Preventive Veterinary Medicine 239 (2025) 106525

Contents lists available at ScienceDirect

Preventive Veterinary Medicine

journal homepage: www.elsevier.com/locate/prevetmed

Time series modelling of swine lung lesion prevalence to predict the temporal dynamics of *Mycoplasma hyopneumoniae* and *Actinobacillus pleuropneumoniae* infections in Spain

Amanda Fernández-Fontelo [a], María Teresa Lasierra-Morales [b], Marta Carmona [b], Marina Sibila [c], Laura Garza-Moreno [b,*]

[a] *Departament de Matemàtiques, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain*
[b] *CEVA Salud Animal, Barcelona, Spain*
[c] *IRTA, Centre de Recerca en Sanitat Animal (CRESA, IRTA-UAB), Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain*

ARTICLE INFO

Keywords:
Actinobacillus pleuropneumoniae Mycoplasma hyopneumoniae
Pig
Prediction
Prevention
Respiratory disease

ABSTRACT

Respiratory diseases are considered one of the most important problems in swine production worldwide due to the significant economic losses associated. Lung lesion evaluation at slaughterhouses by different scoring systems is commonly used to monitor respiratory diseases in swine. Concretely, cranioventral pulmonary consolidation lesions are associated with enzootic pneumonia (EP) caused by *Mycoplasma hyopneumoniae (Mhyo);* whereas haemorrhagic necrotizing pneumonia, mainly in the dorso-caudal lung lobes, and chronic pleuritis (CP) are associated with *Actinobacillus pleuropneumoniae* (*App*). Despite the recent consideration of several statistical methods for modelling the temporal dynamics of diseases and the construction of monitoring systems, none have been applied to lung lesions data collected at slaughterhouses. Thus, this work aimed (1) to describe the temporal patterns of EP and CP-like lesions in Spain using time series methods to model the collected data on lung lesions at slaughterhouses; and (2) to construct and evaluate in quasi-real time a surveillance system for early detection of outbreaks and abnormal trends potentially related to both pathogens. In total, 16 time series were analysed including 3947 audits from 474 Spanish farms associated with 302 companies between 2016 and 2019. The monthly time series of the EP index between 2016 and 2019 in Spain (point estimate for Spain was −0.088 with an associated $p = 0.073$) and different Spanish subregions showed decreasing trend patterns (point estimates for Aragon was −0.028 with an associated $p = 0.000$ and for Catalonia was −0.064 with an associated $p = 0.092$), whereas the monthly time series of the CP index increased (point estimate for Spain was 0.004 with an associated $p = 0.045$ and for Aragon was 0.007 with an associated $p = 0.000$) over the same period. Additionally, the predictive performance of the estimated models was evaluated at quasi-real time using the data between 2020 and 2021. Results from this evaluation showed that overall, the selected models predicted the evolution of both the EP and CP indices in a reasonable manner being between 90 % prediction intervals. Therefore, time series models constructed in this work could be used to prevent and shorten the response time in implementing of control strategies against these respiratory pathogens minimizing their economic impact associated.

1. Introduction

Respiratory diseases are considered one of the most important problems in swine production worldwide due to the significant economic losses associated with them. These losses are mainly due to reduced growth rates, increased feed conversion ratios as well as mortality rates, and elevated treatment costs, among others (Sorensen et al., 2006; Yaeger and Alstine, 2019).

The monitoring of lung lesions in pigs at the slaughterhouse is commonly used to estimate the prevalence and severity of different respiratory diseases as well as to measure their impact on the carcass market price (Fraile et al., 2010; Maes et al., 2023; Sibila et al., 2009). Several studies have identified cranioventral pulmonary consolidation (CVPC) and pleuritis as the most frequent lung lesions observed at the slaughterhouse (Arruda et al., 2024; Fraile et al., 2010; Meyns et al., 2011). Cranioventral pulmonary consolidation lesions are mainly

* Corresponding author.
E-mail address: laura.garza@ceva.com (L. Garza-Moreno).

https://doi.org/10.1016/j.prevetmed.2025.106525
Received 26 July 2024; Received in revised form 26 March 2025; Accepted 27 March 2025
Available online 29 March 2025

associated with enzootic pneumonia (EP), a respiratory disease caused by *Mycoplasma hyopneumoniae (Mhyo)* that has a negative impact on pig welfare and pig growth indicators (Arruda et al., 2024; Paz-Sánchez et al., 2021). CVPC lesions consist of purple to grey consolidated areas, mainly located in the apical, intermediate, accessory, and cranial parts of the diaphragmatic lobes (Maes et al., 2008). In addition, haemorrhagic necrotizing pneumonia, mainly in the dorso-caudal lung lobes, and chronic pleuritis (CP) are associated with *Actinobacillus pleuropneumoniae (App)* infections (Arruda et al., 2024; Merialdi et al., 2008; Ostanello et al., 2007). Both types of lesions are also economically important for slaughterhouses because they can lead to partial condemnations of carcasses and require additional labour on the slaughter line (De Conti et al., 2021; Maes et al., 2023).

Different scoring systems, based on visual examination or using computed techniques, can quantify the affected lung surface to evaluate the severity of EP- and CP-like lung lesions at slaughterhouses (Garcia-Morante et al., 2016; Ostanello et al., 2007). Although such lesions are not pathognomonic to *Mhyo* or *App,* the data generated by these slaughter check audits are commonly used to evaluate the respiratory health status of farms for both respiratory pathogens (Maes et al., 2023; Petri et al., 2023). Additionally, a description of the evolution of EP- and CP-like lesions over time could yield valuable insights into the infection dynamics of these pathogens (i.e., trend components and seasonality) as well as their prevalence (Fano et al., 2007; Petri et al., 2023).

Recently, the use of statistical methods, including methods of time series analysis (Brockwell and Davis, 2016; Cryer and Chan, 2008; Shumway and Stoffer, 2017; Montgomery et al., 2011; Peña et al., 2001), has gained increasing attention for the temporal and/or spatial-temporal monitoring of related data, e.g., certain diseases, mortalities, etc (Alba et al., 2015; Alba-Casals et al., 2020; Fernández-Fontelo et al., 2017, 2020; Ward et al., 2020). These methods, many of which have been comprehensively reviewed by, e.g., Benschop et al., (2008), can help identify potential abnormal peaks as well as help design (or redefine) adequate intervention strategies to mitigate the impact of, e.g., the disease under study (Keeling and Danon, 2009). Additionally, authors such as Benschop et al. (2008) have highlighted the potential of data collected at abattoirs for constructing health monitoring systems over time, primarily aimed at early identification of disease outbreaks and enhancing surveillance designs.

In the swine industry, a number of statistical modelling methods have been used to gain valuable knowledge into the dynamics of different pathologies, including the study of the transmission of *Mhyo* (Nathues et al., 2017), porcine reproductive and respiratory virus (Evans et al., 2010), Salmonella Typhimurium (Ivanek et al., 2004) and African swine fever (Barongo et al., 2016). In the literature, a limited number of works describe the temporal and/or spatial-temporal characteristics of infections related to EP and CP (Neumann et al., 2014; Eze et al., 2015). To the best of the authors' knowledge, there is no information available on the data collected on EP- and CP-like lesions at slaughterhouses from different Spanish regions and their modelling with time series analysis methods to describe the dynamics of both pathologies over time, or even their use for the construction of time series predictive models as surveillance systems for infections related to EP and CP in Spanish porcine populations.

Therefore, this study aimed to (i) describe the temporal dynamics of EP- and CP-like lesions at slaughterhouses at the Spanish level as well as at different sub-regions of Spain from 2016 to 2019 and (ii) construct and evaluate time series models to predict the evolution of both lung lesions as an early warning system to detect potential abnormal peaks that may be related to an increase in the incidence of both pathologies.

2. Materials and methods

2.1. Scoring systems used for the evaluation of EP- and CP-like lesions: the ceva lung program (CLP)

In this work, the Ceva Lung Program (CLP) was used to evaluate EP- and CP-like lung lesions at slaughterhouses (Krejci and Munoz, 2014). This program combines the Madec and Kobisch systems with Christensen's system to evaluate EP-like lesions (Madec and Kobisch, 1982; Christensen et al., 1999) and the Slaughterhouse Pleurisy Evaluation System (SPES) for CP like-lung lesions (Dottori et al., 2007). On the one hand, the CLP-CVPC system provides a score between 0 and 4 according to the percentage of the surface showing CVPC per lobe (i.e., 0: no lesion; 1: less than 25 % of the total lung lobe surface affected; 2: between 25 % and 50 % of the total lung surface affected; 3: between 50 % and 75 % of the total lung surface affected; 4: more than 75 % of the total lung surface affected). Additionally, the CLP program normalizes the scores with respect to the relative volume of each lobe, as proposed by Christensen et al. (1999). Finally, an EP index as an overall measure of the degree of EP-like lesions in a lung is obtained for each lung and pig batch (Table 1).

