd immunity threshold for a disease with R_0 of 3 to 5 is between 67% and 80%. This means that, in theory, vaccinating at least 80% of the small ruminant population in high-risk regions would be sufficient to prevent sustained transmission. Morocco's mass vaccination campaigns between 2008 and 2010 reportedly reached over 20 million animals per year, which likely exceeded this threshold at the national level and contributed to the eventual control of the epidemic. However, our results suggest that vaccination efforts should be prioritized in regions with the highest R_0, where even moderate gaps in coverage could allow the virus to circulate and reduce the impact , to hose those locations identified as member of cluster 1 and 2 to hinder the spatial propagation. This targeted approach

aligns with the risk-based strategy advocated by ElArbi et al. [8] for Mauritania and with the findings reported by Lezaar et al. [11], who showed that most historical PPR outbreaks in Morocco occurred in a limited number of high-risk communes.

Several limitations should be acknowledged. First, the analysis relies on officially reported outbreaks, and under-reporting is likely, particularly in remote pastoral areas with limited access to veterinary services. Missed outbreaks could bias parameter estimates. Second, the binary connectivity matrix is a simplification. Real livestock networks involve directed, weighted flows that vary over time. A survey-based network, such as the one described by Lezaar et al. [11], would provide a more realistic representation of contacts, but such data were not available during the 2008 epidemic period. Third, the model assumes homogeneous mixing within each municipality and constant transmission rates throughout the epidemic. In reality, contact patterns may have changed as control measures were introduced and as farmers altered their behavior during the outbreak. Fourth, the model does not explicitly incorporate vaccination, which began during the epidemic. The estimated parameters therefore reflect a mix of vaccinated and unvaccinated populations. Despite these limitations, the model fits the observed data well and provides plausible estimates of transmission parameters.

Future research should aim to integrate real-time mobility data, such as GPS tracking of transhumant herds or digital market records, into transmission models for Morocco. Explicitly modeling the effects of vaccination, including partial and waning immunity, would also improve the accuracy of simulations and support the design of optimal vaccination strategies. Finally, extending the model to include cross-border connections with neighboring countries could inform regional eradication efforts, as the risk of virus reintroduction across land borders remains a concern.

Conclusions

We reconstructed the 2008 PPR epidemic in Morocco using outbreak notifications from WAHIS and a stochastic metapopulation model. The model was calibrated without explicit livestock movement data. relying instead on a binary connectivity matrix inferred from the spatiotemporal clustering of outbreaks. Despite this sim-

plification, the model reproduced the observed epidemic temporal dynamics and produced regional estimates of transmission rates and basic reproduction number.

These findings carry practical implications for the control of PPR in Morocco and in similar settings. The heterogeneity in R_0 across regions suggests that a uniform national vaccination strategy may be less efficient than a risk-based approach that concentrates resources in high-transmission areas. The effectiveness of Morocco's past mass vaccination campaigns in containing the 2008 epidemic and preventing resurgence until 2015 supports the value of sustained, high-coverage vaccination. However, the 2015 re-emergence of the disease also shows that surveillance and biosecurity cannot be relaxed once vaccination targets are met.

The modeling approach of our study shows that meaningful epidemiological insights can be gained from outbreak data alone, even in the absence of detailed mobility information. The combination of DBSCAN-based and Wallinga–Teunis connectivity inference with neural posterior estimation provides a framework that could be applied to other diseases and other countries where data availability is limited. Future work should focus on integrating detailed mobility data when they become available, explicitly modeling vaccination effects, and extending the spatial scope of the model to include cross-border dynamics. As the global community works towards the 2030 eradication target, such data-driven tools can help veterinary services allocate limited resources more effectively and protect the livelihoods of communities that depend on small ruminants.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The code and processed data generated in this study are available at:
https://github.com/mnhili/ppr_morocco_metapop

Competing interests

The authors declare that they have no competing interests.

Funding

This work received co-funding from the European Union's Horizon Europe, Grant agreement 101136346: European Partnership Animal Health and Welfare, Joint Internal Project SOA10. This work was supported also by the PHC

Toubkal Franco-Moroccan partnership through Campus France and the CNRST [grant number: Toubkal/23/158 – Campus France: 4858YB], awarded to author Manal Nhili.

Authors' contributions

M.N.: Conceptualization, Methodology, Software, Formal analysis, Data curation, Visualization, Writing – original draft, Writing – review & editing. **S.B.**: Formal analysis, Software, Validation, Writing – review & editing. **M.K.**: Data curation, Validation, Writing – review & editing. **N.E.K.**: Validation, Supervision, Writing – review & editing. **M.A.**: Validation, Supervision, Writing – review & editing. **A.A.**: Conceptualization, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing, Project administration, Supervision, Funding acquisition. All authors read and approved the final manuscript.

Acknowledgements

The authors gratefully acknowledge the National Office of Food Safety (ONSSA) for providing the livestock market location data used in this study.

Author details

[1]ENSIAS, Mohammed V University in Rabat, Rabat, Morocco. [2]UMR ASTRE, CIRAD, Montpellier, France. [3]UMR ASTRE, CIRAD, University of Montpellier, Montpellier, France. [4]Ploufragan-Plouzané-Niort Laboratory, the French Agency for Food, Environmental and Occupational Health & Safety (ANSES), Ploufragan, France. [5]National Office of Food Safety (ONSSA), Rabat, Morocco. [6]Faculty of Sciences, Mohammed V University in Rabat, Rabat, Morocco.

