Ann Oper Res (2014) 219:25–47
DOI 10.1007/s10479-012-1257-4

Mathematical formulation and validation of the Be-FAST model for Classical Swine Fever Virus spread between and within farms

Benjamin Ivorra · Beatriz Martínez-López · José M. Sánchez-Vizcaíno · Ángel M. Ramos

Published online: 23 November 2012

Abstract Classical Swine Fever is a viral disease of pigs that causes severe restrictions on the movement of pigs and pig products in the affected areas. The knowledge of its spread patterns and risk factors would help to implement specific measures for controlling future outbreaks. In this article, we describe in detail a spatial hybrid model, called Be-FAST, based on the combination of a stochastic Individual-Based model (modeling the interactions between the farms, considered as individuals) for between-farm spread with a Susceptible-Infected model for within-farm spread, to simulate the spread of this disease and identify risk zones in a given region. First, we focus on the mathematical formulation of each component of the model. Then, in order to validate Be-FAST, we perform various numerical experiments considering the Spanish province of Segovia. Obtained results are compared with the ones given by two other Individual-Based models and real outbreaks data from Segovia and The Netherlands.

Keywords Epidemiological modeling · Individual-Based model · Susceptible-Infected model · Model validation · Classical Swine Fever

1 Introduction

Modeling and simulation are important tools to fight diseases (Anderson and May 1979). Each disease has its own characteristics and, therefore, most of them need a well-adapted simulation model in order to tackle real-life situations (Brauer and Castillo-Chavez 2001).

B. Ivorra (✉) · Á.M. Ramos
Departamento de Matemática Aplicada, Fac. de C.C. Matemáticas & Instituto de Matemática Interdisciplinar, Universidad Complutense de Madrid, Madrid, Spain
e-mail: ivorra@mat.ucm.es

B. Martínez-López · J.M. Sánchez-Vizcaíno
Departamento de Sanidad Animal, Fac. de C.C. Veterinaria, Universidad Complutense de Madrid, Madrid, Spain

In this article, we consider Classical Swine Fever (CSF). CSF is a highly contagious viral disease of domestic and wild pigs caused by the CSF Virus (CSFV) (Moennig 2000). It generates important economic losses (due to, for instance, the devaluation of the pig meat price, the cost of culling and cleaning infected farms, the price of detection campaigns, etc.) in the affected regions (Fernández et al. 2011; Niemi et al. 2008). Despite the efforts to control and eradicate CSF, it remains endemic in many countries of America, Africa and Asia and sporadic outbreaks have affected 19 European countries from 1996 to 2007 (Edwards et al. 2000). Due to the different ways of CSFV spread (airborne, contact with infected animals, etc.) (Elbers et al. 1999; Ribbens et al. 2004), it is difficult to extrapolate the routes of infection and the consequences of a CSF epidemic from one region to another. Furthermore, the magnitude and duration of a CSF epidemic change depending on the epidemiological and demographic characteristics of the infected area and the timing and effectiveness of the applied control measures (Mangen et al. 2002).

The study of the potential spread patterns of CSFV into a region may help to identify risk zones to improve the prevention and management of future outbreaks. In CSF-free areas, a good way to quantify the magnitude of potential CSF epidemics and evaluate the efficiency of control measures is to use simulation models. Recently, some models have been developed to simulate CSFV spread into CSF-free regions such as Belgium, Germany, Australia and The Netherlands (Jalvingh et al. 1999; Kartsen et al. 2005a; Mintiens et al. 2003; Stegeman et al. 1999). In Martínez-López et al. (2009), authors have also described a spatial stochastic model for Spain by using the commercial software InterSpread Plus (Sanson 1993; Massey Univ. 2012). However, most of those models only focus on the between-farm spread of the CSFV, with poor assumptions regarding the within-farm spread and do not explicitly consider the specific farm to farm contact patterns (such as commercial network, shared vehicles, etc.) into the studied region.

To overcome these limitations, a spatial hybrid model, called Be-FAST (Between Farm Animal Spatial Transmission), has been developed (Martínez-López et al. 2011, 2012). Compared to the other models introduced previously, Be-FAST presents interesting novel characteristics. One of the most important is that it is based on the combination of a stochastic Individual-Based model (where the farms are considered as individuals) (DeAngelis and Gross 1991; Kartsen et al. 2005a), simulating the between-farm interactions (such as pig shipments, veterinarian routes, etc.) and CSFV spread, with a Susceptible-Infected model (Brauer and Castillo-Chavez 2001), simulating the within-farm CSFV spread. In particular, Be-FAST uses the proportions of infected pigs in each contaminated farm to calibrate dynamically some of its coefficients (for instance, those referring to the probabilities of infection). Another newsworthy feature of Be-FAST is the use of real database for farms and transports allowing simulating realistic commercial contacts between farms. We note that, in recent literature, other epidemiological models consider a similar approach (for instance, in Lyytikäinen et al. 2011), the authors introduce a model based on a Monte Carlo algorithm which uses real farms and transports databases but does not simulate the number of infected animals within each farm). At the end of a simulation, Be-FAST returns outputs referring to CSF outbreaks characteristics (such as the mean epidemic magnitude, the CSFV introduction risk map, the proportions of CSF infection routes and the proportions of detection due to control measures). This model has been previously described from the veterinarian point of view (i.e., choice of the CSFV transmission routes to be modeled or neglected, interpretation of the results, etc.) in Martínez-López et al. (2011) and an extensive sensitivity analysis of its parameters and outputs behavior was presented in Martínez-López et al. (2012).

This paper has two objectives. The first one is to give a complete mathematical formulation of Be-FAST in order to provide a transparent and understandable model for users. The

second one is to validate our model. To do so, we compare some of the Be-FAST results obtained considering two particular numerical experiments with the outputs given by two other models (InterSpread Plus Sanson 1993 and the Individual-Based model presented in Kartsen et al. 2005a) and real outbreak information. Those experiments simulate the possible evolution of a CSF epidemic inside the particular Spanish province of Segovia (J.C.L. 2008). In the first case, the CSF epidemic spreads freely without any control measure whereas in the second case we consider all the control measures defined by the Spanish administration. Similar validation approach has been considered in previous literature (for instance, see Jalvingh et al. 1999; Kartsen et al. 2007).

Here, after recalling in Sect. 2 the main characteristics of the CSF, we give an extended description of the Be-FAST model from the mathematical perspective (i.e., detailed equations, numerical schemes, etc.) in Sect. 3. Finally, during Sect. 4, we focus on the model validation by considering various numerical experiments based on real databases provided by the Spanish administration (J.C.L. 2008; M.A.P.A. 2006). We also compare the results given by our model with those obtained by other models (Sanson 1993; Kartsen et al. 2007) and data observed during CSF outbreaks occurring in Spain and The Netherlands (Martínez-López 2009; Elbers et al. 1999; Jalvingh et al. 1999).

2 Characteristics of Classical Swine Fever

In order to facilitate the understanding of the Be-FAST model, described in Sect. 3, we briefly explain the CSF evolution process, the routes of CSFV transmission and some control measures used to fight CSFV. A complete justification of the assumptions and simplifications considered during this work can be found in Martínez-López et al. (2011).

2.1 CSF evolution

CSF results from infection by CSFV, a member of the genus Pestivirus and family Flaviviridae (Moennig 2000). CSFV affects both domestic and wild pigs. When a pig is not infected by CSFV, it is categorized in the Susceptible state. Once it is infected, it passes successively through the following states (Elbers et al. 1999):

- Infected: The pig is infected by CSFV but cannot infect other pigs and has no visible clinical signs (i.e., fever, lesions, etc.). The mean duration of a pig in this state is 7 days and is called latent period. Then, the pig passes to be infectious.
- Infectious: The pig can infect other pigs but does not have clinical signs. The mean duration from infectious to the development of clinical signs is 21 days and is called incubation period. After that period, the pig presents clinical signs.
- Clinical Signs: The pig develops visible clinical signs and still may infect other pigs. After a period between two weeks and three months the pig can be recovered or died due to the disease. However, the CSF death and recuperation of pigs are assumed to be neglected because the culling of infected animals is usually accomplished before.

Those four states can be also applied at the farm level by considering that a farm is in the state (in order of priority) (Kartsen et al. 2005a):

- Clinical Signs (denoted by CF): If at least one pig has clinical signs.
- Infectious (denoted by TF): If at least one pig is infectious.
- Infected (denoted by IF): If at least one pig is infected.
- Susceptible (denoted by SF): If all pigs in the farm are susceptible.

A farm either in the state *IF*, *TF* or *CF* is called contaminated farm. Moreover, a farm in the state *TF* or *CF* is also called spreading farm.

2.2 Routes of transmission

The main ways of CSFV transmission (i.e., that a susceptible pig becomes infected) are: the direct contacts with infected animals; and the indirect contacts due to airborne spread of the CSFV and contaminated vehicles, persons or fomites (i.e., material). Historically, those ways of spreading have been reported as the principal routes of CSFV spread (Elbers et al. 1999; Ribbens et al. 2004), although other routes (such as movements of wild animals) have also been described as potential ways of CSFV transmission but with a minor impact on the CSF epidemics (Elbers et al. 1999). Those alternative routes have been neglected here.

2.3 Control measures

Once an animal becomes infected, another important concept in epidemiology is its detection and the application of control measures by the authorities (Moennig 2000).

