



Modelling the 2007 Chikungunya Outbreak in Italy

A. Annunziato, N.I. Stilianakis, T. de Groeve,
T.P. Weber, L. Vernaccini, S. Gadenz

EUR 23065 EN - 2007

The Institute for the Protection and Security of the Citizen provides research-based, systems-oriented support to EU policies so as to protect the citizen against economic and technological risk. The Institute maintains and develops its expertise and networks in information, communication, space and engineering technologies in support of its mission. The strong cross-fertilisation between its nuclear and non-nuclear activities strengthens the expertise it can bring to the benefit of customers in both domains.

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen

Contact information

Address: Via Fermi 1
E-mail: alessandro.annunziato@jrc.it
Tel.: +39 0332 789519
Fax: +39 0332 785584

<http://ipsc.jrc.ec.europa.eu/>
<http://www.jrc.ec.europa.eu/>

Legal Notice

Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use which might be made of this publication.

***Europe Direct is a service to help you find answers
to your questions about the European Union***

Freephone number (*):

00 800 6 7 8 9 10 11

(* Certain mobile telephone operators do not allow access to 00 800 numbers or these calls may be billed.

A great deal of additional information on the European Union is available on the Internet. It can be accessed through the Europa server <http://europa.eu/>

JRC42264

EUR 23065 EN
ISSN 1018-5593

Luxembourg: Office for Official Publications of the European Communities

© European Communities, 2007

Reproduction is authorised provided the source is acknowledged

Printed in Italy

Table of Content

1	General introduction.....	5
2	The outbreak and basic descriptive statistics of symptoms.....	6
2.1	Estimation of the time evolution of the epidemic	10
2.1.1	The basic reproduction number R_0	10
2.1.2	Estimation for the doubling time of human cases	11
3	A simple deterministic compartmental model for chikungunya	13
3.1	Model assumptions.....	13
3.2	Model results	14
3.3	Concluding remarks	16
4	An agent-based approach	17
4.1	Introduction to Agent Based models.....	17
4.2	Application to chikungunya	20
4.2.1	Mosquito model.....	21
4.2.2	Human model	22
4.2.3	Initial condition	23
4.2.4	Actual cases	23
4.3	Calculations results.....	23
4.3.1	Initial case in position (1).....	23
4.3.2	Initial case in position (2).....	24
4.3.3	Initial case in position (3).....	24
4.3.4	Initial case in position (4).....	25
4.3.5	Effective population concept.....	26
4.4	Discussion of the results.....	27
5	Discussion on the two approaches.....	29
5.1	Complementarity of the two approaches.....	29
5.2	Simulation of a large city not possible with an ABM	29
5.3	How is it possible to use the global method in real time modelling	29
6	Conclusions	30
	Appendix A – Geolocalization procedure	33
	Appendix B – Description of the methodology used for retrieve population density in the affected area	35

1 General introduction

Chikungunya fever (CHIK) is a crippling arthritic disease of sudden onset, first recorded in 1952 in Tanzania. The responsible virus (CHIKV) was isolated in 1953 (Ross 1956). The virus cycles between insect vectors (mosquitoes from the genera *Aedes* and *Culex*) and vertebrate reservoir species, principally non-human primates (Diallo et al. 1999). CHIKV causes epidemics in India, Southeast Asia, Indonesia, the Philippines and in most of sub-Saharan Africa. In African rural areas the disease is endemic with a small number of cases every year, whereas in urban areas outbreaks are sporadic and explosive. After an epidemic, large portions of the population have become immune and the disease typically vanishes from the affected area for years.

CHIK has developed into a major public health concern for a number of reasons.

- 1) one important vector, the aggressive mosquito *Aedes albopictus*, has in the past decades successfully invaded many parts of the world, including the United States and several European countries.
- 2) an outbreak in Eastern Africa, the West Indian Ocean islands (including the French overseas territory of La Réunion), India and Indonesia between 2004 and 2007 affected approximately 3 million people (Pialoux et al. 2007) and infected travelers from these regions showed up in considerable numbers in Europe.