References

1. Parida S, Muniraju M, Mahapatra M, Muthuchelvan D, Buczkowski H, Banyard AC. Peste des petits ruminants. Veterinary Microbiology. 2015;181(1–2):90–106.
2. WOAH. Peste des Petits Ruminants. Paris, France: World Organisation for Animal Health; 2015. Available from: https://www.woah.org/fileadmin/Home/eng/Animal_Health_in_the_World/docs/pdf/Disease_cards/PESTE_DES_PETITS_RUMINANTS.pdf.
3. European Parliamentary Research Service. Peste Des Petits Ruminants: A Race against Time for Livelihoods and Livestock. European Parliament; 2025. Available from: https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2025)775865.
4. FAO & WOAH. Global Strategy for the Control and Eradication of PPR. Rome, Italy: Food and Agriculture Organization of the United Nations and World Organisation for Animal Health; 2015. Available from: https://openknowledge.fao.org/handle/20.500.14283/i4460e.
5. Diallo A, Taylor WP, Lefèvre PC, Provost A. Attenuation of a strain of rinderpest virus: potential homologous live vaccine. Revue d'Elevage et de Medecine Veterinaire des Pays Tropicaux. 1989;42(3):311–319.
6. WAHIS. World Animal Health Information System. World Organisation for Animal Health; 2026. Available from: https://wahis.woah.org/.
7. Baazizi R, Mahapatra M, Clarke B, Ait-Oudhia K, Khelef D, Parida S. Peste des petits ruminants (PPR): A neglected tropical disease in Maghreb region of North Africa and its threat to Europe. PLoS ONE. 2017;12.
8. ElArbi AS, Kane Y, Metras R, Hammami P, Ciss M, Beye A, et al. PPR Control in a Sahelian Setting: What Vaccination Strategy for Mauritania? Frontiers in Veterinary Science. 2019;6:242.
9. Lezaar Y, Khayli M, Coste C, Lancelot R, Bouslikhane M. Ovine Network between Fatteners and Breeders in Middle Atlas of Morocco: Where to Act to Prevent the Spread of Epidemics? Epidemiology – Open Journal. 2019;4(1):21–30. Available from: https://www.researchgate.net/publication/338480918_Ovine_Network_between_Fatteners_and_Breeders_in_Middle_Atlas_of_Morocco_Where_to_Act_to_Prevent_the_Spread_of_Epidemics.
10. Fakri F, Embarki T, Parida S, Bamouh Z, Jazouli M, Mahapatra M, et al. Re-Emergence of Peste Des Petits Ruminants Virus in 2015 in Morocco: Molecular Characterization and Experimental Infection in Alpine Goats. Veterinary Microbiology. 2016;197:137–141.
11. Lezaar Y, Boumalik M, Lhor Y, El-Ayachi M, Araba A, Bouslikhane M. Ovine Network in Morocco: Epizootics Spread Prevention and Identification of the At-Risk Areas for "Peste des Petits Ruminants" and "Foot and

Mouth Disease". Epidemiology – Open Journal. 2024;8(1):20–38. Available from: https://www.researchgate.net/publication/380006172_Ovine_Network_in_Morocco_Epizootics_Spread_Prevention_and_Identification_of_the_At-Risk_Areas_for_Peste_des_Petits_Ruminants_and_Foot_and_Mouth_Disease.

12. Akasbi Z, Oldeland J, Dengler J, Finckh M. Social and ecological constraints on decision making by transhumant pastoralists: A case study from the Moroccan Atlas Mountains. Journal of Mountain Science. 2012;9(3):307–321.
13. Birant D, Kut A. ST-DBSCAN: An algorithm for clustering spatial–temporal data. Data & Knowledge Engineering. 2007;60(1):208–221.
14. Wallinga J, Teunis P. Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures. American Journal of Epidemiology. 2004;160(6):509–516.
15. Jombart T, Cori A, Didelot X, Cauchemez S, Fraser C, Ferguson N. Bayesian Reconstruction of Disease Outbreaks by Combining Epidemiologic and Genomic Data. PLoS Computational Biology. 2014;10(1):e1003457.
16. Campbell F, Didelot X, Fitzjohn R, Ferguson N, Cori A, Jombart T. outbreaker2: a modular platform for outbreak reconstruction. BMC Bioinformatics. 2018;19(Supplement 11):363.
17. Cranmer K, Brehmer J, Louppe G. The Frontier of Simulation-Based Inference. Proceedings of the National Academy of Sciences. 2020;117(48):30055–30062.
18. Boelts J, Deistler M, Gloeckler M, Tejero-Cantero Á, Lueckmann JM, Moss G, et al. SBI Reloaded: A toolkit for simulation-based inference workflows. Journal of Open Source Software. 2025;10(108):7754.
19. Wang Y, Chen X, Xue F. A Review of Bayesian Spatiotemporal Models in Spatial Epidemiology. ISPRS International Journal of Geo-Information. 2024;13:97.
20. Beaumont MA, Zhang W, Balding DJ. Approximate Bayesian Computation in Population Genetics. Genetics. 2002;162(4):2025–2035.
21. McKinley TJ, Vernon I, Andrianakis I, McCreesh N, Oakley JE, Nsubuga RN, et al. Approximate Bayesian Computation and Simulation-Based Inference for Complex Stochastic Epidemic Models. Statistical Science. 2018;33(1).
22. Papamakarios G, Murray I. Fast ϵ-Free Inference of Simulation Models with Bayesian Conditional Density Estimation. In: Advances in Neural Information Processing Systems. vol. 29. Curran Associates, Inc.; 2016. p. 1028–1036. Available from: https://proceedings.neurips.cc/paper/2016/hash/6aca97005c68f1206823815f66102863-Abstract.html.
23. Greenberg D, Nonnenmacher M, Macke J. Automatic Posterior Transformation for Likelihood-Free Inference. In: Proceedings of the 36th International Conference on Machine Learning. vol. 97. PMLR; 2019. p. 2404–2414. Available from: https://proceedings.mlr.press/v97/greenberg19a.html.
24. Lueckmann JM, Boelts J, Greenberg DS, Gonçalves PJ, Macke JH. Benchmarking Simulation-Based Inference. arXiv; 2021. arXiv:2101.04653. Available from: https://arxiv.org/abs/2101.04653.
25. Wieland V, Wassmuth N, Contento L, Kühn M, Hasenauer J. Assessment of Simulation-Based Inference Methods for Stochastic Compartmental Models in Epidemiological Research. arXiv; 2026. arXiv:2512.02528. Available from: https://arxiv.org/abs/2512.02528.
26. Legin R, Hezaveh Y, Levasseur LP, Wandelt B. Simulation-Based Inference of Strong Gravitational Lensing Parameters. arXiv; 2021. arXiv:2112.05278. Available from: https://arxiv.org/abs/2112.05278.
27. Bernardo D, Xie X, Verma P, Kim J, Liu V, Numis AL, et al. Simulation-based inference of developmental EEG maturation with the spectral graph model. Communications Physics. 2024;7(1):255.
28. Chatha P, Bu F, Regier J, Snitkin E, Zelner J. Neural Posterior Estimation for Stochastic Epidemic Modeling. arXiv; 2024. arXiv:2412.12967. Available from: https://arxiv.org/abs/2412.12967.
29. Pinotti F, Thézé J, Bailly X, Fournié G. Simulation-Based Inference of Epidemiological and Phylodynamic Models via Neural Posterior Estimation. bioRxiv; 2025. 2025.11.25.690436.
30. Deistler M, Gonçalves PJ, Macke JH. Truncated Proposals for Scalable and Hassle-Free Simulation-Based Inference. arXiv; 2022. arXiv:2210.04815. Available from: https://arxiv.org/abs/2210.04815.
31. Durkan C, Bekasov A, Murray I, Papamakarios G. Neural Spline Flows. In: Advances in Neural Information Processing Systems. vol. 32. Curran Associates, Inc.; 2019. p. 7511–7522. Available from: https:

//proceedings.neurips.cc/paper/2019/hash/7ac71d433f282034e088473244df8c02-Abstract.html.

32. Lemos P, Coogan A, Hezaveh Y, Perreault-Levasseur L. Sampling-Based Accuracy Testing of Posterior Estimators for General Inference. arXiv; 2023. arXiv:2302.03026. Available from: https://arxiv.org/abs/2302.03026.
33. Leung WTM, Rudge JW, Fournié G. Simulating Contact Networks for Livestock Disease Epidemiology: A Systematic Review. Journal of the Royal Society, Interface. 2023;20(202):20220890.
34. Wardle J, Bhatia S, Kraemer MUG, Nouvellet P, Cori A. Gaps in Mobility Data and Implications for Modelling Epidemic Spread: A Scoping Review and Simulation Study. Epidemics. 2023;42:100666.
35. Fournié G, Waret-Szkuta A, Camacho A, Yigezu LM, Pfeiffer DU, Roger F. A Dynamic Model of Transmission and Elimination of Peste Des Petits Ruminants in Ethiopia. Proceedings of the National Academy of Sciences. 2018;115(33):8454–8459.
36. Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton University Press; 2008.

Figures

Figure 1 Methodological workflow for mutli-Scale PPR transmission dynamics inference.

Figure 2 Schematic representation of the metapopulation model structure. Each commune constitutes a patch with local SEIRD dynamics, and between-patch transmission is governed by a stochastic importation process informed by the contact matrix derived from spatio-temporal clustering.

Figure 3 National epidemic dynamics and regional burden in the Morocco 2008 outbreak. (A) National weekly epidemic curve showing total cases (solid red line with shaded area) and outbreak counts (dashed blue line), with peak weeks annotated. **(B)** Share of communes affected by the outbreak in each region, shown as affected communes over total communes. **(C)** Regional onset distribution of outbreak start dates, with violins summarizing temporal spread, individual outbreak records shown as points, and horizontal segments indicating the first-to-last outbreak period per region. Only infected regions are shown in panels B and C.

Tables

Additional Files

Additional file 1 — Supplementary Material

Additional information, figures and tables.

Figure 4 Distance in km of the outbreak location from the closest market. Color corresponds to the cluster, while shade indicates distance from rural or urban markets

Figure 5 Heatmap representing the number of contacts (on the left) and probability (on the right). Matrices are asymmetric and x indicates origin and y destination of the contact

Figure 6 Posterior predictive timeline by region Posterior predictive timelines for communes within each region, sorted by observed time to first detection (T1). For each region, the blue curve shows the posterior mean predicted T1 and the shaded band shows the 95% credible interval. Observed T1 values are shown as median points, with color indicating whether the observation falls within the credible interval; dotted vertical segments connect each observed value to its posterior mean prediction. The seed commune in Fès-Meknès is highlighted in the plot and excluded from the summary metrics reported in each panel title.

Table 1 Summary of the clusters identified using ST-DBSCAN (ID 1-6) and of the outliers nodes (ID 0). For each cluster we show the date of the first and last outbreak detected, the number of cases and deaths, the number of locations in the clusters, and the number of different administrative Units they belong to.

			Duration					
Cluster	**Cases**	**Deaths**	**First Dec.**	**Last Dec.**	**Locations**	**Communes**	**Provinces**	**Regions**
0	181	48	6/12/2008	10/22/2008	15	14	11	7
1	2617	1219	7/10/2008	8/5/2008	101	57	29	9
2	927	391	8/8/2008	9/10/2008	94	63	30	9
3	8	5	8/17/2008	8/18/2008	5	4	2	1
4	376	87	6/30/2008	7/3/2008	5	4	4	2
5	61	24	10/1/2008	10/5/2008	6	5	5	3
6	44	33	9/20/2008	9/24/2008	6	6	5	4