On the other hand, the CLP-CP ranges from 0 to 4 (excluding 1) depending on the presence and severity of caudal pleurisy in the lungs (i.e., 0: no lesions; 2: a single small focus of CP-like lesions in the caudal lobe; 3: a single large focus of CP-like lesions in the caudal lobe, small foci in both caudal lobes or a large focus in one of the lobes and a little focus of a CP-like lesion in the other caudal lobe; 4: a large CP-like lesion in both caudal lobes). Note that the CLP-CP slightly differs from the traditional SPES system since cranio-ventral pleuritis (scored as 1) is not considered for evaluation in the CLP-CP. Given that cranio-ventral pleuritis is also associated with different respiratory pathogens, such as *Pasteurella multocida* (Christensen et al., 1999; Enøe et al., 2002), cranio-ventral pleuritis lesions caused by other pathogens are considered separate binary variables in CLP indicating the presence (coded as 1) or absence (coded as 0) of such lesions in each evaluated lung. Finally, the CP-index is considered as the overall outcome for the degree of infection related to CP per pig and batch (Table 1).

Lung audits were carried out by independent veterinarians trained by Ceva Salud Animal (Spain). The training of the evaluators consisted of assessing and scoring the lesions in each of the lung lobes on the slaughter line, using the CLP system previously described. To avoid potential bias, all evaluators received the same training session from the same person to calibrate lung lesion scores. Evaluators were regularly examined by the trainer to ensure the reliability and reproducibility of the process.

2.2. Population of study and data sources

Lung batches from pigs had reached finishing weight of different Spanish farms and producers were audited for the presence of EP- and

Table 1
Summary of the indices considered for the evaluation of EP- and CP-like lung lesions.

Index	Description
EP	Sum of the CLP-CVPC scores for all lobes of all evaluated lungs in a batch divided by the total number of lungs being evaluated in that batch, i.e., the arithmetic average of the CLP-CVPC scores within the same batch. This index provides information on the presence and severity of EP-like lesions.
CP	Relative frequency of lungs with CLP-CP scores equal to or greater than 2 in a batch multiplied by the arithmetic mean of the CLP-CP scores in that batch. This index provides information on the presence and severity of CP-like lesions.

EP: Enzootic pneumonia; **CP**: Chronic pleuritis; **CLP**: Ceva Lung Program

CP-like lung lesions at different slaughterhouses. Audits were conducted between 2016 and 2021 and included information on audit identification number, date, slaughterhouse, REGA code for farm identification (it provides information related to the farm location), number of evaluated lungs per batch, and the EP and CP indices at the batch level. Information related to vaccination against *Mhyo* of these evaluated pig batches was obtained from the corresponding veterinarian/producer. No information regarding vaccination against *App* or other pathogens was recorded.

The EP and CP indices per batch were aggregated in the audit's dataset at monthly level by averaging the EP and CP indices of the evaluated batches within the same month and region. The time series of this study included all evaluated batches in Spain as well as 7 different Spanish regions (Catalonia, Aragon, Galicia, Norte, Centro, Levante and Sur). The aggregation of the information provided a total of 16 time series (for Spain and each of the seven Spanish regions considered, we had 2 time series at the month level: for the EP and CP indices) that were all independently modelled using classical time series analysis and, more specifically, Autoregressive Integrated Moving Average (ARIMA) processes.

2.3. Data modelling and forecasting

Univariate time series are random variables that are sequences of data points and are usually measured over discrete time intervals. One of the most popular models for continuous-valued time series is the ARIMA process, which is often used for describing and estimating the dynamic properties of a time series, as well as for forecasting.

More formally, let X_t be a univariate time series where $t \in \mathscr{T}$ is a discrete time interval, e.g., $\mathscr{T} = (1, \cdots, T)$ following an ARIMA(p, d, q) process. Let $Z_t = (1-L)^d X_t$ be the d-time differentiated X_t process, where L is the lag operator. Then, the process Z_t for $t \in \mathscr{T}$ follows a weakly stationary ARMA(p, q) process defined with the equation given below:

$$Z_t = \mu + \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \ldots + \varphi_p Z_{t-p} + \varepsilon_t + \vartheta_1 \varepsilon_{t-1} + \ldots + \vartheta_q \varepsilon_{t-q}, \quad 1)$$

where $Z_{t-k} = (1-L)^d X_{t-k}$ and parameters p and q determine the order of the autoregressive (AR) and the moving average (MA) parts, respectively. In addition, the process ε_t is defined as independent and identically distributed (iid) observations coming from a random variable with expectation 0 and constant variance σ^2. We denote it as $\varepsilon_t \sim WN(0, \sigma^2)$, where WN indicates white noise.

A weakly stationary (or only stationary) ARMA(p,q) process Z_t satisfies that its expectation is constant over time, i.e., $E(Z_t) = \mu_Z$ for all t, and its autocovariance function depends only on the temporal lag $|h| = |s-t|$ for all s, t, i.e., $\gamma(s,t) = Cov(Z_s, Z_t) = Cov(Z_{s+r}, Z_{t+r}) = \gamma(s+r, t+r)$ or, in other words, $\gamma(h) = Cov(Z_t, Z_{t+h})$ for all h; thus, $Var(Z_t) = \gamma(0)$ for all t.

In this study, all evaluated 16 time series were not stationary. In general, the monthly time series of the EP index at the batch level at either the country or the region level showed a clear decreasing linear trend over time (Figure S1), while the monthly time series of the CP index at the batch level at either the country or the region level showed an increasing linear trend over time (Figure S2). Thus, with these trend patterns, the distributions of the monthly EP and CP indices over time did not have a constant expectation, as the expectation of the EP index decreased from 2016 to 2019, while the expectation of the CP index increased from 2016 to 2019.

Non-stationary time series are frequent in real-world applications, and different methods for addressing them are available in the literature (Brockwell and Davis, 2016; Shumway and Stoffer, 2017). As most of our time series showed linear trends and seasonal components, we used the method described below to model our non-stationary time series using ARMA processes.

Let Y_t be a univariate time series where $t \in \mathscr{T}$ and $\mathscr{T} = (1, \cdots, T)$. Assume that $Y_t = m_t + s_t + X_t$, where m_t is the trend component of the time series Y_t, s_t is the seasonal component of the time series Y_t, and finally, $Z_t = (1-L)^d X_t$ is a stationary time series following an ARMA structure, as given in Eq. (1). There are a number of methods aimed at estimating the trend component m_t and the seasonal components s_t. For instance, the trend component, m_t, can be fitted using a parametric approach (e.g., estimating a polynomial, exponential, logistic trend) or a non-parametric approach with methods such as moving averages, a local polynomial smoother, an exponential smoother, and kernel smoothing, among others. In addition, we can use a wide range of methods for the estimation of the seasonal component, s_t, including the use of the trigonometric polynomial. Further details on classical methods for non-stationary time series modelling can be found in Brockwell and Davis (2016) and Shumway and Stoffer (2017).

In the current paper, we estimated the trend and seasonal components of our time series using a parametric approach based on a linear regression model with a linear trend and a trigonometric polynomial. More precisely, the following model was considered for our time series Y_t :

$$Y_t = \mu + \beta_1 t + \beta_2 \sin\left(\frac{2\pi t}{12}\right) + \beta_3 \cos\left(\frac{2\pi t}{12}\right) + \beta_4 \sin\left(\frac{2\pi t}{4}\right) + \beta_5 \cos\left(\frac{2\pi t}{4}\right) + \beta_6 \sin\left(\frac{2\pi t}{3}\right) + \beta_7 \cos\left(\frac{2\pi t}{3}\right) + X_t, \quad (2)$$

where $m_t = \beta_1 t$ is the linear trend, $s_t = \beta_2 \sin\left(\frac{2\pi t}{12}\right) + \beta_3 \cos\left(\frac{2\pi t}{12}\right) + \beta_4 \sin\left(\frac{2\pi t}{4}\right) + \beta_5 \cos\left(\frac{2\pi t}{4}\right) + \beta_6 \sin\left(\frac{2\pi t}{3}\right) + \beta_7 \cos\left(\frac{2\pi t}{3}\right)$ are the seasonal components and $Z_t = (1-L)^d X_t$ is a stationary ARMA process, as defined in Eq. (1). Note that part $\beta_2 \sin\left(\frac{2\pi t}{12}\right) + \beta_3 \cos\left(\frac{2\pi t}{12}\right)$ of the trigonometric polynomial considers an annual periodicity, part $\beta_4 \sin\left(\frac{2\pi t}{4}\right) + \beta_5 \cos\left(\frac{2\pi t}{4}\right)$ of such a polynomial considers a fourth-month periodicity, and part $\beta_6 \sin\left(\frac{2\pi t}{3}\right) + \beta_7 \cos\left(\frac{2\pi t}{3}\right)$ of the polynomial considers a quarterly periodicity. These periodicities were considered because they were the most frequently observed on the dynamics over time of the infections related to EP and CP, e.g., the 12-month periodicity was considered to cover the effect of temperature changes related to the seasons on both the time series of the EP and CP indices, while the 3-month and 4-month periodicities were considered to cover the effect on both the time series of the EP and CP indices due to aspects related to logistics and batch management in swine production, as the fattening period generally lasts for 3–4 months. Note also that while the time series Y_t is not stationary, the time series Z_t is stationary. Before choosing the values for the parameters p, d and q, we selected, for each considered time series, the trend and seasonal components to be included as exogenous variables in the ARIMA models. In particular, we estimated a classical linear model with the corresponding EP or CP index as the response variable and including a linear trend and all the seasonal components previously described as explanatory variables, i.e., we estimated the following linear model using the ordinary least squares method: $Y_t = \beta_0 + \beta_1 t + \beta_2 \sin\left(\frac{2\pi t}{12}\right) + \beta_3 \cos\left(\frac{2\pi t}{12}\right) + \ldots + \beta_6 \sin\left(\frac{2\pi t}{3}\right) + \beta_7 \cos\left(\frac{2\pi t}{3}\right) + \varepsilon_t$, $\forall t$, and selected those components that were statistically significant to include as exogenous variables in the ARIMA model. Then, for a given EP or CP index to be modelled, the considered ARIMA models with different values for the parameters p, d and q included the previously selected trend and seasonal components as exogenous variables.