If infected persons and competent vectors in sufficient density come together in a susceptible population an outbreak can conceivably occur anywhere; these factors came in fact together in the summer of 2007 in the Italian region of Emilia Romagna causing the first certain self-sustaining outbreak of CHIK in Europe with 301 confirmed cases so far.

The behavioural, physiological and ecological factors determining the vectorial capacity of European populations of *Aedes albopictus*, the mosquito vector responsible for the outbreak in Italy are not well-known. Host preference and host range, preference for indoor or outdoor feeding, incubation period of the virus in the vector, longevity, length of the gonotrophic cycle, vector density etc. all co-determine the capability of the vector to cause a self-sustaining epidemic. Given the current lack of reliable empirical knowledge concerning these factors, but having at hand the epidemic curve of the Italian outbreak, modelling allows us a preliminary exploration of different scenarios and parameter combinations and even control measures.

2 The outbreak and basic descriptive statistics of symptoms

The list of all the cases has been obtained by the Italian Health Ministry in collaboration with the Istituto Superiore di Sanità and the Regione Emilia Romagna Health Authority and JRC is very grateful¹.

The first case of chikungunya fever occurred on June 23, 2007 and the disease then slowly spread through the two villages of Castiglione di Cervia and Castiglione di Ravenna; the outbreak lasted approximately 3 months. On days 56 and 60 aggressive mosquito control measures were undertaken in the affected area.

Figure 1 shows the daily number of new cases (i.e. daily incidence) in the two villages, while Figure 2 is the cumulative curve (the integral of Figure 1. It will be later explained that for comparison purposes the integral better shows the behaviour.

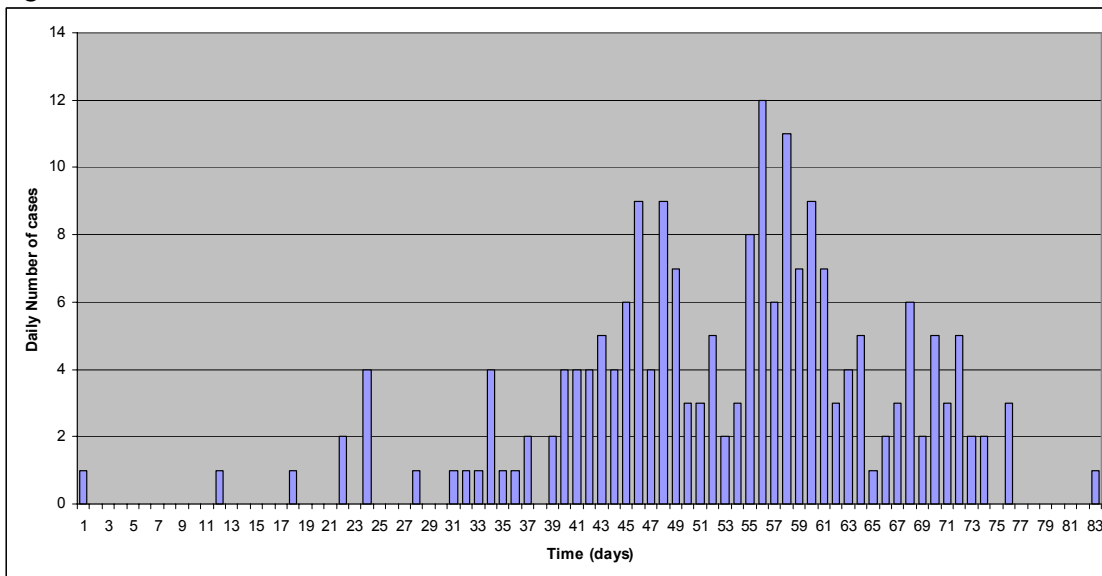


Figure 1. Daily incidence of chikungunya cases in Castiglione di Cervia and Castiglione di Romagna., Day 0 is considered 23/6/2007

¹ The database that we have received is dated 15 October 2007. An updated version of this database exists. A revision of this document will be performed when the new (and final) version of the database will be provided.