When an infected pig is detected in a farm, this farm is classified as detected. Generally, before the first detection of a contaminated farm (called index case), the detection occurs when pigs present clinical signs and is due to the vigilance of the farmers or veterinarians (Koenen et al. 1996). After the detection of the index case, the awareness of the farmers and authorities is widely increased and the time of detection decreases (Kartsen et al. 2005a). Moreover, the detection can be also due to the control measures presented below.

In order to control a potential CSF epidemic, the following control measures defined by the European and Spanish legislations, described in Jalvingh et al. (1999), J.C.L. (2008), Martínez-López (2009), M.A.P.A. (2006), are considered:

- Movement restrictions: Outgoing or incoming movements in farms inside the considered region are limited during a specified time period (in our case, between 1 and 3 months).
- Zoning: Zones, called protection and surveillance zones, are defined around a detected farm (considering a minimum radius of 3 km and 10 km, respectively). Surveillance activities are applied within those zones during 30 and 40 days, respectively.
- Depopulation: All the animals of a detected farm are culled and destructed and the premise is cleaned and disinfected.
- Tracing: Tracing activities involve the process of determining contacts that have left or entered a detected farm during a time interval preceding the detection (here, 2 months). The objective of tracing is to identify potential infectious contacts which may have introduced CSFV into the farm or spread CSFV to other farms.

3 Mathematical formulation of the model

In this section, we describe in detail the Be-FAST model by presenting its general structure, the considered input parameters and the processes related to the CSFV spread and the control measures. The main notations used here are summarized in Table 1.

3.1 General description

The Be-FAST model is used to evaluate the spread of CSFV within and between farms into a specified region during a fixed time interval.

Table 1 Summary of the main notations used during this work to describe Be-FAST

Notation	Description
M	Number of scenarios considered in the Monte Carlo algorithm
SCE_m	m-th scenario of the Monte Carlo algorithm
T	Maximum number of simulation days of each Monte Carlo scenario
NF	Number of farms in the study region
$NP_i(t)$	Number of pigs in farm i at time t
$NSP_i(t)/NIP_i(t)$	Number of susceptible/infected pigs in farm i at time t
$SF/IF/TF/CF$	State of a farm: Susceptible/Infected/Infectious/Clinical Signs

At the beginning of the simulation, the model parameters are set by the user.[1] The parameters referring to farms and transports of pigs are detailed in Sect. 3.2. The other parameters are described in Sects. 3.3–3.6 and are summarized in Table 2. Furthermore, the control measures presented in Sect. 2.3 are also implemented and can be activated or deactivated, when starting the model, in order to quantify their effectiveness to reduce the magnitude and duration of a CSF epidemic.

Be-FAST is based on a Monte Carlo approach that generates $M \in \mathbb{N}$ epidemic scenarios (here, we call scenario to one iteration of the Monte Carlo algorithm that simulates a possible evolution of the CSFV). More precisely, at the beginning (i.e., at time $t = 0$) of each scenario, denoted by SCE_m with $m = 1, 2, \ldots, M$, all the farms are in the susceptible state except one farm, selected randomly by considering a discrete uniform distribution, which has one infectious pig and is classified as infectious. Then, during the time interval $[0, T]$, with $T \in \mathbb{N}$ the maximum number of simulation days, the within-farm and between-farm daily spread routines (described in Sects. 3.3 and 3.4, respectively) are applied. Moreover, the daily processes simulating the detection by authorities of contaminated farms and the activated control measures (presented in Sects. 3.5 and 3.6, respectively) are also run. If at the end of a simulation day all farms are in the susceptible state, the scenario SCE_m is stopped and we start the next scenario SCE_{m+1}. When the scenario SCE_M is over, many kinds of outputs can be generated. In Sect. 4.1.3, we present the most typical ones used to analyze the performance of an epidemiological model.

A diagram summarizing all those steps is presented in Fig. 1.

Remark 1 Currently, Be-FAST does not simulate the possible introduction or spread of CSFV due to contacts with foreigner regions. In fact, we simulate scenarios assuming that: the foreigner sources of CSF have been already identified and cannot send infected pigs to Segovia; and, once the index case in Segovia is detected, zoning is implemented and other regions do not buy animals from Segovia (thus, we neglect this way of infection). This is

[1] The value of the parameters used by Be-FAST should be set in function of the studied region (for example, due to the specific legislations). For instance, the parameters values considered during this work are adapted for their application to the province of Segovia (see Sect. 4). In particular the parameters referenced by J.C.L. (2008) or M.A.P.A. (2006) in Table 2 have been obtained by expert opinions of the Spanish administration. The reliability of those parameters is discussed in Martínez-López et al. (2012).

Table 2 Summary of the parameters used by the Be-FAST model[1] (except those referring to farms and transports of pigs which are presented in Sect. 3.2)

Parameter description	Dist./Value[2]	Reference
Transmission parameters of the SI model: $\beta_{far}/\beta_{fat}/\beta_{ftf}$	0.66/0.4/0.53	Klinkenberg et al. (2002)
Daily PI due to local spread	See Table 3	Kartsen et al. (2005a)
PI due to infectious vehicles transporting pigs	Ber(0.011)	Stegeman et al. (2002)
PI due to infectious INT vehicles	Ber(0.0068)	Stegeman et al. (2002)
PI due to infectious SDA persons	Ber(0.0065)	Stegeman et al. (2002)
Daily PD of the index case due to clinical signs	Ber(0.03)	Kartsen et al. (2005a)
Daily PD due to clinical signs	Ber(0.06)	Kartsen et al. (2005a)
PD due to tracing	Ber(0.95)	M.A.P.A. (2006)
Maximum daily PD due to protection zone	Ber(0.98)	J.C.L. (2008)
Maximum daily PD due to surveillance zone	Ber(0.95)	J.C.L. (2008)
PR of animal movements of detected farms	Ber(0.99)	J.C.L. (2008)
PR of INT movements of detected farms	Ber(0.95)	J.C.L. (2008)
PR of SDA movements of detected farms	Ber(0.80)	J.C.L. (2008)
PR of animal movements of zoned farms	Ber(0.95)	J.C.L. (2008)
PR of INT movements of zoned farms	Ber(0.90)	J.C.L. (2008)
PR of SDA movements of zoned farms	Ber(0.70)	J.C.L. (2008)
PR of all movements of non zoned farms	Ber(0.40)	J.C.L. (2008)
Probability of tracing animal movements	Ber(0.99)	J.C.L. (2008)
Probability of tracing INT movements	Ber(0.70)	J.C.L. (2008)
Probability of tracing SDA movements	Ber(0.40)	J.C.L. (2008)
Duration of general movement restriction	30 days	M.A.P.A. (2006)
Duration of protection/surveillance zones	30/40 days	M.A.P.A. (2006)
Radius of protection/surveillance zones	3/10 km	M.A.P.A. (2006)
Maximum number of traced farms per day	Poi(60)	J.C.L. (2008)
Maximum number of depopulated farms per day	Poi(20)	J.C.L. (2008)
Time to repopulate a depopulated farm	Poi(90) days	J.C.L. (2008)
Time to depopulate a detected farm	See Table 4	Elbers et al. (1999)
Tracing period	60 days	J.C.L. (2008)
Latent period	Poi(7) days	Kartsen et al. (2005b)
Incubation period	Poi(21) days	Kartsen et al. (2005b)
Daily number of contacts with INT vehicles of farms	Poi(0.4)	Kartsen et al. (2005b)
Number of farms visited by an INT vehicle per day	Poi(4)	J.C.L. (2008)
Daily number of contacts with SDA persons of farms	Poi(0.3)	Kartsen et al. (2005b)
Number of farms visited by a SDA person per day	Poi(3)	J.C.L. (2008)

[1] Abbreviations in the table: PI = Probability of Infection; PD = Probability of Detection; PR = Probability of Restriction; SDA = Sanitary Defense Association group; INT = Integrator group; Poi(X) = Poisson distribution with mean X; Ber(X) = Bernoulli distribution with mean X

[2] Distribution (Dist.) or value used in model validation (Sect. 4)

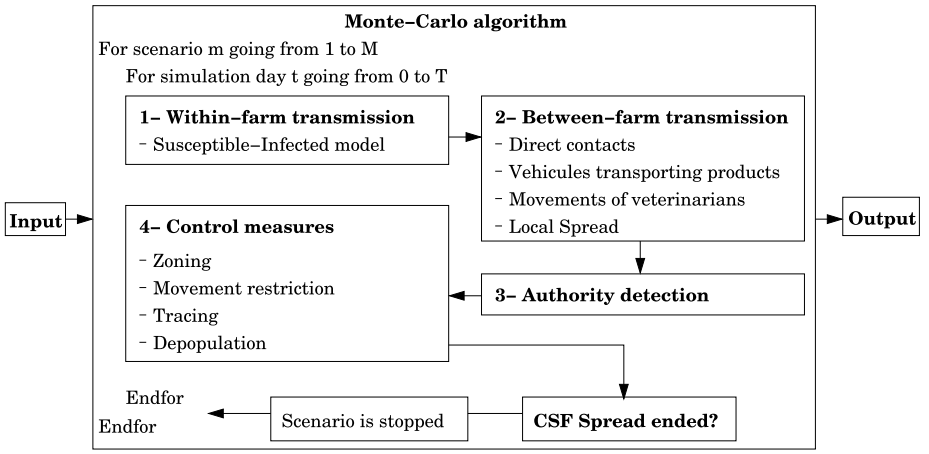

Fig. 1 Diagram summarizing the Be-FAST model steps presented in Sect. 3.1

a limitation of Be-FAST which is common with other epidemiological models (such as the one described in Kartsen et al. 2005a) and is discussed in Martínez-López et al. (2011).