Different versions of ARIMA models were considered by combining different values for the parameters p, d and q and estimating them using the maximum likelihood method (function `arima` from `R`). As described in the documentation for this function, the maximum likelihood method can handle missing values in the time series, which is particularly useful in our case as we have some missing values in some of our analysed time series. Values for parameters p and q ranging from 0 to 5 were explored and for parameter d ranging from 0 (no differentiation of the original time series) to 1 (first-order differentiation of the original time series). These candidate models were compared based on the lowest second-order Akaike Information Criterion (AICc), as our time series were of short length (Burnham and Anderson, 2002), selecting the model among all candidates with a small AICc value accompanied by a good validation of the model assumptions and model properties (e.g., residuals should be uncorrelated, with expectation around 0, constant variance and normally distributed). Therefore, for some geographical areas, we did not select the model with the smallest AICc value as its validation was not good, but a model with an AICc value close to the smallest and a good validation of the model properties based on its residuals. Although there are different methods for order selection in ARMA models, in this paper we adopt the approach described by Brockwell and Davis (2016) [Chapter 5, Section 5.1.3]. The authors there showed that when $p > 0$ and $q > 0$, i.e., under ARMA-like structures, the orders p and q in the empirical autocorrelation function (ACF) and partial autocorrelation function (PACF) are more difficult to identify, and therefore this approach is less valuable compared to cases where either $p = 0$ or $q = 0$. As an alternative, they suggested that in cases where $p > 0$ and $q > 0$, a likelihood-based goodness-of-fit measure such as the AICc for time series of small length should be used. It should also be noted that the monthly time series of both the EP and CP indices were not (weakly) stationary, as they showed trends and seasonal components in many of the geographical areas studied (see also Figures S1 and S2 in the Supplementary Material). In these cases, the evaluation of the empirical ACF and PACF for order selection is neither appropriate nor straightforward, as this practice is only valid when the time series are stationary. At this point, it is also important to highlight that, for a given geographical area, the comparison between ARIMA models with $d = 0$ and ARIMA models with $d = 1$ based on AICc is not an appropriate practice, since the likelihoods on ARIMA models with $d = 0$ and $d = 1$ are not comparable -as the response variables are not the same- and the time series used are not of the same length (if we differentiate once, we lose one observation from the original time series). The comparison can be made within the ARIMA models considered with $d = 0$ and independently within the ARIMA models considered with $d = 1$, taking the best ones in terms of AICc with $d = 0$ and the best ones in terms of AICc with $d = 1$. After examining the properties of these models, such as their predictive ability and complexity, the behaviour of their residuals, the statistical significance of their parameter estimates and the size of the corresponding standard errors of the parameter estimates, etc., the final model selected is the one that offers the best compromise between all these properties. In our case here, we have studied these properties from the best ARIMA models based on AICc from the group with $d = 0$ and $d = 1$ and accordingly selected the best one.

ARIMA models can be used for short-term predictions. In our case here, this is especially interesting since the models proposed in the current work can be used as tools for the early detection of potentially abnormal peaks in the EP and CP indices, anticipating an unexpected change in the temporal dynamics of such indices, and thus providing information on changes in the incidence of diseases related to EP- and CP-like lung lesions in pigs. In the current work, we have considered a 1-step ahead dynamic prediction procedure for the time series of the EP and CP indices for Spain and Catalonia for the period 2020 and 2021. Consider the time series Y_t for $t = 1, \ldots, T$ modelled following an ARIMA structure such as the one described in Eqs. (1) and (2). We want to predict the expected value for the time series at time $T + 1$, i.e., $\widehat{Y}_{T+1}$ (point prediction) based on the estimated ARIMA model, which was estimated with the training time series observed from time for $t = 1$ to time for $t = T$. For example, for a zero-mean stationary AR(1) process, the point prediction of the time series at time $T+1$ (1-step ahead prediction) can be computed as $\widehat{Y}_{T+1} = \widehat{\varphi} Y_T$. In addition, we want to provide this point prediction together with prediction intervals as a measure of error prediction. For more details on the 1-step ahead prediction in MA(q) and ARMA(p,q) processes, as well as the construction of prediction intervals, see Section 3.3 of Chapter 3 of Brockwell and Davis (2016). Once we have the 1-step ahead prediction and the associated prediction intervals for time $T + 1$, we compare the observed value of the time series at time $T+1$ to the 1-step ahead prediction and evaluate whether the observed value of the time series lies inside the corresponding prediction intervals. If so, the observation a time $T+1$ is considered normal under the estimated model; otherwise, the observation is considered potentially abnormal and replaced by the predicted value. After that, the observed values at time $T+1$ is included in the training time series and the estimated ARIMA model (keeping the same structure) is re-estimated with the time series from time $t = 1$ to time $t = T + 1$, and 1-step ahead point prediction and prediction intervals are computed for the next observation of the time series. Again, once the new prediction and prediction intervals are computed, the observation of the corresponding time is compared to the corresponding prediction and the observed value is evaluated if it is potentially abnormal or not. This process is recursively computed for a given period of time, and at each prediction and prediction intervals step the observed value of the time series is compared to the corresponding prediction and it is investigated whether this observation is abnormal or not based on the estimated ARIMA model.

3. Results

3.1. Period 2016–2019: construction of the time series models

The analysis was conducted on a monthly aggregated basis for all evaluated batches at different Spanish geographical locations between 2016 and 2019. In particular, we averaged the recorded EP and CP indices by month and by geographical area in order to construct the monthly time series of the EP and CP indices that were later used to construct the ARIMA models. These geographical locations included the whole country as well as the following sub-regions of Spain: Catalonia, Aragon, Galicia, Norte (including Castilla y León, País Vasco, La Rioja, Navarra and Cantabria), Centro (including Madrid, Castilla-La Mancha and Extremadura), Levante (Comunitat Valenciana and Murcia) and Sur (Andalucía).

A total of 3947 audits obtained from a total of 474 farms associated with 302 different companies in Spain between 2016 and 2019 were included on an aggregated basis (Table S1). Of these, a total of 1140 audits (29.0 %) from 201 farms (42.4 %) belonging to 102 different companies (33.8 %) and a total of 708 audits (18.0 %) from 110 farms (23.3 %) associated with 41 different companies (13.6 %) were from Catalonia and Aragon, the most important Spanish regions in terms of swine production. A total of 604,469 individual lungs from the 3947 audits were evaluated. In addition, the number of pigs evaluated in each audit ranged from 14 to 501, with a mean (standard deviation, SD) of 153.2 (56.1). However, in Catalonia and Aragon, the number of evaluated pigs in each batch ranged from 20 to 406 and 18–465, respectively, with a mean (SD) of 149.5 (45.6) in Catalonia and 164.4 (49.0) in Aragon.