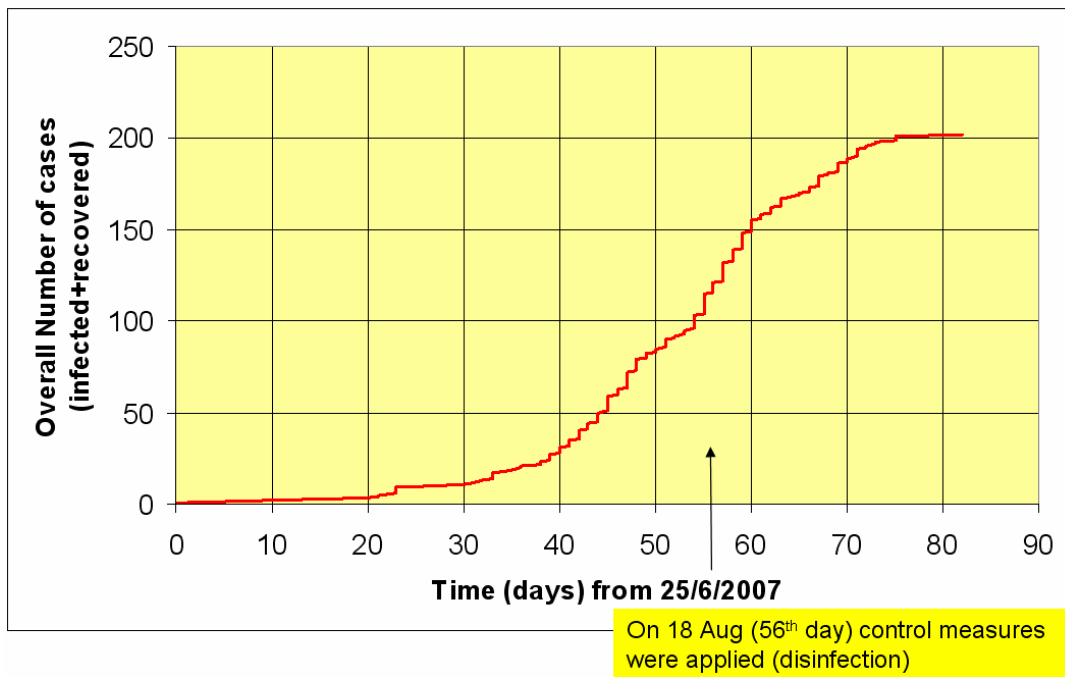


Figure 2. – Temporal distribution of the Chikungunya cases

Chikungunya virus infection can cause a debilitating illness, most often characterized by fever, limb and muscle pain and a range of other symptoms. It is likely that also asymptomatic infections occur, but how commonly that happens is not yet known. Table 1 shows a summary of the major symptoms and the proportion of persons affected.

Table 1. Summary of symptoms

Symptoms	Yes	No
Fever	181 (92%)	16 (8%)
Joint pain	182 (92%)	15 (7%)
Skin rash	98 (50%)	99 (50%)
Myalgia	92 (47%)	105 (53%)
Asthenia	183 (93%)	14 (7%)
Haemor. symptoms	3 (2%)	194 (98%)
Headache	100 (51%)	97 (41%)
Pruritus	36 (18%)	161 (82%)
Nausea	6 (3%)	191 (97%)
Vomiting	32 (16%)	165 (84%)
Diarrhoea	38 (19%)	159 (81%)
Photophobia	27 (14%)	170 (86%)
Conjunctivitis	6 (3%)	191 (97%)
Meningoencephalitis	1 (1%)	196 (99%)

Figure 3a-e shows the age distribution of cases and the distribution of the duration of some major symptoms (table 2 gives a summary of the data).