3.2 Farms and transports of pigs inputs

We consider a study region containing $NF \in \mathbb{N}$ farms. For each farm, identified as farm number i (also called farm i) with $i = 1, \ldots, NF$, the following data are given: the geographical coordinates of the farm centroid; $NP_i(0) \in \mathbb{N}$ the number of pigs at the first day of the simulation (i.e., $t = 0$); the type of pig production of the farm (here, farrowing, fattening or farrow-to-finish) (Klinkenberg et al. 2002); the Integrator (INT) group (i.e., group of farms who share material and vehicles transporting products) identifier; the Sanitary Defense Association (SDA) group (i.e., group of farms who share veterinarians) identifier.

Furthermore, the following information of all farm to farm pig shipments occurring during a specified time interval is also provided: the number of pigs shipped; the date of the shipment; and the farms of origin and destination of the shipment.

3.3 Within-farm CSFV spread

The daily CSFV spread within a particular contaminated farm i is modeled by using a discrete time stochastic Susceptible-Infected (SI) model (Brauer and Castillo-Chavez 2001; Klinkenberg et al. 2002). We assume that the pigs in this farm are characterized to be in one of those two states: Susceptible or Infected, described in Sect. 2.1. In order to reduce the computational complexity of our model (see Remark 2), the Infectious and Clinical Signs states are simulated only at the farm level (more details are given in Sect. 3.4). The natural pig mortality is also neglected (Martínez-López et al. 2011).

Under those assumptions, the evolution of $NSP_i(t)$ and $NIP_i(t)$, denoting the number of susceptible and infected pigs in farm i at time t, respectively, is given (in a continuous version) by

$$\frac{\mathrm{d}NSP_i(t)}{\mathrm{d}t} = -\beta_i \frac{NSP_i(t)NIP_i(t)}{NP_i(t)}, \qquad \frac{\mathrm{d}NIP_i(t)}{\mathrm{d}t} = \beta_i \frac{NSP_i(t)NIP_i(t)}{NP_i(t)}, \tag{1}$$

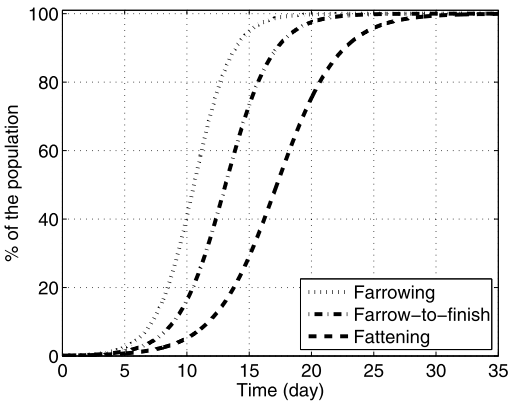

Fig. 2 Time evolution of the percentage of infected pigs obtained by considering System (1) and a farm containing 1000 pigs and with one infected pig at time $t = 0$, in function of the farm type: farrowing, fattening and farrow-to-finish

where $NP_i(t) = NSP_i(t) + NIP_i(t)$ is the number of pigs in farm i at time t, $\beta_i \in \mathbb{R}$ is the transmission parameter set to $\beta_{\text{far}} = 0.66$, $\beta_{\text{fat}} = 0.40$ or $\beta_{\text{ftf}} = 0.53$ depending on the farm type: farrowing, fattening or farrow-to-finish pig farms, respectively (Klinkenberg et al. 2002). The evolution of the proportion of infected pigs governed by System (1) and obtained by considering a farm containing 1000 pigs and with one infected pig at time $t = 0$, in function of the farm type, is presented in Fig. 2.

Moreover, we are interested in obtaining integer values of infected and susceptible pigs in farms and in introducing some randomness in System (1) (the within-farm CSFV spread may be slightly different from one farm to another) but respecting its general behavior. Thus, we consider the following discrete version of System (1) (Klinkenberg et al. 2002)

$$NSP_i(t+1) = NSP_i(t) - \min\bigl(\Gamma_i(t), NSP_i(t)\bigr),$$
$$NIP_i(t+1) = NIP_i(t) + \min\bigl(\Gamma_i(t), NSP_i(t)\bigr),$$

where t corresponds to the simulation day and $\Gamma_i(t) \in \mathbb{N}$ follows a Poisson distribution with mean $\beta_i(NSP_i(t)NIP_i(t))/NP_i(t)$.

Remark 2 Although the SI model presented in this section seems to be too simple (with only two pig states) to simulate the within-farm CSFV spread, in practice it gives a good ratio between spread modeling accuracy and computational time. Indeed, we have to consider that this model is applied to each farm infected during a Monte Carlo scenario, which can dramatically increase the computational time needed by the Be-FAST model. For example, we have tried to consider the Infectious and Clinical Signs states at the pig level. In that case, we have obtained results similar to the ones given by the model described here (2 % of variation) with a significant increase of the computational time (+30 %). For the same reasons, considering a model more complex than a SI model (such as one that simulates the spatial diffusion of the CSFV within a farm) does not appear to be a reasonable choice in terms of model efficiency (for instance, if an epidemiological model is used to support decisions in cases of real outbreaks, its outputs should be given within a day, J.C.L. 2008).

3.4 Between-farm CSFV spread

The CSFV spread between farms is modeled by using a spatial stochastic Individual-Based model (DeAngelis and Gross 1991). In this model, farms are classified in one of those four

states (see Sect. 2.1): Susceptible (*SF*), Infected (*IF*), Infectious (*TF*) and Clinical signs (*CF*). The daily transition from a particular farm state to another one is modeled by considering the direct and indirect contacts and the natural evolution of the CSF presented in Sects. 2.1 and 2.2. Those transition processes are described in Sects. 3.4.1–3.4.3.

3.4.1 State transition due to direct contacts

The CSFV spread by direct contacts is assumed to occur due to the movements of infected pigs between farms. Those movements are estimated by using the data of the shipments of pigs introduced in Sect. 3.2. Although this database usually contains previous years' information (see Sect. 4.1.1), the transports of pigs are similar from one year to another with slight variations (J.C.L. 2008). Thus, we have decided to use random movements, generated from the pig transports data, instead of using the exact ones (using exact or random shipments does not impact dramatically the model outputs, see Martínez-López et al. 2012). Due to the construction process described below, those random movements exhibit similar characteristics (i.e., farms of origin and destination, date and number of moved pigs) than the real ones. However, this process allows to consider, with a low probability, transports to or from farms not included in the shipment database (for instance, due to errors). More precisely, at each simulation day t, those shipments are simulated as following:

We compute the number of movements occurring during the simulation day t by considering a Poisson distribution with mean being the number of all movements occurring at day t in our database. Then, for each simulated movement:

- We select randomly the farm of origin of the movement $i \in [1, \ldots, NF]$ and the farm of destination of the movement $j \in [1, \ldots, NF]$, with $i \neq j$, by considering the discrete probability $\mathbb{P}_M((i,j) = (k,l)) = M_{\text{mov}}(k,l)/(\sum_{m=1}^{NF} \sum_{n=1, n\neq m}^{NF} M_{\text{mov}}(m,n))$, where k and $l \in [1, \ldots, NF]$, $k \neq l$ and $M_{\text{mov}}(k,l) \in \mathbb{R}$ is the number of movements from farm k to farm l in the database plus 10^{-6} (to take into account with a small probability possible movements not occurring in our database). We note that $\mathbb{P}_M$ is computed once before the simulations and only each time we get a new database (other parameters related to the database may be also calculated once before running the model).
- We compute $w_{(i,j)} = \min\{\text{Ceil}(\overline{w_{(i,j)}}(NP_i(t)/NP_i(0))), NP_i(t)\}$, the number of pigs shipped during this movement from farm i to farm j. In the previous expression, $\overline{w_{(i,j)}} \in \mathbb{R}$ denotes the mean number of pigs moved when considering all the shipments between farms i and j in our database and $\text{Ceil}(x)$ returns the nearest integer greater than or equal to $x \in \mathbb{R}$. In the case of no movement from farm i to farm j in the database, $\overline{w_{(i,j)}}$ is set to the mean number of pigs moved taking into account all the database shipments.
- Next, we move $w_{(i,j)}$ pigs from the origin farm i to the destination farm j. Those pigs are selected randomly in $NSP_i(t)$ and $NIP_i(t)$, considering a discrete uniform distribution. We denote by $w^S_{(i,j)} \in \mathbb{N}$ and $w^I_{(i,j)} \in \mathbb{N}$ the number of susceptible and infected pigs that are moved during this simulated shipment, respectively. Thus, the evolution of pigs in farm i and j is governed by $NSP_i(t) = NSP_i(t) - w^S_{(i,j)}$, $NIP_i(t) = NIP_i(t) - w^I_{(i,j)}$, $NSP_j(t) = NSP_j(t) + w^S_{(i,j)}$ and $NIP_j(t) = NIP_j(t) + w^I_{(i,j)}$.
- Finally, if $w^I_{(i,j)} > 0$, the state of farm j is set to the state of farm i in one of the following cases: The state of farm j is *SF*; the state of farm j is *IF* and the state of farm i is *TF* or *CF*; the state of farm j is *TF* and the state of farm i is *CF*. In other cases, the state of farm j remains unchanged.