A total of 25.38 % of the pig lungs examined in this study had EP-like lung lesions and 29.97 % had CP-like lesions. Between 2016 and 2019, the mean (SD) EP index among all evaluated batches at the country level was 1.70 (1.58), and the mean (SD) of the monthly time series of EP indices in Spain was 1.93 (0.90). The mean EP index of all evaluated batches ranged from 1.6 in the northeast and east of Spain (Levante [1.64]; Catalonia [1.64] and Aragon [1.64]) to 1.86 in the Norte (1.86) and Centro (1.86) regions (Table S2). The maximum EP index recorded during this study period among all evaluated batches was 10.66 (December 2016) at the Spanish level, while in Catalonia and Aragon it reached 9.34 (December 2016) and 8.36 (August 2016), respectively. By averaging the EP indices by month, the resulting time series for Spain, Catalonia and Aragon (and the rest of the geographical areas) became more reliable and no dramatic abnormal peaks were observed, except for the time series of the EP index in Aragon in April 2016, which was eventually eliminated. In contrast, the mean (SD) CP index among all evaluated batches recorded between 2016 and 2019 was 0.51 (0.43) in Spain and 0.42 (0.37) and 0.43 (0.39) in Catalonia and Aragon, respectively. By region, we observed a mean (SD) CP index among all evaluated batches ranging from 0.29 (0.29) to 0.83 (0.48) in Sur and Galicia, respectively (Table S3). Finally, the maximum CP index among all evaluated batches recorded during the study period was 2.54 (December 2019) in Spain and 2.52 (May 2018) and 2.54 (December 2019) in Catalonia and Aragon, respectively (Table S3). As for the monthly time series of the EP indices, the monthly time series of the CP indices become more reliable and only one abnormal peak was detected for the monthly time series of the CP indices in Catalonia in May 2016, which was also eliminated. For a more detailed description of the evolution of the time series of the EP and CP indices in Spain and its sub-regions, see Figures S1 and S2 in the Supplementary Material.

The estimated ARIMA models for the time series of the EP and CP indices averaged by month between 2016 and 2019 at the country and region levels are shown in Tables 2 and 3. These tables also provides the point estimates and associated standard errors for all estimated parameters of the selected ARIMA models for each analysed sub-region of Spain. As described in Section 2.3, to incorporate a linear trend and seasonal components into the ARIMA models, we first fitted a classical linear model, regressing the corresponding EP or CP index on covariates, including a linear trend and seasonal components with periods of 3, 4, and 12 months. We then selected the components that were statistically significant at the 10 % significance level for inclusion in the ARIMA model. Finally, to determine the values of the ARIMA parameters p, d, and q, we chose the model with a small AICc value while ensuring a good validation of model assumptions and properties.

We will focus here on a detailed description of the models for Spain and Catalonia and Aragon, which are the most important Spanish sub-regions in terms of production (MAPA, 2022). For additional details on the estimated models for the other regions, see Tables 2 and 3 and the Supplementary Material. In addition, for the validation of the selected models, see Figures S4, S5, S6 and S7 in the Supplementary Material. In general, the empirical ACF and the empirical PACF of the residuals of the selected time series models show a behaviour compatible with white noise (see Figures S4 and S5), and the distribution of these residuals does not deviate significantly from the normality assumption, as they all pass the Shapiro-Wilk normality test at a 1 % significance level (see Figures S6 and S7).

First, for the time series of the EP index, there was a clear decreasing trend in the period of study in almost all studied sub-regions of Spain, and we identified in almost all these sub-regions that the temporal distribution of the EP index was autocorrelated. In more detail, at the Spanish level, we can write the estimated equation of the ARIMA model as follows:

$$\widehat{Y}_{S,t} = -0.088t - 0.048\sin\left(\frac{2\pi t}{12}\right) + 0.355\cos\left(\frac{2\pi t}{12}\right) + \widehat{Z}_{S,t},$$

where the process $Y_{S,t}$ is the observation of the EP index at time t in

Table 2
Estimated ARIMA models for the EP index in Spain and different Spanish sub-regions.

Index	Geographical area	Linear trend $\widehat{\beta}_1$ ($\widehat{\sigma}_{\widehat{\beta}_1}$)	Seasonal components: 12-period $\widehat{\beta}_2$ ($\widehat{\sigma}_{\widehat{\beta}_2}$) $\widehat{\beta}_3$ ($\widehat{\sigma}_{\widehat{\beta}_3}$)	4-period $\widehat{\beta}_4$ ($\widehat{\sigma}_{\widehat{\beta}_4}$) $\widehat{\beta}_5$ ($\widehat{\sigma}_{\widehat{\beta}_5}$)	3-period $\widehat{\beta}_6$ ($\widehat{\sigma}_{\widehat{\beta}_6}$) $\widehat{\beta}_7$ ($\widehat{\sigma}_{\widehat{\beta}_7}$)	Estimated ARIMA (p,d,q) parameters		
EP	Spain	−0.088 (0.049)	−0.048 (0.063)			ARIMA (2,1,2)	$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	0.939 (0.143)
			0.355 (0.065)				$\widehat{\varphi}_2$ ($\widehat{\sigma}_{\widehat{\varphi}_2}$)	−0.496 (0.182)
							$\widehat{\vartheta}_1$ ($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	−1.570 (0.095)
							$\widehat{\vartheta}_2$ ($\widehat{\sigma}_{\widehat{\vartheta}_2}$)	1.000 (0.099)
	Aragon	−0.028 (0.003)	−0.451 (0.099)				$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	2.315 (0.065)
			0.007 (0.099)			ARIMA (1,0,1)	$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	0.648 (0.134)
							$\widehat{\vartheta}_1$ ($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	−1.000 (0.093)
	Catalonia	−0.064 (0.038)				ARIMA (1,1,2)	$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	−0.709 (0.126)
							$\widehat{\vartheta}_1$ ($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	0.413 (0.193)
							$\widehat{\vartheta}_2$ ($\widehat{\sigma}_{\widehat{\vartheta}_2}$)	−0.587 (0.150)
	Centro					ARIMA (3,0,2)	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	2.189 (0.325)
							$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	0.744 (0.166)
							$\widehat{\varphi}_2$ ($\widehat{\sigma}_{\widehat{\varphi}_2}$)	−1.071 (0.073)
							$\widehat{\varphi}_3$ ($\widehat{\sigma}_{\widehat{\varphi}_3}$)	0.383 (0.148)
							$\widehat{\vartheta}_1$ ($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	−0.565 (0.122)
							$\widehat{\vartheta}_2$ ($\widehat{\sigma}_{\widehat{\vartheta}_2}$)	1.000 (0.117)
	Galicia					ARIMA (3,0,1)	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	1.261 (0.073)
							$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	0.708 (0.191)
							$\widehat{\varphi}_2$ ($\widehat{\sigma}_{\widehat{\varphi}_2}$)	0.207 (0.266)
							$\widehat{\varphi}_3$ ($\widehat{\sigma}_{\widehat{\varphi}_3}$)	−0.503 (0.198)
							$\widehat{\vartheta}_1$ ($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	−1.000 (0.182)
	Levante	−0.026 (0.015)				ARIMA (1,0,0)	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	2.610 (0.437)
							$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	0.336 (0.167)
	Norte	−0.021 (0.008)	0.308 (0.138)			ARIMA (0,0,0)	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	2.153 (0.223)
			0.296 (0.136)					
	Sur					ARIMA (2,1,2)	$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	−1.198 (0.037)
							$\widehat{\varphi}_2$ ($\widehat{\sigma}_{\widehat{\varphi}_2}$)	−1.198 (0.037)
							$\widehat{\vartheta}_1$ ($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	1.246 (0.192)
							$\widehat{\vartheta}_2$ ($\widehat{\sigma}_{\widehat{\vartheta}_2}$)	0.994 (0.222)

Table 3
Estimated ARIMA models for the CP index in Spain and different Spanish sub-regions.