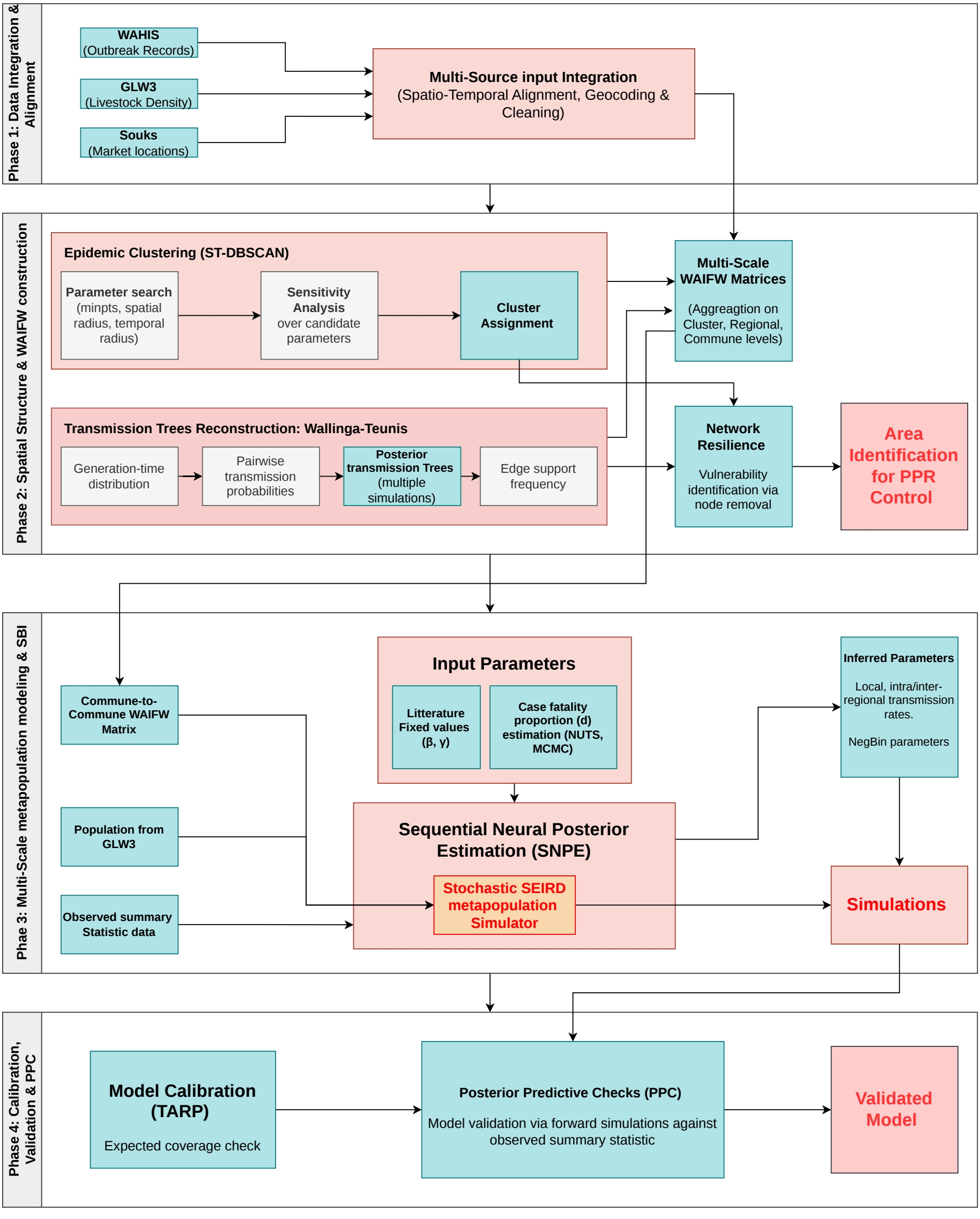
Table 2 Summary of network properties when nodes are removed according to specific criteria

	Network Scale		Topological Properties			
Network	**Size**	**Edges**	**Density**	**Components**	**GCC**	**Mean Distance**
Original	237	2148	0.0384	1	237	2.5090
Nocluster0	223	1900	0.0384	1	223	2.2929
Nocluster1	133	670	0.0382	4	99	2.2131
Nocluster2	140	761	0.0391	2	139	2.2315
Nocluster3	232	2021	0.0377	1	232	2.5276
Nocluster4	232	2057	0.0384	1	232	2.4994
Nocluster5	231	2057	0.0387	1	231	2.4260
Nocluster6	231	2081	0.0392	2	230	2.3718
NoRegionBeni Mellal-Khenifra	172	1258	0.0428	3	169	2.3247
NoRegionCasablanca-Settat	217	1796	0.0383	1	217	2.2979
NoRegionDaraa-Tafilalet	226	1946	0.0383	1	226	2.5144
NoRegionFes-Meknes	172	1085	0.0369	2	171	2.4053
NoRegionMarrakech-Safi	226	1969	0.0387	1	226	2.4679
NoRegionOriental	212	1698	0.0380	1	212	2.4187
NoRegionRabat-Sale-Kenitra	222	1909	0.0389	1	222	2.4773
NoRegionSouss-Massa	232	2049	0.0382	1	232	2.4992
NoRegionTanger-Tetouan-Al Hoceima	217	1789	0.0382	1	217	2.5702

Table 3 Posterior estimates of the model parameters and derivation of the basic reproduction number (R_0). Regional parameters are grouped according to the model's spatial stratification. Estimates are presented with their mean, standard deviation (SD), median, and 95% credible interval (CI). Parameters that do not depend on the region are global.