3.4.2 State transition due to indirect contacts

As specified in Sect. 2.2, the CSFV spread due to indirect contacts is assumed to occur by either movements of vehicles transporting pigs, movements of INT vehicles (i.e., vehicles transporting products from or to farms in the same INT group), movements of SDA persons (i.e., persons visiting farms in the same SDA group) or the so called local spread (i.e., spread due to contacts with the neighborhood which include: airborne spread and contaminated persons or fomites in the vicinity).

In Paragraphs A–C, we describe those four kinds of indirect contacts and the way they contribute to the CSFV spread from farm to farm. Then, in Paragraph D, we show how this spread affects the farms at the levels of pig numbers and state.

A. Movements of vehicles transporting pigs We consider the pig shipments generated in Sect. 3.4.1. If the farm of origin of the transport is either in the state *TF* or *CF*, the truck transporting pigs is considered as contaminated and, thus, can infect the farm of destination. In that case, we assume that the probability of CSFV infection in the farm of destination due to the contact with the contaminated vehicle is modeled by using a Bernoulli distribution with mean 0.011 (Stegeman et al. 2002).

B. Movements of INT vehicles and SDA persons The CSFV spread by contact with INT vehicles (SDA persons, respectively) visiting farms is assumed to occur only among the farms belonging to the same INT group (SDA group, respectively) and with the following assumptions:

- The daily number of contacts with INT vehicles (SDA persons, respectively) per farm is assumed to be Poisson distributed with a mean of 0.4 (0.3, respectively) (Kartsen et al. 2005b).
- An INT vehicle (SDA person, respectively) can visit a maximum of 4 (3, respectively) farms per day (J.C.L. 2008).
- An INT vehicle (SDA person, respectively) is contaminated if, previously, it has visited a spreading farm (i.e., a farm either in the state *TF* or *CF*, see Sect. 2.1) (Kartsen et al. 2005a; Stegeman et al. 2002).
- The probability of CSFV infection in a farm per contact with a contaminated INT vehicle (contaminated SDA person, respectively) is modeled by using a Bernoulli distribution with mean 0.0068 (0.0065, respectively) (Stegeman et al. 2002).

Thus, for each simulation day and for each integrator group INT, we build the routes of the INT vehicles and we simulate the way they spread CSFV by considering the process described below:

- For each farm in INT, we compute the number of INT vehicles visiting it by using a Poisson distribution with mean 0.4.
- Then, we list the farms that are visited by the INT vehicles and we rearrange this list, denoted by L_{INT}, randomly (taking into account that a farm cannot be visited twice consecutively).
- Next, a first INT vehicle visits the first four farms in L_{INT}, following the list order. Each fourth farm, until the end of L_{INT}, we consider a new INT vehicle (non-contaminated) starting from the next farm in L_{INT}.
- During each simulated trip, an INT vehicle becomes contaminated at the moment it visits a spreading farm and can infect other farms by considering a Bernoulli distribution with mean 0.0068.

Table 3 Interpolation points used to compute the function $\overline{PIL}$ (Kartsen et al. 2005a)

Distance x (meters)	0	150	250	500	1000	2000
Value of $\overline{PIL}(x)$	0.02	0.014	0.009	0.0038	0.0019	0

The same method is used to model the itineraries and the CSFV spread of SDA persons, but with the corresponding parameters values.

C. Local spread We assume that the CSFV local spread occurs to farms in the proximity of a farm either in the state *TF* or *CF*. It is mainly due to the airborne spread and contacts with contaminated neighborhood persons or fomites.

During this work, the daily probability of CSFV infection in farm j, due to the local spread from spreading farm i at simulation day t, is modeled by considering a Bernoulli distribution with mean $\overline{PIL}(d(i,j)) \times (NIP_i(t)/\overline{N(0)})$, where $\overline{N(0)} = (\sum_{i=1}^{NF} NP_i(0))/NF$ is the mean number of pigs per farm at day 0, $d(i,j)$ is the distance (in meters) between the centroid of farms i and j and $\overline{PIL}(x) \in [0,1]$ is the mean daily probability of CSFV infection due to local spread between two farms at a distance x. Moreover, $\overline{PIL}(x)$ is built by interpolating the data presented in Table 3 (Kartsen et al. 2005a).

D. New infection and state transition For each new CSFV infection occurring at day t in farm i during the processes described in Paragraphs *A* to *C*, if $NSP_i(t) \geq 1$, we infect one new pig in farm i by considering $NSP_i(t) = NSP_i(t) - 1$ and $NIP_i(t) = NIP_i(t) + 1$. Furthermore, if the state of farm i is *SF*, we change it to *IF*.

3.4.3 State transition due to the CSF natural evolution

According to the characteristics of the CSF evolution described in Sect. 2.1, we consider the following changes in the farm state: when a farm reaches the state *IF* (or *TF*), it will pass at state *TF* (*CF*) after a latent period (incubation period) that follows a Poisson distribution with mean 7 (21) days (Kartsen et al. 2005b).

3.5 Detection of contaminated farms

As specified in Sect. 2.3, a contaminated farm is generally detected by the observation of clinical signs of its pigs (i.e., the farm is in state *CF*) (Koenen et al. 1996). Thus, for each farm in the state *CF*, its daily probability of detection is modeled by using a Bernoulli distribution with mean 0.03 (0.06, respectively), before (after, respectively) the detection of the index case (Jalvingh et al. 1999; Kartsen et al. 2005a). Furthermore, a contaminated farm can be also detected due to the control measures presented in Sect. 3.6.

3.6 Control measures

We now describe the control measures, introduced in Sect. 2.3, implemented in Be-FAST.

3.6.1 Movement restrictions

A drastic restriction on movements (outgoing or incoming in farms) is applied to detected farms. Reductions on the transports of animals, the INT vehicles movements and the SDA

Table 4 Probability of the number of days to wait before depopulating a detected farm (Elbers et al. 1999)

Number of days	0	1	2	3	4	5	6	7	8
Probability	0.11	0.58	0.2	0.06	0.04	0.004	0.003	0.0015	0.0015

persons movements in the detected farms are assumed to be Bernoulli distributed with a mean of 0.99, 0.95 and 0.8, respectively.

Furthermore, after each detection, a general movement restriction considering those three kinds of movements, which follows a Bernoulli distribution with mean 0.4, is applied to all farms[2] during 90 days (J.C.L. 2008; M.A.P.A. 2006).

3.6.2 Zoning

The farms at a distance of less than 3 km from a detected farm are set in a protection zone during 30 days, whereas the farms at a distance between 3 km and 10 km from a detected farm are set in a surveillance zone during 40 days (M.A.P.A. 2006). Overlapping of the period of a farm in protection and surveillance zones is allowed (i.e., if a farm is already within a zone, we add the days of the new zone to those of the old zone).

A movement restriction is applied to farms within those zones (called zoned farms) (M.A.P.A. 2006). Pig transports, movements of INT vehicles and movements of SDA persons are randomly reduced by considering a Bernoulli distribution with mean 0.95, 0.9 and 0.7, respectively (M.A.P.A. 2006; J.C.L. 2008). Furthermore, we apply another surveillance process to zoned farms, in addition to the one described in Sect. 3.5. The daily probability of detection of farm i in the state CF due to this new surveillance is assumed to be dependent on the proportion of infected animals and is modeled by considering a Bernoulli distribution with mean $(\alpha \times NIP_i(t))/NP_i(t)$, where $\alpha = 0.98$ ($\alpha = 0.95$, respectively) if farm i is within a protection (surveillance, respectively) zone (J.C.L. 2008).

3.6.3 Depopulation

The depopulation (i.e., the culling of all animals) of detected farm i occurs after a random time period, generated by using the data provided by Table 4 (Elbers et al. 1999), starting from the day of its detection. However, the maximum number of farms that can be depopulated per day is assumed to follow a Poisson distribution with mean 20 (J.C.L. 2008). Thus, if this limit is reached, the farm is depopulated the following days. When farm i is depopulated, its number of pigs is set to 0 and it is not considered anymore by the model. Then, after a period following a Poisson distribution with mean 90 days (M.A.P.A. 2006), the farm is repopulated (i.e., new pigs are introduced): its number of susceptible pigs is $NP_i(0)$, its state is set to SF and it is again taken into account by Be-FAST.

3.6.4 Tracing

The objective of tracing is to identify infectious contacts which may have introduced CSFV into a detected farm or spread CSFV to other farms. We include the tracing of all contacts (i.e., farms sending/receiving animals or sharing SDA persons/INT vehicles) of a detected

[2]This control measure is adapted for studying CSFV spread in the province of Segovia (see Sect. 4). For larger areas (e.g., a country), the movement restrictions should be limited to a part of the studied region.

farm occurring 60 days before the detection (M.A.P.A. 2006). However, due to registration failures (e.g., errors in databases), tracing all the contacts is not always possible.

More precisely, when farm i is detected, we list all the farms who have shared, 60 days before the detection, at least one INT vehicle, one SDA person or one animal shipment with this farm i. Then, for each farm in this list, we decide if it is traced or not according to the probabilities of tracing a farm due to animal transport, INT vehicle movement or SDA person movement which are assumed to be Bernoulli distributed with a mean of 0.99, 0.7 and 0.4, respectively (J.C.L. 2008). Next, for each farm to be traced, we select its day of tracing taking into account, as in Sect. 3.6.3, that the maximum number of traced farms per day is governed by a Poisson distribution with mean 60. Finally, we perform a detection process to the traced farms, the day of their tracing, by considering that the probability of detecting a contaminated traced farm follows a Bernoulli distribution with mean 0.95 (M.A.P.A. 2006).