Index	Geographical area	Linear trend $\widehat{\beta}_1$ ($\widehat{\sigma}_{\widehat{\beta}_1}$)	Seasonal components: 12-period $\widehat{\beta}_2$ ($\widehat{\sigma}_{\widehat{\beta}_2}$) $\widehat{\beta}_3$ ($\widehat{\sigma}_{\widehat{\beta}_3}$)	4-period $\widehat{\beta}_4$ ($\widehat{\sigma}_{\widehat{\beta}_4}$) $\widehat{\beta}_5$ ($\widehat{\sigma}_{\widehat{\beta}_5}$)	3-period $\widehat{\beta}_6$ ($\widehat{\sigma}_{\widehat{\beta}_6}$) $\widehat{\beta}_7$ ($\widehat{\sigma}_{\widehat{\beta}_7}$)	Estimated ARIMA (p,d,q) parameters		
CP	**Spain**	0.004 (0.002)				**ARIMA (2,0,0)**	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	0.352 (0.060)
							$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	0.212 (0.152)
							$\widehat{\varphi}_2$ ($\widehat{\sigma}_{\widehat{\varphi}_2}$)	0.273 (0.162)
	Aragon	0.007 (0.001)				**ARIMA (2,0,3)**	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	0.246 (0.014)
							$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	−0.550 (0.154)
							$\widehat{\varphi}_2$ ($\widehat{\sigma}_{\widehat{\varphi}_2}$)	0.329 (0.165)
							$\widehat{\vartheta}_1$($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	0.814 (0.160)
							$\widehat{\vartheta}_2$($\widehat{\sigma}_{\widehat{\vartheta}_2}$)	−0.814 (0.129)
							$\widehat{\vartheta}_3$($\widehat{\sigma}_{\widehat{\vartheta}_3}$)	−0.999 (0.170)
	Catalonia					**ARIMA (3,0,1)**	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	0.425 (0.010)
							$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	1.157 (0.139)
							$\widehat{\varphi}_2$ ($\widehat{\sigma}_{\widehat{\varphi}_2}$)	0.218 (0.244)
							$\widehat{\varphi}_3$ ($\widehat{\sigma}_{\widehat{\varphi}_3}$)	−0.471 (0.149)
							$\widehat{\vartheta}_1$($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	−1.000 (0.075)
	Centro				−0.137 (0.047) −0.010 (0.048)	**ARIMA (0,0,1)**	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	0.515 (0.051)
							$\widehat{\vartheta}_1$($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	0.363 (0.128)
	Galicia	0.014 (0.007)	0.020 (0.077) 0.154 (0.074)			**ARIMA (4,0,2)**	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	0.057 (0.207)
							$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	−0.931 (0.339)
							$\widehat{\varphi}_2$ ($\widehat{\sigma}_{\widehat{\varphi}_2}$)	−0.608 (0.323)
							$\widehat{\varphi}_3$ ($\widehat{\sigma}_{\widehat{\varphi}_3}$)	0.102 (0.347)
							$\widehat{\varphi}_4$ ($\widehat{\sigma}_{\widehat{\varphi}_4}$)	0.111 (0.372)
							$\widehat{\vartheta}_1$($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	1.440 (0.392)
							$\widehat{\vartheta}_2$($\widehat{\sigma}_{\widehat{\vartheta}_2}$)	1.000 (0.431)
	Levante	0.005 (0.001)	0.089 (0.017) −0.085 (0.018)	0.019 (0.014) −0.054 (0.016)		**ARIMA (4,0,2)**	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	0.253 (0.011)
							$\widehat{\varphi}_1$ ($\widehat{\sigma}_{\widehat{\varphi}_1}$)	0.588 (0.190)
							$\widehat{\varphi}_2$ ($\widehat{\sigma}_{\widehat{\varphi}_2}$)	0.674 (0.407)
							$\widehat{\varphi}_3$ ($\widehat{\sigma}_{\widehat{\varphi}_3}$)	0.105 (0.185)
							$\widehat{\varphi}_4$ ($\widehat{\sigma}_{\widehat{\varphi}_4}$)	−0.702 (0.154)
							$\widehat{\vartheta}_1$($\widehat{\sigma}_{\widehat{\vartheta}_1}$)	−1.556 (0.367)
							$\widehat{\vartheta}_2$($\widehat{\sigma}_{\widehat{\vartheta}_2}$)	0.556 (0.324)
	Norte	0.003 (0.002)	0.093 (0.033) 0.022 (0.033)	0.027 (0.033) −0.086 (0.032)		**ARIMA (0,0,0)**	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	0.419 (0.053)
	Sur		0.011 (0.032) −0.098 (0.032)		0.064 (0.031) 0.020 (0.033)	**ARIMA (0,0,0)**	$\widehat{\mu}$ ($\widehat{\sigma}_{\widehat{\mu}}$)	0.266 (0.023)

Spain and the process $X_{S,t}$ was once differentiated, resulting in the process $Z_{S,t} = (1 - L)X_{S,t} = X_{S,t} - X_{S,t-1}$ with L as the lag operator. The differentiated process $Z_{S,t}$ is estimated with the following equation:

$$\widehat{Z}_{S,t} = 0.939Z_{S,t-1} - 0.496Z_{S,t-2} + \varepsilon_{S,t} - 1.570\varepsilon_{S,t-1} + 1.000\varepsilon_{S,t-2},$$

which corresponds to the equation of a stationary ARMA process with $p = 2$ and $q = 2$, i.e., ARMA (2,2). Figure S8 in the Supplementary Material shows how the estimated linear trend and the annual seasonality for the EP index in Spain behave over the period of study. This figure clearly shows a decreasing linear trend, but it also shows an annual periodic behaviour with an increase in the mean EP index over in fall-winter and a decrease in the mean EP index over in spring-summer, approximately.

Focusing on Aragon and Catalonia, as these are the two most important geographical areas in Spain in terms of swine production, we still observed a decreasing linear trend as well as a 12-period seasonal component (Aragon) of the monthly time series of the EP index between 2016 and 2019. As for the EP index in Spain, Figure S8 in the Supplementary Material again shows how the estimated linear trend and the annual seasonality for the EP index in Aragon over the study period behave similarly to the EP index in Spain. The monthly time series of the EP index in Aragon also showed an autocorrelation structure, as in the case of the time series of the EP index in Spain. Finally, the time series of the EP index in Catalonia also showed a decreasing linear trend and an autocorrelation structure. For more details on the models for these time series in Spain and in each of the geographical areas analysed, see Table 2. In addition, for the validation of the models, refer to Figures S4 and S6 of the Supplementary Material.

On the other hand, the monthly time series of the CP index generally showed an increasing trend over the period of study, and particularly at the country level, this index slightly linearly increased between 2016 and 2019. After taking this trend component into account, the remaining process showed an autocorrelated structure over time, and more precisely, we modelled such a process with an AR(2) structure. More particularly, consider the process $Y'_{S,t}$ as the CP index at time t in Spain, which was modelled as follows: $\widehat{Y}'_{S,t} = 0.004t + \widehat{Z}'_{S,t}$, where $Z'_{S,t}$ is modelled with the following stationary AR(2) process: $\widehat{Z}'_{S,t} = 0.352 + 0.212Z'_{S,t-1} + 0.273Z'_{S,t-2} + u_{S,t}$. That is, the time series of the CP index in Spain increased slightly from 2016 to 2019, and after considering this trend component, the process at time t was linearly associated with itself at times $t-1$ and $t-2$, i.e., the observed CP index in a given month over the period of study is partially explained by what happened to such an index in the immediately preceding two months. In addition, while the time series of the CP index for Aragon still showed an increasing linear trend, in line with what was observed for the time series of the CP index for Spain, this trend was not observed for Catalonia. However, the time series of the CP index showed a clear second and third order autocorrelation structure for Aragon and Catalonia respectively. For the remaining sub-regions of Spain, Table 3 also provides detailed information on each selected model for the time series of the CP index and Figures S5 and S7 in the Supplementary Material show the results of the validation of these models.

Looking at Tables 2 and 3, it is easy to see that there were common characteristics in the estimated models for the time series of the EP and CP indices in most of the geographical areas. For example, for the time series of the EP index, in general, a linear decreasing trend was observed over the period studied in almost all the geographical areas of Spain (except for Centro, Galicia and Sur), and for some of these areas (Spain, Aragon and Norte) an annual seasonal component is also detected. Regarding the autocorrelation structure of the time series of the EP index for each geographical area, it is generally observed that these time series very often show a second or even third order autocorrelation structure, accompanied by a complex structure for the random error process (i.e., the error process itself is correlated with a generally first and second order temporal dependence). On the other hand, the time series of the CP index shows a generally increasing linear trend over the period of study in most of the geographical areas studied (except for Catalonia, Centro and Sur), together with an annual seasonal component (except for Spain, Aragon, Catalonia and Centro). Four- and three-period seasonal components are also observed in the time series of the CP index for some geographical areas, but these are not very frequent. Finally, the time series of the CP index generally shows a second order autocorrelation structure, as well as an autocorrelation structure for the random error process.

Some of the differences in the model structures between the different geographical areas when modelling either the time series of the EP index or the CP index can be linked to the differences in management and logistics in each of these areas, as well as to differences in climate traits, as there are areas with a typical Mediterranean climate, while others have a more typical continental climate.

3.2. *Period 2020–2021: dynamic predictions with estimated time series models*

After modelling both the time series of the EP and CP indices at the Spanish level and in different sub-regions of Spain to describe their temporal behaviour and understand their evolution in terms of trends, seasonal components, and autocorrelation structures during the study period, we used the selected time series models (Tables 2 and 3) to evaluate their predictive capability. We also wanted the models to be used as alarm tools for the early detection of potentially abnormal changes in the evolution of the studied indices over time. Our focus was primarily on Spain and Catalonia, mainly because these are the most reliable time series. While Catalonia is one of the most important Spanish regions in terms of pig production, Spain aggregates information of EP-like and CP-like lesions from different subregions of the country, giving a picture of the situation of EP and CP related pathologies at the country level.

Following the description of the prediction procedure conducted in the current work (see Section 2.3), to obtain monthly dynamic predictions, the estimated ARIMA models for Spain and Catalonia for both the time series of the EP and CP indices was used. We forecasted the expected EP and CP indices for Spain and Catalonia for January 2020. These point predictions were accompanied by their corresponding 90 % prediction intervals. More precisely, Table S4 shows that the predicted EP index value and the associated 90 % prediction intervals for Spain and Catalonia for January 2020 were 1.434 (0.699, 2.168) and 1.189 (0.319, 2.058), respectively, while the predicted CP index value and their 90 % prediction intervals for Spain and Catalonia for January 2020 were 0.482 (0.300, 0.664) and 0.369 (0.152, 0.585), respectively. When data from both Spain and Catalonia were available for January 2020, the observed EP and CP indices were compared to the estimated values and evaluated whether the observed indices fell within the corresponding 90 % prediction intervals. It was considered a normal value if the observed index fell within the respective 90 % prediction intervals. Otherwise, it was considered an abnormal observation and reported as such. In the example above, while we observed in January 2020 a value for the EP index of 1.199 for Spain and a value of 1.434 for Catalonia, for the CP index we observed in January 2020 values of 0.525 and 0.573 for Spain and Catalonia, respectively. In all the cases before, except for the CP value for Catalonia, the observed values are within the corresponding 90 % prediction intervals.