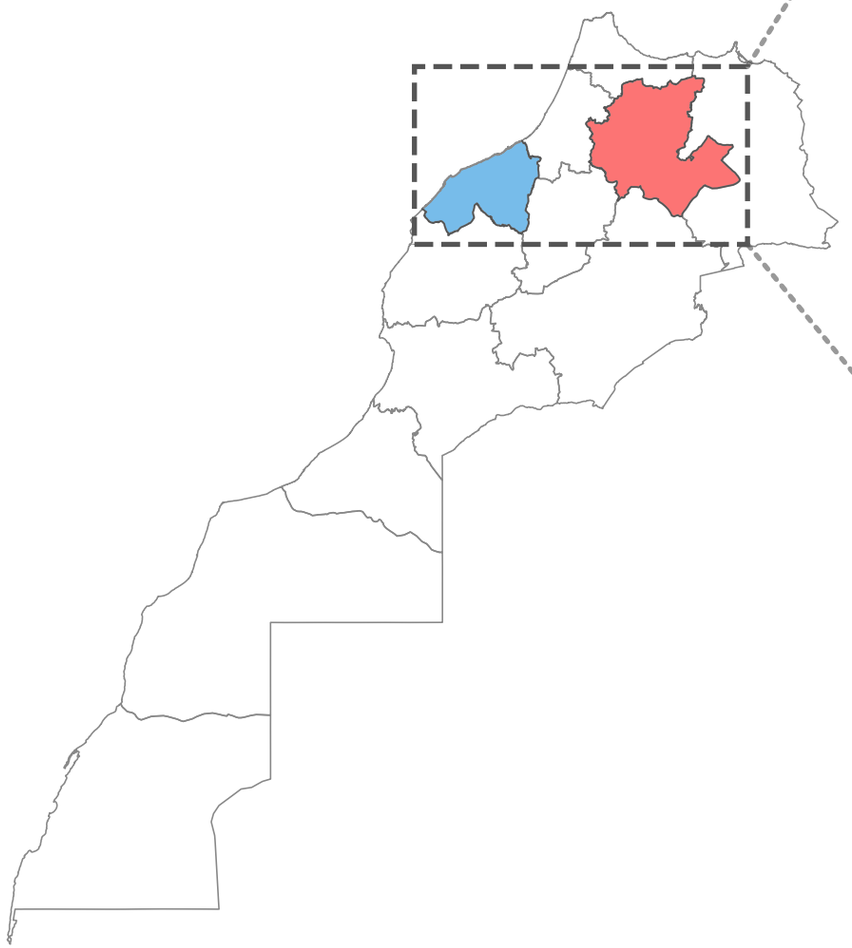
Parameter	Mean	SD	Median	[95% CI]
Global Parameters				
Inter-regional transmission rate (β_{inter})	0.030	0.026	0.022	[0.002 ; 0.092]
Spatial dispersion parameter (k)	2.000	1.543	1.536	[0.140 ; 4.882]
Scaling parameter (ρ)	0.180	0.117	0.143	[0.053 ; 0.462]
Region: Casablanca-Settat				
Local transmission rate (β_{local})	0.614	0.166	0.668	[0.208 ; 0.796]
Intra-regional transmission rate (β_{intra})	0.129	0.103	0.096	[0.013 ; 0.377]
Basic reproduction number (R_0)	4.299	1.165	4.678	[1.453 ; 5.573]
Region: Fès-Meknès				
Local transmission rate (β_{local})	0.613	0.143	0.644	[0.269 ; 0.794]
Intra-regional transmission rate (β_{intra})	0.189	0.122	0.177	[0.016 ; 0.399]
Basic reproduction number (R_0)	4.294	1.004	4.505	[1.883 ; 5.555]
Regions: Béni Mellal-Khénifra & Rabat-Salé-Kénitra				
Local transmission rate (β_{local})	0.701	0.099	0.732	[0.435 ; 0.798]
Intra-regional transmission rate (β_{intra})	0.307	0.095	0.339	[0.063 ; 0.408]
Basic reproduction number (R_0)	4.906	0.691	5.125	[3.047 ; 5.583]
Region: Other regions[§]				
Local transmission rate (β_{local})	0.441	0.213	0.407	[0.157 ; 0.789]
Intra-regional transmission rate (β_{intra})	0.127	0.103	0.095	[0.013 ; 0.377]
Basic reproduction number (R_0)	3.088	1.489	2.851	[1.102 ; 5.525]

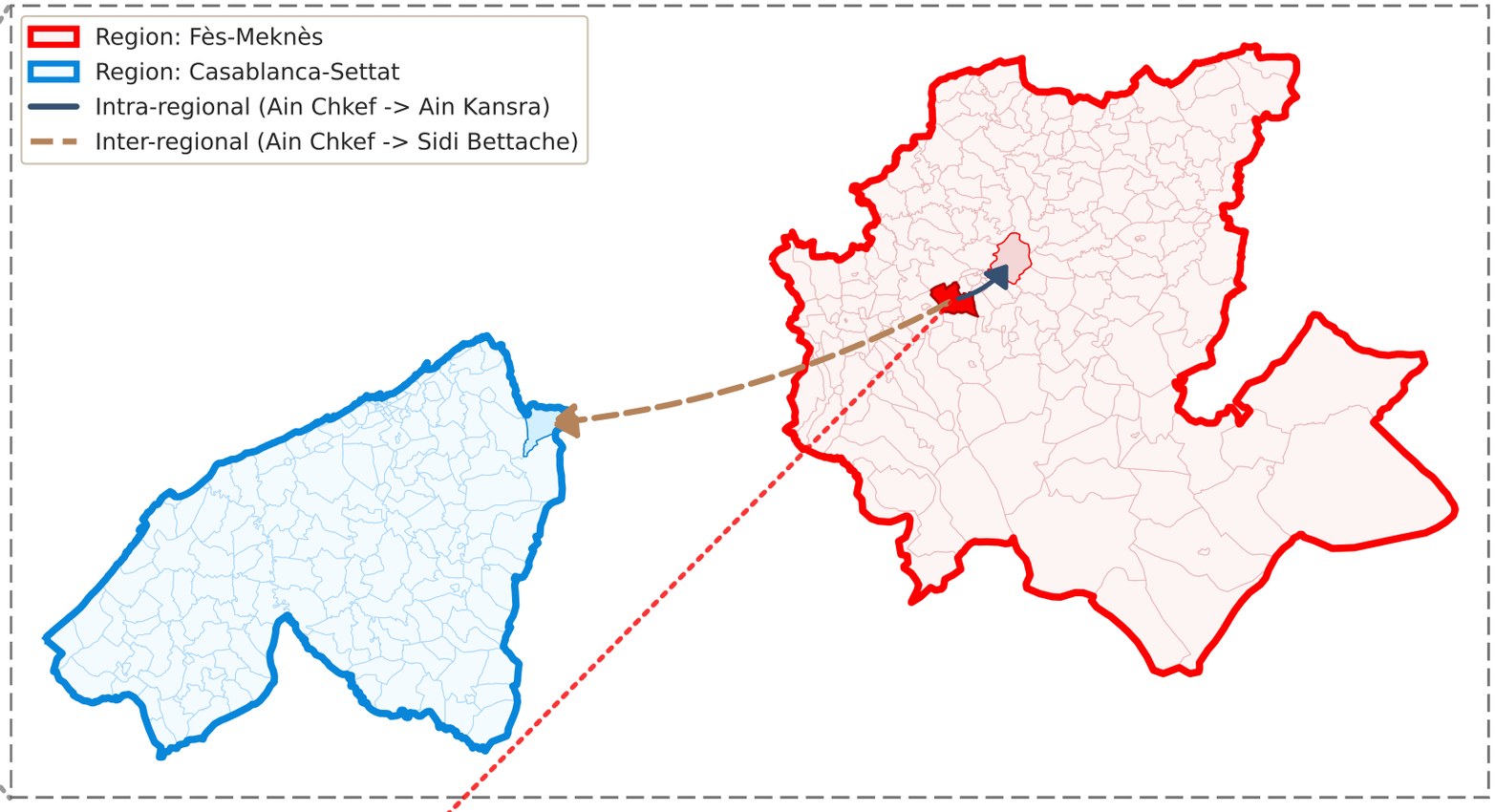
[§] Contains Drâa-Tafilalet, Marrakech-Safi, Oriental, Souss-Massa, and Tanger-Tétouan-Al Hoceima.