4 Model validation

In order to validate our model, we perform various numerical experiments, described in Sect. 4.1. The results obtained by Be-FAST, presented in Sect. 4.4, are compared with those generated by two other epidemiological models considering similar experiments (see Sect. 4.2) and with data observed during real CSF outbreaks (see Sect. 4.3).

4.1 Numerical experiments

In the following, we present the numerical experiments used for the model validation. In particular, we detail the inputs, the scenarios parameters and the considered outputs.

4.1.1 Farms and pig transports inputs

We consider the Spanish province of Segovia,[3] one of the most important pig production areas in Spain, which has a surface of 6796 km^2. We use real databases of the year 2008, provided by the Spanish Regional Government of Castilla and Leon (J.C.L. 2008) and the Spanish Ministry of the Environment and Rural and Marine Affairs (M.A.P.A. 2006), corresponding to the inputs described in Sect. 3.2. In 2008, $NF = 1400$ pig farms, containing a total of 1403800 pigs, were located in Segovia. Those farms were divided in 11 INT groups and 34 SDA groups. Furthermore, 208 farms were of the type farrowing, 510 fattening and 682 farrowing-to-finish. Finally, during this year, there were 10046 pig shipments. The locations of those pig farms and the province of Segovia are shown in Fig. 3.

4.1.2 Scenarios parameters

We have considered two kinds of simulations:

In the first one, we do not consider any control measure and we run the model during $T = 200$ days (this value gives a good ratio between results precision and computational time, see Martínez-López et al. 2011, 2012). This case is denoted by **NM** (No Measure). The interest of this experiment is to evaluate the principal routes of CSFV spread.

[3]This reduced region has been already used to illustrate, with success, the behavior of our model in Martínez-López (2011, 2012) and, thus, is also applied here to validate our model. However, this region could be replaced by any larger region when the required databases are available.

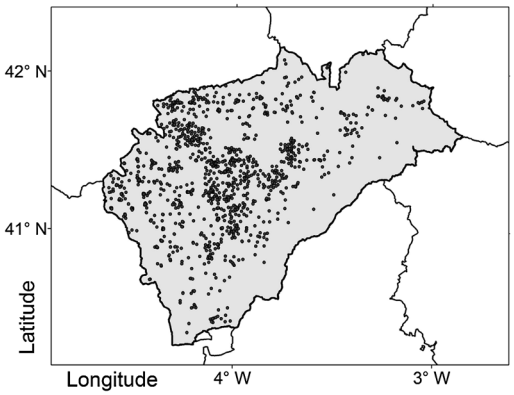

Fig. 3 Coordinates and boundaries of the province of Segovia (*in gray*). The locations of the considered pig farms are represented by *black spots* (•)

In the second one, all the control measures described in Sect. 3.6 are activated and the model is run during a maximum period of three years (i.e., $T = 1095$ days), which is large enough to ensure the end of a CSF epidemic (Martínez-López et al. 2011). This case is denoted by **WM** (With Measures). In this experiment, which is more realistic and classical (i.e., this experiment is also considered in other works, such as Kartsen et al. 2007 and Mangen et al. 2002) than **NM**, we are interested in evaluating the magnitude of the epidemic and the efficiency of the control measures.

In both cases, we set $M = 1000$ scenarios. This value gives a good compromise, in the particular experiments considered here for the Be-FAST model, between the stability (with variations less than 3 %) of the outputs presented in Sect. 4.1.3 (including risk maps) and the computational complexity (Martínez-López et al. 2012). Furthermore, we want to point out that, in the literature, some models are run with a much lower value. For instance, in Jalvingh et al. (1999), Kartsen et al. (2005a, 2005b, 2007), the number of scenarios of the considered Monte Carlo algorithm is 100. This reinforces the idea of precision of the results obtained during this work. We also remark that this value of M is adapted to our particular number of farms $NF = 1400$. When considering a greater value of NF, M should be increased proportionally (Lyytikäinen et al. 2011).

4.1.3 Considered outputs

When the scenario SCE_M is over, many types of outputs can be obtained. Here, we consider the most typical ones used to evaluate the performances of an epidemiological model (Jalvingh et al. 1999; Kartsen et al. 2007; Martínez-López et al. 2009).

More precisely, for each scenario SCE_i with $i = 1, \ldots, M$, we compute: the number of infected farms, denoted by NIF; and the duration of the epidemic (i.e., the number of days between the beginning and the ending of the scenario), denoted by DUR. For those both quantities, we calculate, regarding all the scenarios, their mean, minimum and maximum values; their 95 % prediction interval; their quartiles; and their discrete distribution functions.

In addition, taking into account the M scenarios, we evaluate: the percentages of infection due to local spread, INT vehicles, SDA persons and transports of pigs; and the percentages of detection of contaminated farms, after detecting the index case, due to observation of the clinical signs, zoning and tracing.

Furthermore, for each farm i, we compute its risk of CSFV introduction, denoted by $RI(i)$. It is defined as the number of times that farm i becomes contaminated during the

Table 5 Weight coefficients of the contaminated farms, in function of the number of days after the farm infection, used in our simulations with the InterSpread Plus model

Number of days	0	5	10	13	15	20	22
Weight coefficient	0.001	0.03	0.17	0.55	0.7	0.95	1

whole Monte Carlo simulation. In particular, in order to identify the risk zones in the studied region, we are interested in obtaining the geographical distribution of RI. Typically (Mintiens et al. 2003), the risk zones are classified in three categories: high, medium and low risk. This is useful, for instance, to design preventive control measures to fight CSFV (Fernández et al. 2011). To do so and to compare the values of RI given by the models presented in Sect. 4.2, we first normalize $RI(i)$ by considering $\bar{RI}(i) = \hat{RI}(i)/(\max_i \hat{RI}(i))$, where $\hat{RI}(i) = RI(i)/(\sum_i RI(i))$. Then, we obtain the spatial distribution of $\bar{RI}$ in Segovia by interpolating the values of $\bar{RI}(i)$ considering an inverse distance weighted method. Finally, the identification of the three risk zones is done by applying the Jenks Natural Breaks (JNB) classification method. Those both last steps are done by using ArcGIS 9.1 (E.S.R.I. 2012).

4.2 Comparison with other models

In order to validate the Be-FAST model, we perform the experiments presented in Sect. 4.1 by using the two following models:

(A) **Be-FAST model (BF)**: A Matlab 7.8 (MathWorks 2012) implementation of Be-FAST.
(B) **InterSpread Plus model (IS)**: We also consider InterSpread Plus 1.0.49.5 (Massey Univ. 2012), a commercial C++ implementation of a state transition Individual-Based model presented in Sanson (1993). It is one of the most popular epidemiological modeling software used worldwide. However, in our opinion, it has several drawbacks, such as the scarce transparency of the code (i.e., it is a black-box program) and the difficulty to incorporate complex databases with real farm-to-farm movements or contacts.

We intend to reproduce the same processes as the one used by **BF**. As it is out of the scope in this paper, we do not describe the **IS** model in detail, but we only present here the differences between both models (other parts are similar).

IS does not allow simulating the within-farm transmission (it is a purely between-farm spread model), so it is not possible to compute the number of infected or susceptible pigs of a particular farm i at a given time. For this reason, the model coefficients cannot be expressed as a function of $NSP_i(t)$ or $NIP_i(t)$. Thus, all the coefficients depending on those numbers are set to constant values: the daily probability of infection due to local spread is set to $\overline{PIL}(x)$ without interpolation; the probability of infection due to a pig shipment coming from a contaminated farm is 1; the daily probabilities of detection of a contaminated farm in a protection or surveillance zone follow a Bernoulli distribution with mean 0.98 and 0.95, respectively. However, for each contaminated farm i, **IS** allows to assign to this farm a weight coefficient, depending on the number of days after the farm infection, that is multiplied to all probabilities of infection due to contact with farm i. This process simulates the fact that the infectiousness of a farm increases with the time. The weight coefficients are reported in Table 5 and fit the SI evolution of a farrow-to-finish farm, depicted in Fig. 2.

Moreover, the real commercial network (i.e., pig shipments, SDA and INT groups) cannot be integrated directly in **IS**. First, the swine transport process is simulated as follows.

Table 6 Probability of selecting a farm of destination for pig shipments, in function of the distance (in kilometers) with the farm of origin, used in our simulations with the InterSpread Plus model

Distance interval (kilometers)	[0–15]	[15–30]	[30–45]	[45–60]	[60–80]	>80
Probability of selection	0.4	0.29	0.18	0.07	0.04	0.02

For each simulation day and farm i, we have to compute the number of pig transports sent by this farm. This is done by considering a Poisson distribution with mean given by the daily mean number of pig transports sent by farm i. Then, for each shipment, we select randomly a farm of destination according to the distance with farm i and the probability distribution given in Table 6. The values of the daily mean number of pig transports sent by each farm and Table 6 are obtained from our database. Secondly, the SDA and INT contacts are simulated as in our model but considering only one SDA and INT group. The farms to be visited are randomly selected in function of their distances according to Table 6. From those simplifications, we see that **IS** does not allow to incorporate the real commercial contacts between farms.