After this first prediction for the EP and CP indices for January 2020, for the EP index in Spain, the selected ARIMA (2,1,2) was re-estimated but over the monthly observations from January 2016 to January 2020, i.e., the observation corresponding to January 2020 is incorporated in the training time series. Similarly, we used the data from the same period to re-estimate the selected ARIMA (2,0,0) for the CP index in Spain. Related to Catalonia, we re-estimated the selected ARIMA (1,1,2) for the EP index and the selected ARIMA (3,0,1) for the CP index using the data collected between January 2016 and January 2020. The re-estimated models for both the EP and CP indices for Spain and Catalonia were used to predict again the expected values of such indices for February 2020 as well as the associated 90 % prediction intervals, as previously described.

Once the information on the observed EP and CP indices for February 2020 was available, these observed values were compared to the predicted ones and were classified as normal or abnormal values with respect to the 90 % prediction intervals constructed over the re-estimated ARIMA models. This process was repeated until September 2021. Predicted values (green) and the associated 90 % prediction intervals (red) for EP and CP indices in Spain and Catalonia from January 2020 to September 2021 are shown in Fig. 1. It also depicts the evolution of the observed EP and CP indices from January 2016 to September 2021, showing whether the observed values between January 2020 and September 2021 fall within the estimated 90 % prediction intervals (Fig. 1). Additionally, the observed values from January 2020 to September 2021 for both the EP and CP indices, respectively, as well as the predicted values and the respective 90 % prediction intervals, are detailed in Tables S4 and S5 in the Supplementary Material. Note here that we have considered 90 % prediction intervals rather than 95 % prediction intervals, as our time series lengths are not particularly long, and the precision of the predictions is not as ideal. For prediction methods in ARMA models, please see Brockwell and Davis (2016).

4. Discussion

In this study, we described the prevalence and dynamics of EP- and CP-like lesions at slaughterhouses in different Spanish regions between 2016 and 2019, as well as a system to predict the prevalence of such lung lesions in pigs at slaughter age in 2020 and 2021.

The presence of EP- and CP-like lung lesions are typically associated with *Mhyo* and *App* infections, respectively (Arruda et al., 2024). High prevalence of these lung lesions has been reported worldwide, except for Switzerland, Finland and Norway, where eradication programmes were performed (Maes et al., 2023). Specifically, previous studies carried out in Spain reported that the percentages of pigs showing infections related to EP ranged from 31 % to 56 % and those related to CP ranged from 19 % to 37 % (Fraile et al., 2010; Pallarés et al., 2021). This study indicated lower prevalence for EP-like lung lesions (25.38 %) in Spain over the years, while lesions related to CP (29.97 %) seems to be within the range. Thus, these results confirm that these lesions are still widespread in Spanish herds. Interestingly, a statistically significant decreasing trend in the monthly time series of the EP index was observed in both Spain and Catalonia from 2016 to 2019 (Fig. 1). This significant decrease could be associated with the implementation of different measures within their control programs, such as improvements in management and biosecurity practices and/or differences in the vaccines used against *Mhyo*.

Vaccination is the most used management practice to control *Mhyo* infections (Maes et al., 2023). In this study, the EP index observed in Spain at time t was inversely linearly related to the proportion of batches vaccinated with Hyogen® at time $t - 1$, with a statistically significant

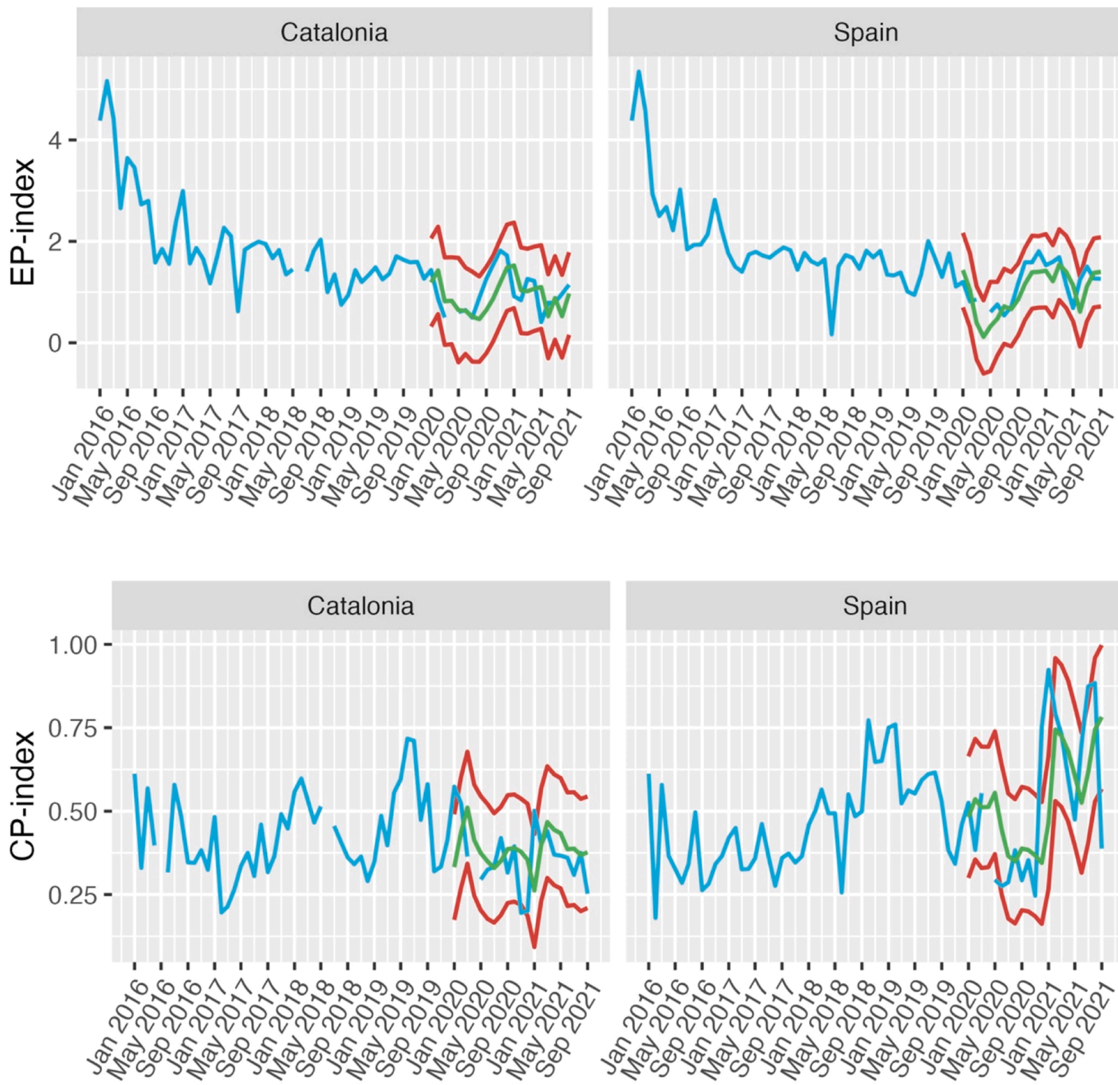

Fig. 1. Evolution of the observed (blue), predicted (green) and 90 % prediction intervals (red) of the EP and CP indices from 2016 to 2021.

difference of 5 % (Figure S3). Thus, an increase in the proportion of vaccinated batches seemed to impact the immediate future evolution of the EP index at slaughterhouses and the prevalence of infections related to *Mhyo*.

In contrast, our results showed that the prevalence of lesions related to CP increased during the same study period, especially at the national level. This finding could be related to different factors previously reported in different studies. First, the intensification and type of pig production and the decrease in the use of antibiotics may increase the occurrence of lesions related to CP at the slaughterhouse (Enøe et al., 2002; Maes et al., 2023). Moreover, management practice of increasing the fattening time of pigs to reach the slaughter weight could also contribute to increase the risk of pleuritis (Hartley et al., 1988; Jirawattanapong et al., 2010). Additionally, the presence of diverse serotypes of *App* with different virulence and/or seroconversion to other primary pathogens (e.g., porcine reproductive and respiratory syndrome virus and swine influenza virus) has also been identified as a risk factor for pleuritis occurrence (Dottori et al., 2007; Fraile et al., 2010; Merialdi et al., 2012). These findings, together with the fact that infections related to CP are among the main causes of condemnation of carcasses in slaughterhouses (De Conti et al., 2021; Szeredi et al., 2015), highlight the importance of lung lesion assessment at slaughterhouses as part of control monitoring programs. Finally, CVPC and pleurisy lesions are not pathognomonic to *Mhyo* and *App* infections, respectively. In fact, pleuritis is recognized as a multifactorial lesion associated with *Pasteurella multocida* type *A*, a secondary agent of EP caused by *Mhyo* infection (Park et al., 2016; Petri et al., 2023), and/or with other agents causing polyserositis such as *Glaesserella parasuis* or *Mycoplasma hyorinis* in case of weaned piglets Salogni et al., 2020). Therefore, an additional diagnostic analysis is required to confirm the pathogen(s) involved in both lung lesions.