Phase 1: Data Integration & Alignment
WAHIS
(Outbreak Records)
GLW3
(Livestock Density)
Souks
(Market locations)
Multi-Source input Integration
(Spatio-Temporal Alignment, Geocoding & Cleaning)
Phase 2: Spatial Structure & WAIFW construction
Epidemic Clustering (ST-DBSCAN)
Parameter search
(minpts, spatial radius, temporal radius)
Sensitivity Analysis
over candidate parameters
Cluster Assignment
Multi-Scale WAIFW Matrices
(Aggreagtion on Cluster, Regional, Commune levels)
Transmission Trees Reconstruction: Wallinga-Teunis
Generation-time distribution
Pairwise transmission probabilities
Posterior transmission Trees
(multiple simulations)
Edge support frequency
Network Resilience
Vulnerability identification via node removal
Area Identification for PPR Control
Phae 3: Multi-Scale metapopulation modeling & SBI
Commune-to-Commune WAIFW Matrix
Population from GLW3
Observed summary Statistic data
Input Parameters
Litterature Fixed values (β, γ)
Case fatality proportion (d) estimation (NUTS, MCMC)
Sequential Neural Posterior Estimation (SNPE)
Stochastic SEIRD metapopulation Simulator
Inferred Parameters
Local, intra/inter-regional transmission rates.
NegBin parameters
Simulations
Phase 4: Calibration, Validation & PPC
Model Calibration (TARP)
Expected coverage check
Posterior Predictive Checks (PPC)
Model validation via forward simulations against observed summary statistic
Validated Model

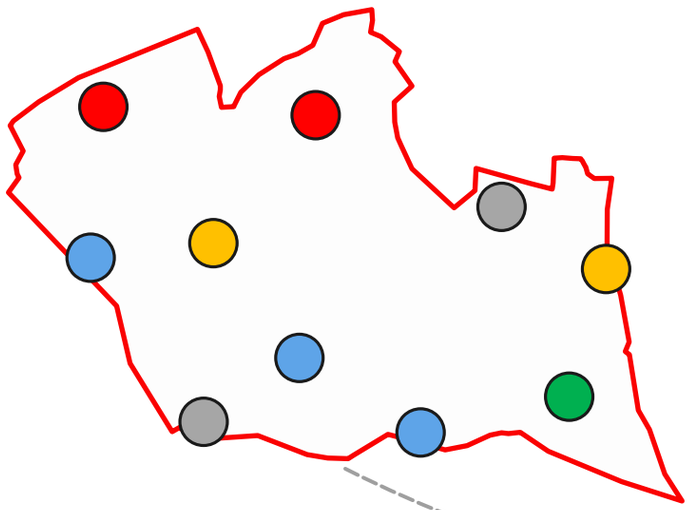
Transmission Scales

Geographic Context (Morocco)

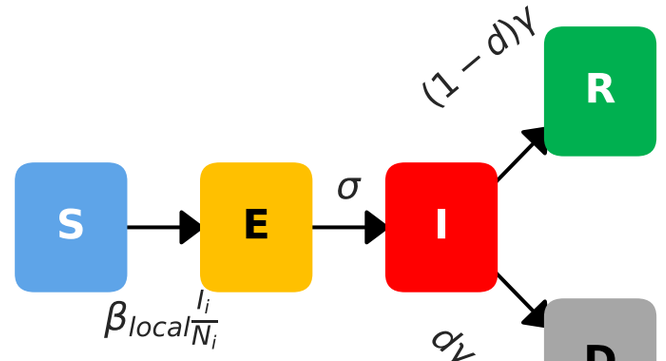




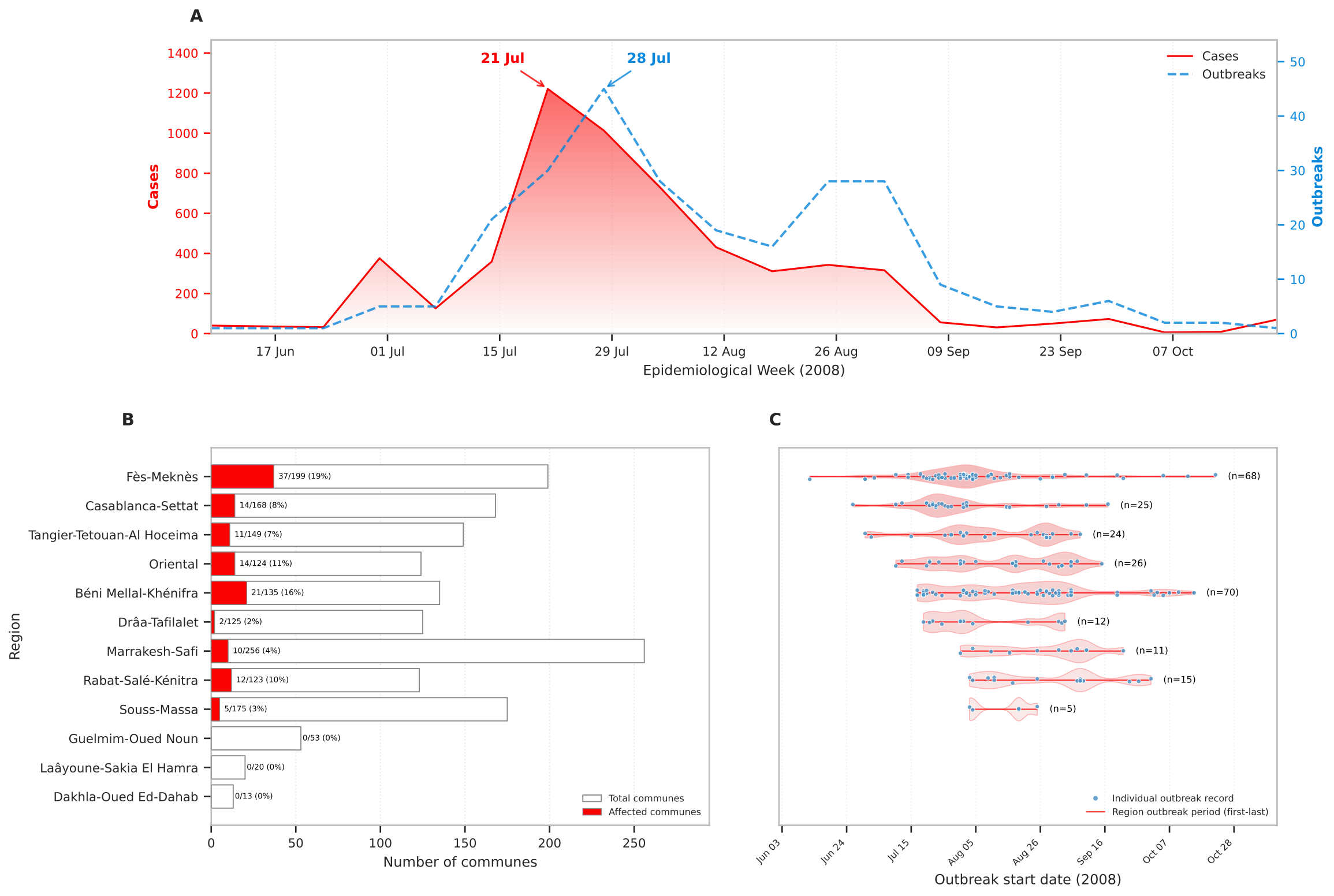
Local Population (Ain Chkef)