Finally, as a last difference with **BF**, **IS** considers that when a farm is infected by a pig shipment its state is set to *IF*.

Furthermore, we also compare some results obtained by **BF** and **IS** with those given by the model described below on a region different than Segovia by considering experiments similar to the ones presented in Sect. 4.1.

(C) **Kartsen et al. model (KM)**: This Individual-Based model, presented in Kartsen et al. (2005a, 2005b), simulates only the between-farm spread and is applied to a fictitious German region, but based on real statistics, in Kartsen et al. (2007). The considered region, which has an area of 2230 km^2, contains 2986 farms (1896 fattening farms, 543 farrowing farms, 546 farrow-to-finish farms). No data on the commercial network between farms are considered. The simulated processes for the CSFV spread and the control measures are similar to the ones used by **BF** (except the spread due to artificial insemination which is neglected in **BF** and **IS** due to its low probability of infection, see Stegeman et al. 1999). Furthermore, the model parameters are adapted to the German country and can be found in Kartsen et al. (2005a). However, they are close to the ones used in this work. The interesting results obtained by **KM**, reported in Kartsen et al. (2007), are the mean, minimum and maximum values of the variables *NIF* and *DUR* for the **WM** experiment.

4.3 Comparison with real epidemic data

In order to compare the validity of the results given by **BF** and **IS**, we consider data observed during the following two real CSF epidemics occurring in 1997–98:

(A) **Segovia**: A description of this epidemic can be found in Martínez-López (2009) and data were provided by J.C.L. (2008). During this event, 22 farms were infected and the epidemic duration was approximatively 60 days. As Segovia is the region studied during our experiments, we use the spatial location of those infected farms to validate the risk maps generated by **BF** and **IS**. However, data regarding the proportions of infection routes and control measures efficiency are not available. Thus, we have to consider another real case.

Table 7 Results obtained when solving the **NM** (No Measure) and **WM** (With Measures) experiments considering models **BF** (Be-FAST) and **IS** (InterSpread Plus): computational time (Comp. Time), in seconds, needed to solve each case; percentages of infection due to each CSFV route (Route) obtained by considering the **NM** case: local spread (*LS*), Integrator vehicles (*INT*), Sanitary Defense Association persons (*SDA*) and transport of animals (*TA*); percentages of detection of contaminated farms due to each control measure (Measure) obtained by considering the **WM** case: observation of clinical signs (*CS*), zoning (*ZO*) and tracing (*TR*). Row **Real** corresponds to the percentages observed during the real epidemics occurring in 1997–98 in The Netherlands (Elbers et al. 1999; Jalvingh et al. 1999)

Model	Comp. Time		Route				Measure		
	NM	**WM**	*LS*	*INT*	*SDA*	*TA*	*CS*	*ZO*	*TR*
BF	28000	14500	54	26	14	6	47	30	23
IS	4000	11000	51	13	10	26	38	50	12
Real	–	–	52	24	20	4	55	27	18

(B) **The Netherlands**: This case, detailed in Elbers et al. (1999), Jalvingh et al. (1999), is interesting as this country has similar pigs production characteristics than Segovia. Due to the high epidemic magnitude (429 infected farms) and the quality of the experiments done in situ, the results provided in this literature are assumed to be valid reference values. Here, we are interested in the proportions of infection due to each CSFV route and detection of each control measure observed during this epidemic and reported in Table 7 in the **Real** row.

Remark 3 We note that real epidemic data should be considered as only one possible evolution of the disease. However, they give some interesting reference values in order to interpret the output given by simulation models. Of course, this interpretation should be considered as an indicator of the validity of a model but not as a rigorous fact. This approach is assumed to be appropriate in literature (for instance, see Jalvingh et al. 1999; Kartsen et al. 2007).

4.4 Results

All the simulations presented in this Section are run on a computer with a 2.4 GHz Core2 Duo P8600 CPU, 4 GB of DDR3 memory and Windows Vista 32 bit OS. The outputs, described in Sect. 4.1.3, obtained during those experiments are reported in Tables 7 and 8 and, some of them, are depicted in Figs. 4 and 5.

The computational time needed by models **BF** and **IS** to solve the **NM** and **WM** experiments are presented in Table 7—(Comp. Time columns). We can see that the **IS** model is the fastest one. In particular, for the **NM** case (with highest numbers of infected farms), the difference between both models is quite high (**BF** is 7 times slower). In the **WM** case, which is more realistic, the difference is reasonable. This can be explained, in part, by the fact that our model has a more complex and complete structure than **IS** (for instance, the use of SI model for each contaminated farm, dynamic coefficients, etc.), and therefore, requires more computations. Another explanation is the choice of the program languages used by **BF** and **IS**. Our model is implemented in Matlab script, an interpreted language known to be slow in comparison to compiled languages (DeRose and Padua 1999), such as C++, which is used to implement **IS**. Matlab was chosen in order to obtain quickly a first implementation of our model and for the easiness to process the outputs. The computational time required by **BF** could be significantly decreased by programming it in C++.

Table 8 Some statistical values on the number of infected farms (*NIF*) and the epidemic duration (*DUR*) obtained by considering the **NM** (No Measure) and **WM** (With Measures) experiments and models **BF** (Be-FAST) and **IS** (InterSpread Plus): mean value (Mean); minimum value (Min.); maximum value (Max.); 95 % Prediction Interval lower (PI[2.5 %]) and upper (PI[97.5 %]) bounds; and quartiles (Q1, Q2 and Q3). Some data available for the model **KM** in the **WM** case are also reported (Kartsen et al. 2007)

Model	Output	Mean	Min.	PI [2.5 %]	Q1	Q2	Q3	PI [97.5 %]	Max.
NM									
BF	*NIF*	32	1	1	4	16	40	122	339
IS	*NIF*	58	1	1	7	33	84	255	523
WM									
BF	*NIF*	3.3	1	1	1	2	3	14	53
BF	*DUR*	63	14	25	38	51	80	178	428
IS	*NIF*	4.6	1	1	1	1	3	34	68
IS	*DUR*	79	14	28	38	54	81	326	729
KM	*NIF*	7.5	1	–	–	–	–	–	56
KM	*DUR*	84	20	–	–	–	–	–	230

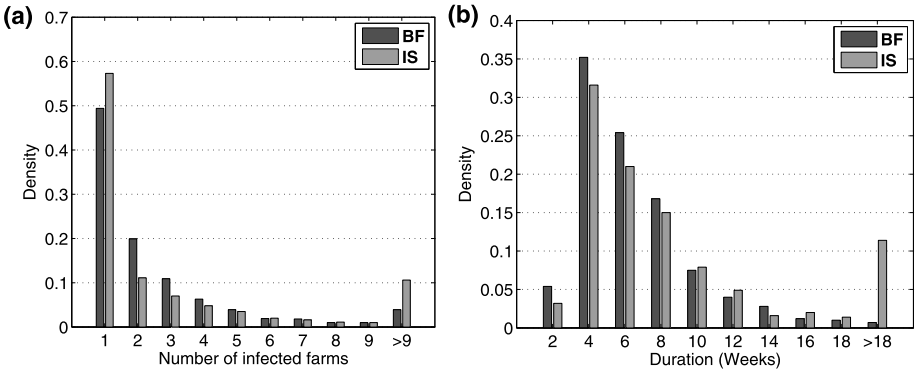

Fig. 4 Discrete distributions of (**a**) the number of infected farms (*NIF*) and (**b**) the epidemic duration (*DUR*) (in weeks) obtained by models **BF** (Be-FAST, *dark gray*) and **IS** (InterSpread Plus, *light gray*) for the **WM** (With Measures) experiment

The percentages of infection in function of the CSFV routes are given in Table 7—(Route columns). We can see that both models identify the local spread as the main source of infection. Moreover, the proportions due to SDA persons are quite similar in the two cases. The main differences between models **BF** and **IS** are obtained when regarding the transports of animals and the INT vehicles proportions. In that case, **IS** considers the animal shipments as the second most important cause of CSFV infection instead of the INT vehicles. This discordance may be due to the fact that **IS** does not take into account the number of infected pigs when performing the animal transports, whereas **BF** uses this information. Thus, when the number of infected animals in the farm of origin of the transport is low, many of the pig shipments simulated by **BF** do not infect the destination farms, decreasing the proportion of infection due to this route. When we compare those results with the proportions observed