Like many other respiratory diseases, environmental factors could play an important role in the transmission and severity of lung lesions (Fraile et al., 2010; Meyns et al., 2011; Merialdi et al., 2012). In this sense, a seasonal effect was observed for infections related to EP and CP in Catalonia and Aragon. Although most swine production is based on confinement systems, the impact of environmental conditions, such as temperature and humidity, in fattening units with natural ventilation could be especially significant since pigs are exposed to different external conditions depending on the season.

The time series models considered in this study were constructed based on data collected at different slaughterhouses in Spain during a period of 4 years, from 2016 to 2019. Although the coverage of the target swine population varied between regions, north-eastern Spain (Aragon and Catalonia) was the most represented since it covered more than half of the Spanish pig census (MAPA, 2022). One of the original parts of this study was the evaluation of the predictive performance of the estimated model at a quasi-real time with data from 2020 to 2021.

The predictive values obtained for the EP and CP indices based on the estimated ARIMA models did not significantly differ from the observed values collected at the slaughterhouse and were within the corresponding 90 % prediction intervals. Therefore, these results showed that the ARIMA model constructed in this study has high performance in predicting EP- and CP-like lesions in pigs slaughtered at the national level and in the considered sub-regions of Spain.

The estimated time series models provide a strong starting point for future research on infection dynamics and outbreak forecasting. The prediction of lung lesion prevalence using this powerful tool can facilitate the early detection of potential respiratory outbreaks on a farm by veterinarians. Likewise, the management practices focused to minimize the pathogen spread (i.e. pig density reduction, age of weaning) and/or the application of preventive strategies as vaccination against these pathogens could be implemented in advance, minimizing the impact of the disease/outbreak. Thus, this predictive approach using time series models can have a significant impact reducing economic losses associated to respiratory diseases and outbreaks in swine industry. Notwithstanding, further investigations into the application of automatic image analysis methods for lesion evaluation as well as the artificial intelligence-based systems could be explored to enhance the predictive efficiency and accuracy of these models.

5. Conclusions

In conclusion, the study of the evolution of EP- and CP-like lesions at slaughterhouses using time series models is a useful tool to be routinely implemented as part of control programmes for *Mhyo* and *App*. First, the temporal evolution of infections associated with EP and CP in Spain was described, and second, the evolution of such infections over time was predicted. Likewise, the proposed models constructed in the current study could be used to early identify potential critical points such as respiratory outbreaks within herds and thus design appropriate intervention strategies to minimize their impact.

Funding

This work was supported by Ceva Salud Animal.

CRediT authorship contribution statement

Sibila Marina: Writing – review & editing, Supervision, Conceptualization. **Carmona Marta:** Writing – review & editing, Supervision, Conceptualization. **Lasierra Maria Teresa:** Writing – review & editing, Supervision, Data curation, Conceptualization. **Fernandez-Fontelo Amanda:** Writing – review & editing, Writing – original draft, Validation, Software, Formal analysis, Data curation, Conceptualization. **Garza-Moreno Laura:** Writing – review & editing, Validation, Supervision, Project administration, Formal analysis, Conceptualization.

Declaration of Competing Interest

The authors declare no conflict of interest. MTL, MC and LGM are employees of Ceva Salud Animal.

Acknowledgement

The authors are extremely grateful to the slaughterhouses and integrators involved in this study

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.prevetmed.2025.106525.

Data availability

All data generated and analysed in this work are included in this published article and its supplementary information files.

References

Alba, A., Dórea, F.C., Arinero, L., Sanchez, J., Cordón, R., Puig, P., Revie, C.W., 2015. Exploring the surveillance potential of mortality data: nine years of bovine fallen stock data collected in Catalonia (Spain). PLOS ONE 10 (4), e0122547. https://doi.org/10.1371/journal.pone.0122547.

Alba-Casals, A., Allue, E., Tarancon, V., Baliellas, J., Novell, E., Napp, S., Fraile, L., 2020. Near Real-Time Monitoring of Clinical Events Detected in Swine Herds in Northeastern Spain. Front. Vet. Sci. 7, 68. https://doi.org/10.3389/fvets.2020.00068.

Arruda, L.P., Malcher, C., Petri, F.A.M., da Silva, D.G., Storino, G.Y., Almeida, H.M., de, S., Sonalio, K., Toledo, L.T., de Oliveira, L.G., 2024. Pathological analysis and etiological assessment of pulmonary lesions and its association with pleurisy in slaughtered pigs. Vet. Microbiol. 292, 110039. https://doi.org/10.1016/j.vetmic.2024.110039.

Barongo, M.B., Bishop, R.P., Fèvre, E.M., Knobel, D.L., Ssematimba, A., 2016. A Mathematical model that simulates control options for african swine fever virus (ASFV). PLOS ONE 11 (7), e0158658. https://doi.org/10.1371/journal.pone.0158658.

Benschop, J., Stevenson, M.A., Dahl, J., Morris, R.S., French, N.P., 2008. Temporal and longitudinal analysis of Danish Swine Salmonellosis Control Programme data: implications for surveillance. Epidemiol. Infect. 136 (11), 1511–1520. https://doi.org/10.1017/S0950268807000234.

Brockwell, P.J., Davis, R.A., 2016. Introduction to Time Series and Forecasting. Springer International Publishing. https://doi.org/10.1007/978-3-319-29854-2.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-theoretic Approach, second ed. Springer.

Christensen, G., Sørensen, V., Mousing, J., 1999. Diseases of the respiratory system. In: Straw, En.B.E., D'Allaire, S., Mengeling, W.L., Taylor, D.J. (Eds.), Diseases of Swine. Iowa State University Press, pp. 913–940.

Cryer, J.D., Chan, K.-S., 2008. Time Series Analysis. Springer. https://doi.org/10.1007/978-0-387-75959-3.

De Conti, E.R., Resende, T.P., Marshall-Lund, L., Rovira, A., Vannucci, F.A., 2021. Correction: De Conti et al. Histological lesions and replication sites of PCV3 in naturally infected pigs. Animals 2021, 11, 1520. Anim. Open Access J. MDPI 11 (11), 3277. https://doi.org/10.3390/ani11113277.

Dottori, M., Nigrelli, A.D., Bonilauri, P., Merialdi, G., Gozio, S., Cominotti, F., 2007. Proposta per un nuovo sistema di punteggiatura delle pleuriti suine in sede di macellazione. La griglia SPES (Slaughterhouse Pleurisy Evaluation System). Large Anim. Rev. 13, 161–165.

Enøe, C., Mousing, J., Schirmer, A.L., Willeberg, P., 2002. Infectious and rearing-system related risk factors for chronic pleuritis in slaughter pigs. Prev. Vet. Med. 54 (4), 337–349. https://doi.org/10.1016/S0167-5877(02)00029-6.

Evans, C.M., Medley, G.F., Creasey, S.J., Green, L.E., 2010. A stochastic mathematical model of the within-herd transmission dynamics of porcine reproductive and respiratory syndrome virus (PRRSV): Fade-out and persistence. Prev. Vet. Med. 93 (4), 248–257. https://doi.org/10.1016/j.prevetmed.2009.11.001.

Eze, J.I., Correia-Gomes, C., Borobia-Belsué, J., Tucker, A.W., Sparrow, D., Strachan, D. W., Gunn, G.J., 2015. Comparison of respiratory disease prevalence among voluntary monitoring systems for pig health and welfare in the UK. PLOS ONE 10 (5), e0128137. https://doi.org/10.1371/journal.pone.0128137.

Fano, E., Pijoan, C., Dee, S., Deen, J., 2007. Effect of Mycoplasma hyopneumoniae colonization at weaning on disease severity in growing pigs. Can. J. Vet. Res. 71 (3), 195–200.

Fernández-Fontelo, A., Puig, P., Caceres, G., Romero, L., Revie, C., Sanchez, J., Dorea, F. C., Alba-Casals, A., 2020. Enhancing the monitoring of fallen stock at different hierarchical administrative levels: an illustration on dairy cattle from regions with distinct husbandry, demographical and climate traits. BMC Vet. Res. 16 (1), 110. https://doi.org/10.1186/s12917-020-02312-8.

Fraile, L., Alegre, A., López-Jiménez, R., Nofrarías, M., Segalés, J., 2010. Risk factors associated with pleuritis and cranio-ventral pulmonary consolidation in slaughter-aged pigs. Vet. J. 184 (3), 326–333. https://doi.org/10.1016/j.tvjl.2009.03.029.