Figure 3a Age distribution

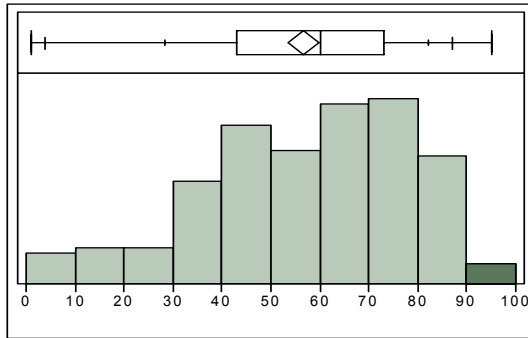


Figure 3b: Fever duration

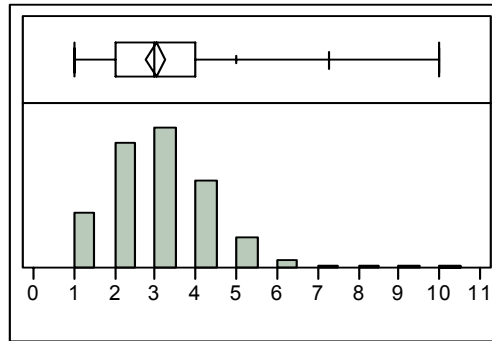


Figure 3c: Limb pain duration

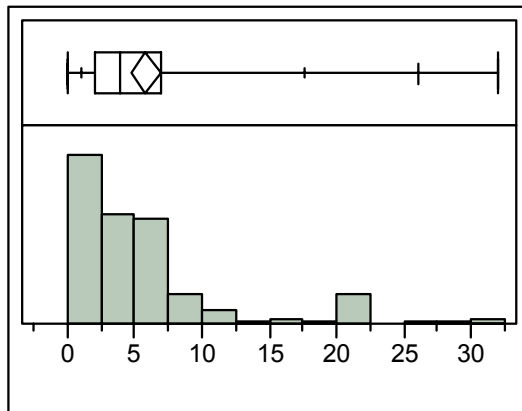


Figure 3d: Asthenia duration

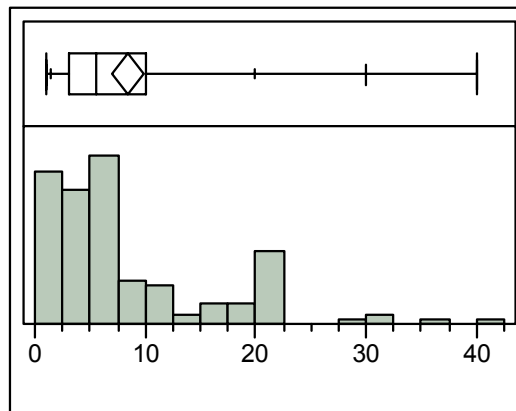


Figure 3e: Muscle pain duration

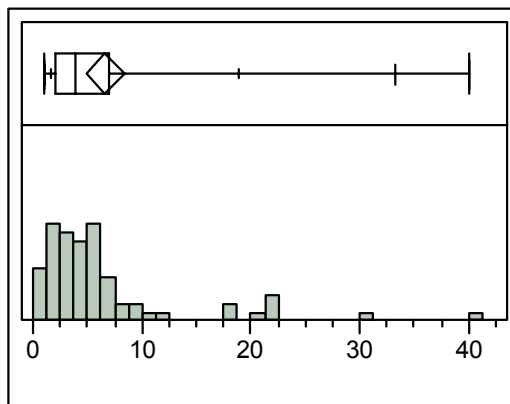


Table 2. Summary statistics for symptoms.