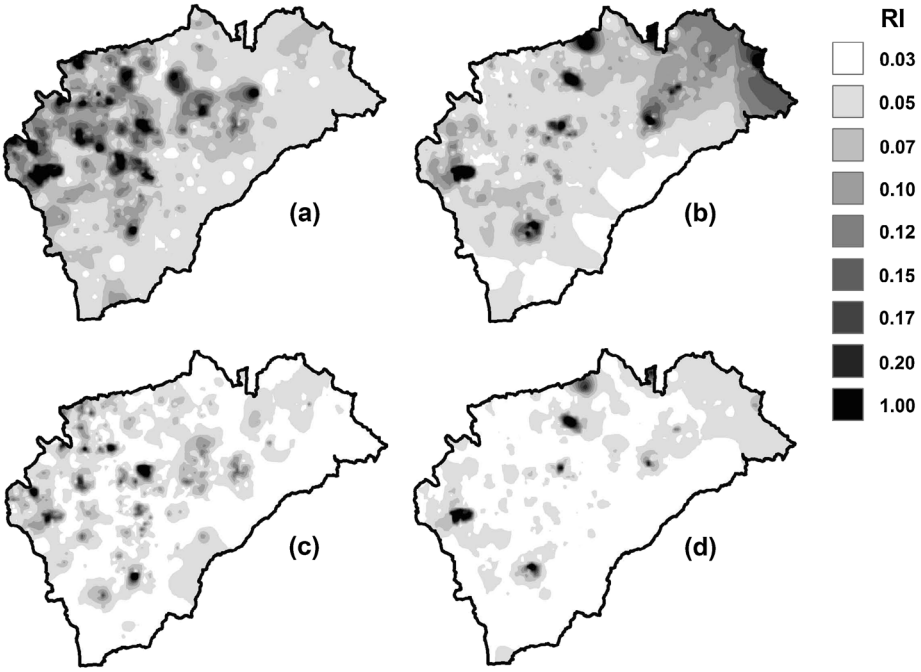

Fig. 5 Normalized risk of CSFV introduction ($\bar{RI}$) interpolated maps obtained by models (**a**) **BF** (Be-FAST) for the **NM** (No Measure) case, (**b**) **IS** (InterSpread Plus) for the **NM** case, (**c**) **BF** for the **WM** (With Measures) case and (**d**) **IS** for the **WM** case. The grayscale colormap corresponding to the considered Jenks Natural Breaks classification, described in Sect. 4.1.3, is also reported

during the 1997–98 epidemic in The Netherlands (Jalvingh et al. 1999), we see that the **BF** outputs fit better those real data (in particular the *INT* and *TA* columns) than the **IS** ones.

The percentages of detection of contaminated farms in function of the control measures are reported in Table 7—(Measure columns). On one hand, **IS** considers that the zoning is the most efficient control measure, then the observation of the clinical signs and, finally, the tracing. On the other hand, **BF** returns the observation of clinical signs as the best detection technique, the zoning and the tracing presenting similar efficiency. This can be explained by the fact that, as described in Sect. 3.6.2, **BF** uses the number of infected animals in zoned farms to generate the probabilities of detection due to the zoning process (thus, the efficiency of this control measure may be reduced for farms with a low proportion of infected animals), whereas **IS** does not allow this possibility. As previously, the percentages generated with **BF** are closer to the real data reported during the 1997–98 epidemic in The Netherlands (Elbers et al. 1999) than those given by **IS**.

Both results in the proportions of infection and detection seem to indicate that our approach, which consists in simulating the number of infected animals in farms and using it in the formulas of the **BF** model coefficients, provides better results than not considering it and is suitable for generating epidemiological models presenting a realistic behavior.

The statistical values associated to the number of infected farms (*NIF*) and the duration of the epidemic (*DUR*), obtained by the **BF** and **IS** models during the **NM** and **WM** experiments and some of those values available for the **KM** model, are reported in Table 8. The discrete distribution functions of *NIF* and *DUR* for the **WM** case are presented in Fig. 4. We can observe on the table, that **IS** generates slightly larger values of *NIF* and *DUR*, but

of the same order, than **BF**. This is an expected result when considering the differences between the models coefficients, in particular, the use (or not) of the proportions of infected animals which increases or decreases the probabilities of infection and detection. In fact, the main discrepancy is observed on the amplitude of the extreme scenarios (i.e., scenarios with numerous infected farms) which is higher for **IS** than **BF**. This can be observed in Fig. 4, where the discrete densities are quite similar for both models except for the highest values (i.e., $NIF > 9$ and $DUR > 18$). This is confirmed, in the **WM** case, by the fact that the minimum, PI[2.5 %], Q1, Q2 and Q3 values of both models are close and the PI[97.5 %] and the maximum values are twice higher for **IS** than **BF**. However, in the **NM** experiments, since the mean value of **NIF** is higher (i.e., there are more extreme scenarios) than in **WM** ones, the difference between both models is amplified: Q2 and Q3 are also more than twice greater for **IS** than **BF**. When regarding the results produced by **KM** in the **WM** case, and taking into account that the considered region has a double number of farms and the area is smaller than Segovia, they can be consider as similar to those produced by **BF** and **IS**. Regarding the effect of applying or not the control measures, in both models we observe a similar behavior: the epidemic is reduced by ten when comparing the **NM** and **WM** experiments.

When considering the amplitude of the 1997–98 epidemic in Segovia, which consisted in 22 infected farms and had a duration of 60 days, it is difficult to compare it with the **BF** and **IS** results obtained in the *WM* case: the values of *DUR* are close to the real outbreak length, but the values of *NIF* are much lower. We have to take into account that 10 years separate the 2008 database used in our experiments and the real situation in Segovia in 1997. During this period, more than half of the farms have disappeared, due to an economic crisis in 2006 (Martínez-López 2009), and the control measures have been highly reinforced after the tremendous CSF epidemic in Europe during the years 1997 and 1998 (Ribbens et al. 2004). Moreover, it is possible that this particular epidemic represents an extreme scenario of the model. A good way to compare those real data with the **BF** and **IS** outputs, is to analyze the risk maps and see if the 1997–98 infected farms are in high risk zones.

The $\bar{RI}$ risk maps generated by models **BF** and **IS** for the **NM** and **WM** experiments are presented in Fig. 5. The Jenks Natural Breaks (JNB) classification, divided in 9 intervals corresponding to 9 gray colors (for a better understanding of the maps), is also reported in this figure: the first three intervals [0–0.03], [0.03–0.05] and [0.05–0.07] correspond to the low risk areas; the intervals [0.07–0.10], [0.10–0.12] and [0.12–0.15] correspond to the medium risk areas; and the last three intervals [0.15–0.17], [0.17–0.20] and [0.20–1] correspond to the high risk areas. This classification is obtained by considering the **NM** case (i.e., the worst case) with the **BF** model and is extended to the other maps. We point out that the JNB classifications obtained by **IS** are similar to the **BF** ones. As we can observe on those maps, the risk distributions obtained by both models decrease drastically from the **NM** cases to the **WM** ones. We can also see that, although both models identify similar high risk zones in the South-West of Segovia, **IS** concentrates the risk in some specific areas in the North and the East parts, whereas **BF** identifies the center of the region as presenting a high risk of CSFV spread. This is particularly visible on the **WM** maps.

Focusing on the **WM** case, we consider the farms infected during the 1997–98 epidemic in Segovia and see the risk zones where they are included. In Fig. 6, we incorporate those farms to the **BF** and **IS** risk maps and we detail the area where most of the farms are included. We can see that, in the **BF** case, a large majority of the infected farms is situated in a dark gray (high risk) zone and the other farms in medium or low risk zones. In the **IS** case, the high risk zone does not include those farms which are mainly located in low risk areas. The mean $\bar{RI}$ value of the 1998–97 infected farms given by the **BF** model is 0.201, which is included in the highest risk interval for the considered JNB classification. In the **IS** model,

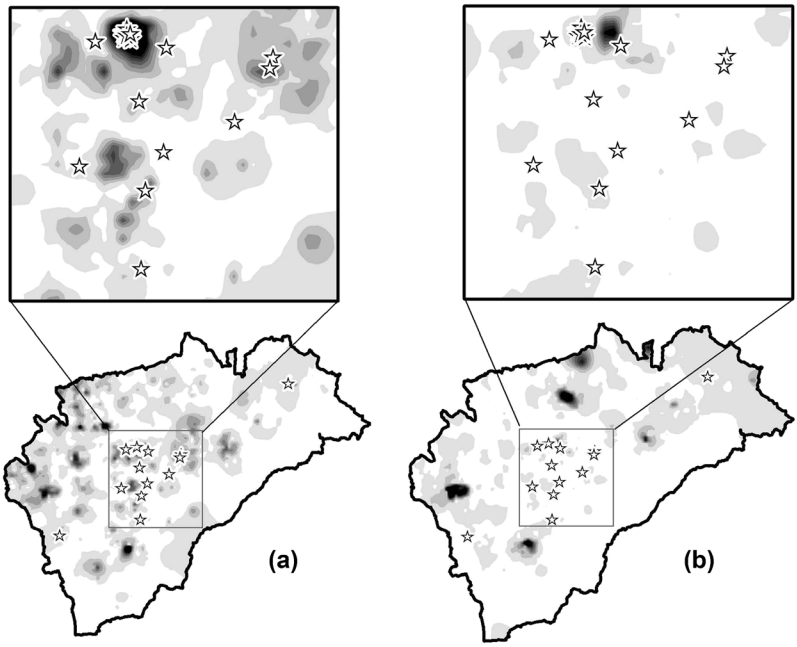

Fig. 6 Normalized risk of CSFV introduction ($\bar{RI}$) interpolated maps obtained by models **(a) BF** (Be-FAST) and **(b) IS** (InterSpread Plus) for the **WM** (With Measures) experiment. We also report, with *white stars* (☆), the location of the farms infected during the 1997–98 CSF epidemic in Segovia. Furthermore, we present in the square region a zoom of the zone where most of those farms were situated (except two of them)

the mean risk value of those farms is 0.032, which corresponds to a low risk. This result tends to show that the maps generated by **BF** are more consistent with real data than those given by **IS**. This can be explained by the fact that our model uses the real commercial network between farms (i.e., transports of animals and SDA and INT groups), whereas this information is not suitably processed by **IS**. This shows the importance of using this database to obtain a fine representation of the risk areas, and one should include this input in an epidemiological model if it is available. As previously, we point out that 10 years separate the used databases from the 1997–98 outbreaks in Segovia and the real data considered for the **BF** and **IS** inputs, explaining why some farms are included in low risk zones, even in the **BF** map. However, this also shows the robustness of the **BF** risk maps, which seem to be valid for years different from those of the databases.