Garcia-Morante, B., Segalés, J., Fraile, L., Pérez de Rozas, A., Maiti, H., Coll, T., Sibila, M., 2016. Assessment of Mycoplasma hyopneumoniae-induced pneumonia using different lung lesion scoring systems: a comparative review. J. Comp. Pathol. 154 (2), 125–134. https://doi.org/10.1016/j.jcpa.2015.11.003.

Hartley, P.E., Wilesmith, J.W., Bradley, R., 1988. Prevalence of pleurisy in pigs at slaughter. Vet. Rec. 123 (7), 173–175. https://doi.org/10.1136/vr.123.7.173.

Ivanek, R., Snary, E.L., Cook, A.J.C., Gröhn, Y.T., 2004. A mathematical model for the transmission of *Salmonella typhimurium* within a grower-finisher pig herd in great Britain. J. Food Prot. 67 (11), 2403–2409. https://doi.org/10.4315/0362-028X-67.11.2403.

Jirawattanapong, P., Stockhofe-Zurwieden, N., van Leengoed, L., Wisselink, H., Raymakers, R., Cruijsen, T., van der Peet-Schwering, C., Nielen, M., van Nes, A., 2010. Pleuritis in slaughter pigs: relations between lung lesions and bacteriology in 10 herds with high pleuritis. Res. Vet. Sci. 88 (1), 11–15. https://doi.org/10.1016/j.rvsc.2009.06.007.

Keeling, M.J., Danon, L., 2009. Mathematical modelling of infectious diseases. Br. Med. Bull. 92 (1), 33–42. https://doi.org/10.1093/bmb/ldp038.

Krejci R., Munoz R. Proceedings of the 23 rd International Pig Veterinary Society Congress, Cancun, Mexico. 2014. Lung scoring program as a monitoring tool in Asia.
Madec, F., Kobisch, M., 1982. Bilan lésionnel des poumons de porcs charcutiers. Journées De. la Rech. Porc. 14 (8).
Maes, D., Segales, J., Meyns, T., Sibila, M., Pieters, M., Haesebrouck, F., 2008. Control of Mycoplasma hyopneumoniae infections in pigs. Vet. Microbiol. 126 (4), 297–309. https://doi.org/10.1016/j.vetmic.2007.09.008.
Maes, D., Sibila, M., Pieters, M., Haesebrouck, F., Segalés, J., de Oliveira, L.G., 2023. Review on the methodology to assess respiratory tract lesions in pigs and their production impact. Vet. Res. 54 (1), 8. https://doi.org/10.1186/s13567-023-01136-2.
Merialdi, G., Boniliari, P., Dottori, M., Nigrelli, A., Martelli, P., 2008. Monitoring Respiratory Disease at Slaughterhouse Using Lung and Pleural Lesions Score and Serology. 20th International Pig Veterinary Society.
Merialdi, G., Dottori, M., Bonilauri, P., Luppi, A., Gozio, S., Pozzi, P., Spaggiari, B., Martelli, P., 2012. Survey of pleuritis and pulmonary lesions in pigs at abattoir with a focus on the extent of the condition and herd risk factors. Vet. J. (Lond. Engl. 1997) 193 (1), 234–239. https://doi.org/10.1016/j.tvjl.2011.11.009.
Meyns, T., Van Steelant, J., Rolly, E., Dewulf, J., Haesebrouck, F., Maes, D., 2011. A cross-sectional study of risk factors associated with pulmonary lesions in pigs at slaughter. Vet. J. 187 (3), 388–392. https://doi.org/10.1016/j.tvjl.2009.12.027.
, 2014Ministerio de Agricultura, Pesca y Alimentación (MAPA), Gobierno de España. “El sector de la carne de cerdo en cifras: Principales indicadores económicos”, Subdirección General de Producciones Ganaderas y Cinegéticas, Dirección General de Producciones y Mercados Agrarios. 2022. [(accessed on 15th November 2023)].
Montgomery, D.C., Jennings, C.L., Kulahci, M., 2011. Introduction to Time Series Analysis and Forecasting. John Wiley & Sons.
Nathues, H., Alarcon, P., Rushton, J., Jolie, R., Fiebig, K., Jimenez, M., Geurts, V., Nathues, C., 2017. Cost of porcine reproductive and respiratory syndrome virus at individual farm level—an economic disease model. Prev. Vet. Med. 142, 16–29. https://doi.org/10.1016/j.prevetmed.2017.04.006.
Neumann, E., Hall, W., Stevenson, M., Morris, R., Ling Min Than, J., 2014. Descriptive and temporal analysis of post-mortem lesions recorded in slaughtered pigs in New Zealand from 2000 to 2010. N Z Vet J. 62 (3), 110–116. https://doi.org/10.1080/00480169.2013.853278. Epub 2013 Dec 17.
Ostanello, F., Dottori, M., Gusmara, C., Leotti, G., Sala, V., 2007. Pneumonia disease assessment using a slaughterhouse lung-scoring method. J. Vet. Med. Ser. A 54 (2), 70–75. https://doi.org/10.1111/j.1439-0442.2007.00920.x.
Pallarés, F., Añón, J., Rodríguez-Gómez, I., Gómez-Laguna, J., Fabré, R., Sánchez-Carvajal, J., Ruedas-Torres, I., Carrasco, L., 2021. Prevalence of mycoplasma-like lung lesions in pigs from commercial farms from Spain and Portugal. Porc. Health Manag. 7 (1), 26. https://doi.org/10.1186/s40813-021-00204-3.
Park, C., Jeong, J., Kang, I., Choi, K., Park, S.-J., Chae, C., 2016. Increased fucosyl glycoconjugate by Mycoplasma hyopneumoniae enhances adherences of Pasteurella multocida type A in the ciliated epithelial cells of the respiratory tract. BMC Vet. Res. 12, 25. https://doi.org/10.1186/s12917-016-0650-7.
Paz-Sánchez, Y., Herráez, P., Quesada-Canales, Ó., Poveda, C.G., Díaz-Delgado, J., Quintana-Montesdeoca, M. del P., Plamenova Stefanova, E., Andrada, M., 2021. Assessment of lung disease in finishing pigs at slaughter: pulmonary lesions and implications on productivity parameters. Article 12. Animals 11 (12). https://doi.org/10.3390/ani11123604.
Peña, D., Tiao, G.C., Tsay, R.S., 2001. A Course in Time Series Analysis. Wiley.
Petri, F.A.M., Ferreira, G.C., Arruda, L.P., Malcher, C.S., Storino, G.Y., Almeida, H.M., de, S., Sonalio, K., Silva, D.G. da, Oliveira, L.G. de, 2023. Associations between pleurisy and the main bacterial pathogens of the porcine respiratory diseases complex (PRDC. Anim. Open Access J. MDPI 13 (9), 1493. https://doi.org/10.3390/ani13091493.
Salogni, C., Lazzaro, M., Giovannini, S., Vitale, N., Boniotti, M.B., Pozzi, P., Pasquali, P., Alborali, G.L., 2020. Causes of swine polyserositis in a high-density breeding area in Italy. J. Vet. Diagn. Invest. 32 (4), 594–597. https://doi.org/10.1177/1040638720928973.
Shumway, R.H., Stoffer, D.S., 2017. Time Series Analysis and Its Applications: With R Examples. Springer International Publishing. https://doi.org/10.1007/978-3-319-52452-8.
Sibila, M., Pieters, M., Molitor, T., Maes, D., Haesebrouck, F., Segalés, J., 2009. Current perspectives on the diagnosis and epidemiology of Mycoplasma hyopneumoniae infection. Vet. J. 181 (3), 221–231. https://doi.org/10.1016/j.tvjl.2008.02.020.
Sorensen, V., Jorsal, S.E., Mousing, J., 2006. Diseases of the respiratory system. In: Straw, B., Zimmermann, W., D'Allaire, S., Taylor, D.J. (Eds.), Diseases of Swine, ninth ed. Iowa State University Press, Ames, Iowa.
Szeredi, L., Cságola, A., Dán, Á., Dencső, L., 2015. Vascular lesions and pneumonia in a pig fetus infected by porcine circovirus type 2. Acta Vet. Hung. 63 (2), 215–222. https://doi.org/10.1556/004.2015.019.
Ward, M.P., Iglesias, R.M., Brookes, V.J., 2020. Autoregressive Models Applied to Time-Series Data in Veterinary Science. Front. Vet. Sci. 7, 604. https://doi.org/10.3389/fvets.2020.00604. PMID: 33094106; PMCID: PMC7527444.
Yaeger, M.J., Alstine, W.G., 2019. Respiratory system. In: Zimmerman, J.J., Karriker, L. A., Ramirez, A., Schwartz, K.J., Stevenson, G.W., Zhang, J. (Eds.), Diseases of swine, eleventh edition. Wiley-Blackwell.