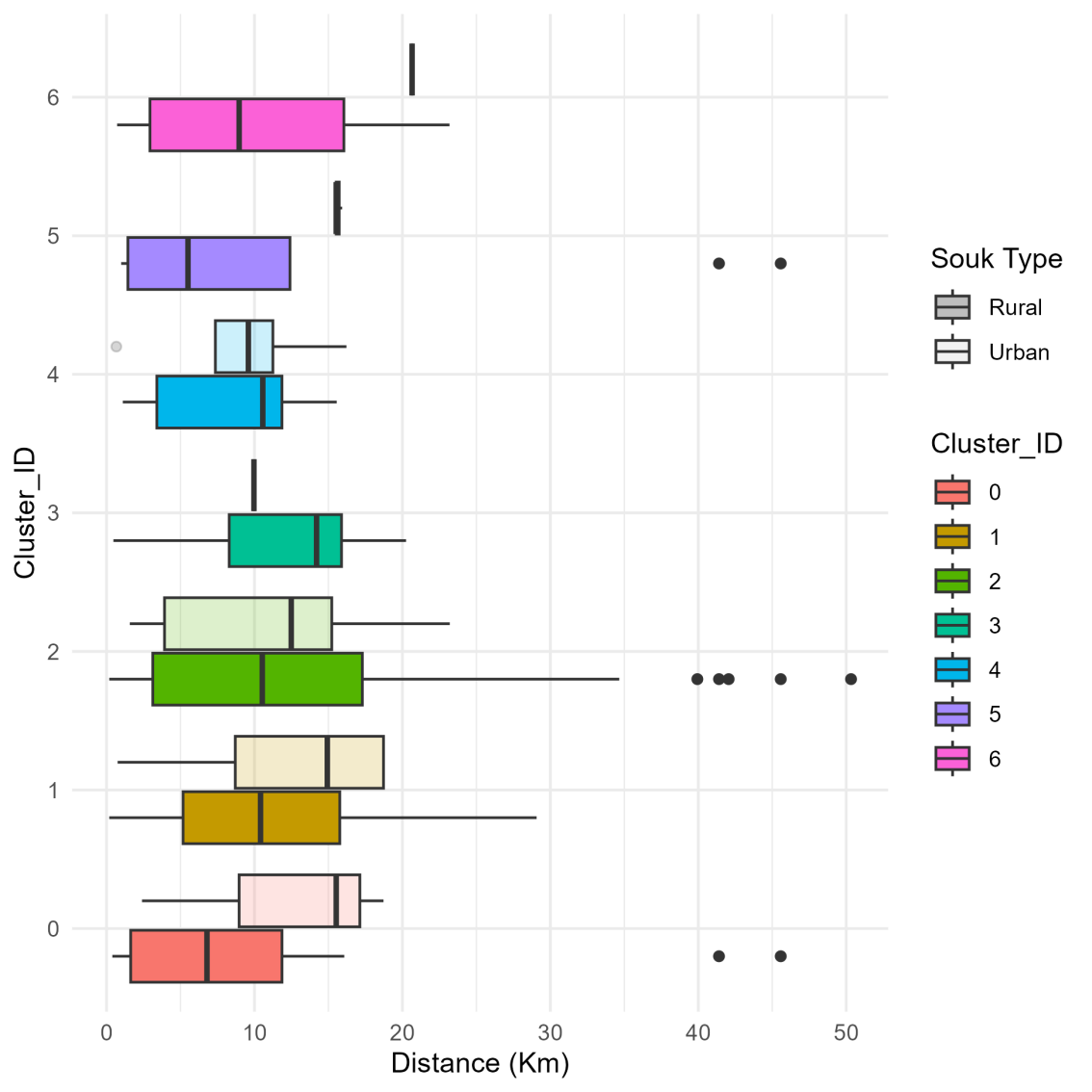


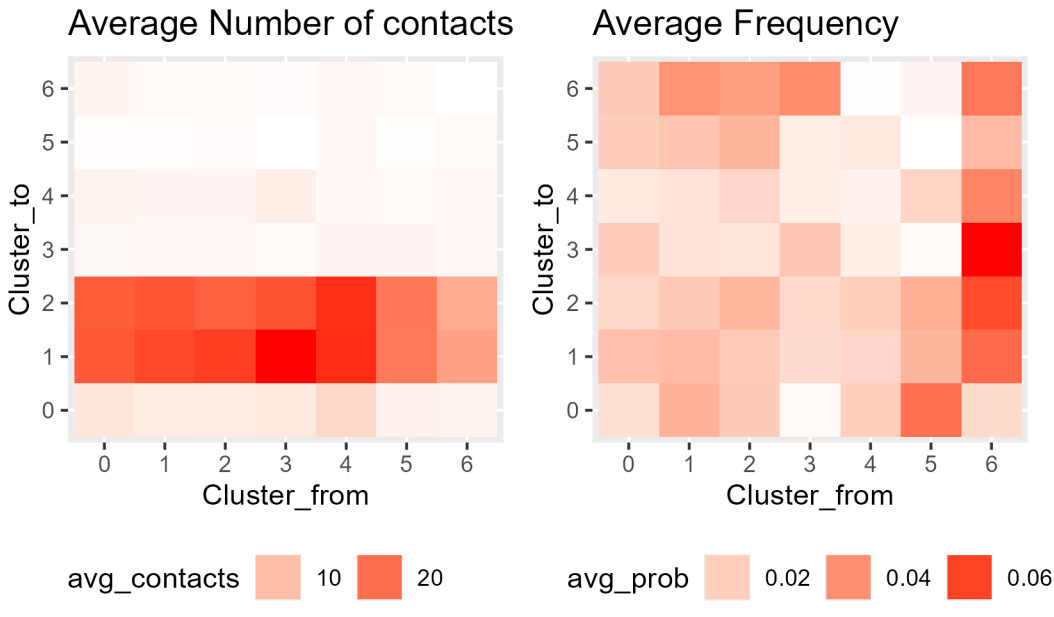
Local Dynamics (SEIRD)



Within-commune transmission


A
21 Jul
28 Jul
Cases
Outbreaks
Epidemiological Week (2008)
B
Fès-Meknès
Casablanca-Settat
Tangier-Tetouan-Al Hoceima
Oriental
Béni Mellal-Khénifra
Drâa-Tafilalet
Marrakesh-Safi
Rabat-Salé-Kénitra
Souss-Massa
Guelmim-Oued Noun
Laâyoune-Sakia El Hamra
Dakhla-Oued Ed-Dahab
37/199 (19%)
14/168 (8%)
11/149 (7%)
14/124 (11%)
21/135 (16%)
2/125 (2%)
10/256 (4%)
12/123 (10%)
5/175 (3%)
0/53 (0%)
0/20 (0%)
0/13 (0%)
Total communes
Affected communes
Number of communes
C
(n=68)
(n=25)
(n=24)
(n=26)
(n=70)
(n=12)
(n=11)
(n=15)
(n=5)
Individual outbreak record
Region outbreak period (first-last)
Outbreak start date (2008)


Cluster_ID
6
5
4
3
2
1
0
0
10
20
30
40
50
Distance (Km)
Souk Type
Rural
Urban
Cluster_ID
0
1
2
3
4
5
6


Average Number of contacts
Cluster_to
Cluster_from
0
1
2
3
4
5
6
avg_contacts
10
20
Average Frequency
Cluster_to
Cluster_from
avg_prob
0.02
0.04
0.06

Béni Mellal-Khénifra
95% Credible interval
50% Credible interval
Predicted τ_i^{sim} (median)
Observed τ_i^{obs}
Seed (initial commune)
50
40
30
Aguelmouss
Bni Oukil
El Hammam
Oulad Zmam
Tighassaline
Khmis Bni Chegdale
Sidi Yacoub
Ait Saadelli
El Borj
Oulad Boughadi
Oulad Abdoune
Lagnadiz
Bni Ykhlef
Had Bouhssoussen
Krifate
Fkih Ben Salah
Sidi Hcine
Bni Ayat
Oum Rabia
Oulad M'Barek
Sidi Amar
Casablanca-Settat
Sidi Bettache