5 Conclusions

During this work, we first have presented an extended mathematical formulation of the spatial hybrid model called Be-FAST, previously introduced from a veterinarian point of view in Martínez-López et al. (2011, 2012), used for the study of CSFV spread into a region. In particular, we have described the processes considered to simulate the infection routes within and between farms and the control measures applied by authorities to fight this disease.

Then, in order to validate our model, we have performed various numerical experiments by considering farms and transports real databases of the Spanish province of Segovia. We

have compared the results given by Be-FAST with those obtained by two other models (i.e., InterSpread Plus and the Individual-Based model presented in Kartsen et al. 2005a) and with real outbreaks data from Spain and The Netherlands. We have seen that, due to some of the characteristics of Be-FAST (i.e., the combination of a Susceptible-Infected model with an Individual-Based model, the use of the proportions of infected animals in farms to calibrate some model coefficients and the consideration of the real commercial network between farms), the outputs generated by our model fit better the real CSF epidemics information than the ones produced by InterSpread Plus. Those results tend to show the validity and the efficiency of our approach.

One of the next steps of this work should be the incorporation in Be-FAST of a model to study the economic impact of a CSF epidemic in the considered region. It will be also interesting to use the risk map distribution to design CSF preventive campaigns, in order to reduce the economic losses and the risk of spread of future outbreaks. Those two last ideas are currently a work in progress and preliminary results can be found in Fernández et al. (2011). Finally, as specified in the article, other newsworthy studies will be the implementation of Be-FAST by using a fast programming language and to take into account the possible interactions with foreigner regions.

Acknowledgements This work was carried out thanks to the financial support of the Spanish Ministry of Science and Innovation under projects MTM2008-04621 and MTM2011-22658; the project CONS-C6-0356 of the I-MATH Proyecto Ingenio Mathematica; the Research Group MOMAT supported by the "Banco Santander" and the "Universidad Complutense de Madrid" (Ref. 910480); and the "Comunidad de Madrid" and "European Social Fund" through project S2009/PPQ-1551. We gratefully acknowledge the assistance of Olga Minguez, her team, the Regional Government of Castilla and Leon Region and the Spanish Ministry of the Environment and Rural and Marine Affairs for providing us data and technical assistance. A.M. Ramos has also been Funded by Fundación Caja Madrid.

References

Anderson, R., & May, R. (1979). Population biology of infectious diseases: Part 1. *Nature*, *280*, 361–367.

Brauer, F., & Castillo-Chavez, C. (2001). *Mathematical models in population biology and epidemiology*. Berlin: Springer.

DeAngelis, D., & Gross, L. (1991). *Individual-based models and approaches in ecology*. New York: Chapman and Hall.

DeRose, L., & Padua, D. (1999). Techniques for the translation of Matlab programs into Fortran 90. *ACM Trans. Program. Lang. Syst.*, *21*(2), 285–322.

Edwards, S., Fukusho, A., Lefevre, P., Lipowski, A., Pejsak, Z., Roehe, P., & Westergaard, J. (2000). Classical Swine Fever: the global situation. *Vet. Microbiol.*, *27*(3), 103–109.

E.S.R.I. (2012). *ArcGIS*. http://www.esri.com/software/arcgis/.

Elbers, A., Stegeman, A., Moser, H., Ekker, H., Smak, J., & Pluimers, H. (1999). The CSF epidemic 1997–1998 in The Netherlands: descriptive epidemiology. *Prev. Vet. Med.*, *4*, 157–184.

Fernández, E., Ivorra, B., Ramos, A., Martínez-Lopez, B., & Sánchez-Vizcaíno, J. (2011). *Diseño de un modelo económico y de planes de control para una epidemia de Peste Porcina Clásica*. Preprint de la Universidad Complutense de Madrid. http://www.ucm.es/centros/cont/descargas/documento28250.pdf.

Jalvingh, A., Nielen, M., Maurice, H., Stegeman, A., Elbers, A.R., & Dijkhuizen, A. (1999). Spatial and stochastic simulation to evaluate the impact of events and control measures on the 1997–1998 CSF epidemic in The Netherlands. *Prev. Vet. Med.*, *42*, 271–295.

J.C.L. (2008). *Junta de Castilla y Leon—Agricultura y Ganadería, expert opinion elicitation performed for FMD and CSF*. http://www.jcyl.es.

Kartsen, S., Rave, G., & Krieter, J. (2005a). Monte Carlo simulation of CSF epidemics and control I. General concepts and description of the model. *Vet. Microbiol.*, *108*, 187–198.

Kartsen, S., Rave, G., & Krieter, J. (2005b). Monte Carlo simulation of CSF epidemics and control II. Validation of the model. *Vet. Microbiol.*, *108*, 199–205.

Kartsen, S., Rave, G., Teuffert, J., & Krieter, J. (2007). Evaluation of measures for the control of CSF using a simulation model. *Arch. Tierz.*, *50*, 92–104.

Klinkenberg, D., DeBree, J., Laevens, H., & DeJong, M. C. M. (2002). Within- and between-pen transmission of Classical Swine Fever Virus: a new method to estimate the basic reproduction ratio from transmission experiments. *Epidemiol. Infect.*, *128*, 293–299.

Koenen, F., Van Caenegem, G., Vermeersch, J., Vandenheede, J., & Deluyker, H. (1996). Epidemiological characteristics of an outbreak of Classical Swine Fever in an area of high pig density. *Vet. Rec.*, *139*(15), 367–371.

Lyytikäinen, T., Niemi, J., Sahlström, L., Virtanen, T., & Lehtonen, H. (2011). *The spread of Foot-and-Mouth Disease (FMD) within Finland and emergency vaccination in case of an epidemic outbreak*. Finnish Food Safety Authority Evira, Helsinki. Research Reports, 1/2011. http://www.evira.fi/portal/en/evira/publications/?a=view&productId=240.

Mangen, M., Nielen, M., & Burrell, A. (2002). Simulated effect of pig-population density on epidemic size and choice of control strategy for Classical Swine Fever epidemics in The Netherlands. *Prev. Vet. Med.*, *56*(2), 141–163.

M.A.P.A. (2006). *Ministerio de Agricultura, Pesca y Alimentación, Manual práctico de actuaciones contra la PPC*. http://rasve.mapa.es/Publica/InformacionGeneral/Documentos/Manuales/Manual%20PPC%20enero%202011.pdf.

Martínez-López, B. (2009). *Desarrollo de modelos epidemiológicos cuantitativos para el análisis del riesgo de introducción y difusión potencial de los virus de la Fiebre Aftosa y de la Peste Porcina Clásica en España*. PhD thesis, Univ. Complutense de Madrid, Spain.

Martínez-López, B., Ivorra, B., Ramos, A. M., & Sánchez-Vizcaíno, J. (2011). A novel spatial and stochastic model to evaluate the within and between farm transmission of CSF Virus: 1. General concepts and description of the model. *Vet. Microbiol.*, *147*(3), 300–309.

Martínez-López, B., Ivorra, B., Ramos, A. M., & Sánchez-Vizcaíno, J. (2012). A novel spatial and stochastic model to evaluate the within and between farm transmission of CSF Virus: 2. Model sensitivity analysis. *Vet. Microbiol.*, *155*(1), 21–30.

Martínez-López, B., Pérez, A., & Sánchez-Vizcaíno, J. (2009). A stochastic model to quantify the risk for Classical Swine Fever Virus introduction through import of domestic and wild boars into Spain. *Epidemiol. Infect.*, *137*(10), 1505–1515.

Massey Univ. (2012). *InterSpread Plus*. http://www.interspreadplus.com.

MathWorks (2012). *Matlab*. www.mathworks.com/products/matlab/.

Mintiens, K., Laevens, H., Dewulf, J., Boelaert, F., Verloo, D., & Koenen, F. (2003). Risk analysis of the spread of CSF virus through neighborhood infections for different regions in Belgium. *Prev. Vet. Med.*, *60*(1), 27–36.

Moennig, V. (2000). Introduction to Classical Swine Fever: virus, disease and control policy. *Vet. Microbiol.*, *73*(2), 93–102.

Niemi, J., Lehtonen, H., Pietola, K., Lyytikäinen, T., & Raulo, S. (2008). Economic implications of potential Classical Swine Fever outbreaks for Finnish pig production sector. *Prev. Vet. Med.*, *84*, 194–212.

Ribbens, S., Dewulf, J., Koenen, F., Laevens, H., & de Kruif, A. (2004). Transmission of Classical Swine Fever. A review. *Vet. Q.*, *26*, 146–155.

Sanson, R. (1993). *The development of a decision support system for an animal disease emergency*. PhD thesis, Massey University, New Zealand.

Stegeman, A., Elbers, A., Bouma, A., & DeJong, M. (2002). Rate of inter-farm transmission of Classical Swine Fever Virus by different types of contact during the 1997–1998 epidemic in The Netherlands. *Epidemiol. Infect.*, *128*, 285–291.

Stegeman, A., Elbers, A., Smak, J., & DeJong, M. (1999). Quantification of the transmission of Classical Swine Fever Virus between farms during the 1997–1998 epidemic in The Netherlands. *Prev. Vet. Med.*, *42*, 219–234.