Had Soualem
Sahel Oulad H'Riz
Koudiat Bni Dghough
Sidi Hajjaj Oued Hassar
Oulad Yahya Louta
Bouznika
Sidi Hajjaj
Laaounate
Moualine El Oued
Guisser
El Borouj
Sidi Bernoussi
Lambarkiyine
Drâa-Tafilalet
Aghbalou
Sidi Yahya ou Youssef
Fès-Meknès
Ain Chkef
Ain Kansra
Ouislane
Matmata
Dayat Aoua
Fès Médina
Aghbalou Agorar
Ben Smim
Sidi El Makhfi
Moulay Idriss Zarhoun
Ait Naamane
Oued Jemaa
Ain Leuh
Oulad Mimoun
Timahdite
Bni Snous
Ain Bida
Sebaa Rouadi
Mezraoua
Ait Boubidmane
Ain Bou Ali
Oued Ifrane
Rghioua
Ain Karma- Oued Rommane
Ait Sebaa Lajrouf
Tigrigra
Louadaine
Laanoussar
Kandar Sidi Khiar
Outat El Haj
Majjate
Almis Marmoucha
Ouizeght
Sidi Boutayeb
Enjil
Zrarda
Oulad Ayyad
Marrakech-Safi
Oued L'Bour
Lamzoudia
Oulad Zarrad
Labrikiyne
Jaafra
Benguerir
Oulad Hassoune Hamri
Saidate
Taouloukoult
M'Nabha
Oriental
Houara Oulad Raho
Aghbal
Lamrija
Bougriba
Sidi Moussa Lamhaya
Zegzel
Oujda
Ain Lehjer
Bni Khaled
Ain Sfa
Guenfouda
Ahl Angad
Sidi Boulenouar
Madagh
Rabat-Salé-Kénitra
El Kansera
Sidi Yahya Zaer
Boumaiz
Ain Attig
Brachoua
Sidi Mohamed Lahmar
Ait Ikkou
Lalla Mimouna
Zaggota
Arbaoua
Maaziz
Agdal Riyad
Souss-Massa
Igoudar Mnabha
Ouled Teima
Ouad Essafa
Sidi Bibi
Arazene
Tanger-Tétouan-Al Hoceima
Ouazzane
Moulay Ahmed Cherif
Sidi Boutmim
Beni Boufrah
Bni Faghloum
Chefchaouen
Tétouan
Souk L'Qolla
Zouada
Issaguen
Lamjaara
ISO weeks)
Time (
Commune

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterial.pdf