	Mean	95%CI mean	SD	Median
Age N =197	57	54 - 60	21	60
Symptoms duration in days				
Fever N=148	3.0	2.7 - 3.3	1.5	3.0
Joint pain N=156	5.8	4.8 - 6.9	6.5	4.0
Asthenia N=45	8.4	6.9 - 9.8	8.0	5.5
Myalgia N=66	6.6	4.8 - 8.4	7.4	4.0

The median age of the diseased persons is close to the mean value, indicating a rather normal distribution of the age. The mean duration of symptoms is, as shown in table 2, in the range of 6.5 - 8.5 days and agrees well with the value of 7 days for the infectious period assumed by Bacaër (2007) for the La Réunion outbreak, although even the La Réunion data show high variability for infectious period in humans and incubation period in vectors.

The distributions of symptom durations are all left-skewed; a few individuals show some symptoms for more than 30 days. An estimation of the infectivity period based on symptoms is thus not straightforward. Some long-lasting symptoms such as joint pain will most likely not correlate with infectiousness. Parameter estimates for models (i.e. the infectious period) should, however, possibly still be based on median values, not means.

The median values for the main symptoms are in the range of 3 - 5 days.

The list of the cases in the database contained several fields, including the home address of the affected persons, from which it has been possible to derive the geo-location of the cases, as described in Appendix A.



Figure 4 – Geographical distribution of the cases. The red dot represents the index case location

2.1 Estimation of the time evolution of the epidemic

2.1.1 The basic reproduction number R_0

In epidemiology, the **basic reproduction number** of an infection is the mean number of secondary cases a typical single infected case will cause in a population with no immunity to the disease in the absence of interventions to control the infection. It is often denoted R_0 . This metric is useful because it helps determine whether or not an infectious disease will spread through a population (Anderson & May 1991). It was originally used by George MacDonal in 1952, who constructed population models of the spread of malaria.

When $R_0 < 1$ the infection will die out in the long run (provided infection rates are constant). But if $R_0 > 1$ the infection will be able to spread in a population.

Large values of R_0 may indicate the possibility of a major epidemic.

There are several ways to estimate R_0 . A formula to estimate R_0 for vector born diseases derived from the above modelling approach and being ‘classic’ is:

$$R_0 = \frac{\beta^2 q q' p}{\alpha \mu P}$$

where

β is the biting rate of the vector,

q, q' are the transmission probabilities per bite from vector to human and from human to vector

p is the vector population

P is the human population

$1/\alpha$ is the average infectious period in humans

$1/\mu$ is the life expectation of adult vectors

R_0 is proportional to the vector population p . A straightforward calculation of R_0 based on the above formula would lead to an overestimate of R_0 . The reason for that is that there are parameters of which the values are unknown, such as the population of the vector and the transmission probability from vector to human und vice versa. Thus, rough estimates and missing information on essential components of the epidemic don't allow a straightforward estimation of the R_0 base on the above formula. A different approach is needed (work in progress).

If surveillance followed the evolution of vector density before and during the epidemic R_0 could be estimated more reliably from the above formula

Values of R_0 of well-known infectious diseases		
Disease	Transmission	R_0
HIV/AIDS	Sexual contact	2-5
Diphtheria	Saliva	6-7
Influenza (1918 pandemic strain)	Airborne droplet	2-3
Measles	Airborne	12-18
Mumps	Airborne droplet	4-7
Pertussis	Airborne droplet	12-17
Polio	Faecal-oral route	5-7
Rubella	Airborne droplet	5-7
SARS	Airborne droplet	2-5
Smallpox	Social contact	6-7

2.1.2 Estimation for the doubling time of human cases

A preliminary evaluation based on the information about the incidence of the human cases can be done by estimating the increase of the number of case over time before any intervention. The only available information for that purpose is the times of symptom onset. A first approximation uses the assumption of an exponential increase in the number of case over time. The doubling time (T_d) is the time needed to double the number of cases during an epidemic. Doubling times are a measure of the rate of spread of disease and also indicate the magnitude of control efforts required to curtail spread.

Although this is not the same as the basic reproduction number, since the infectivity period of mosquitoes and humans aren't included, it provides a first estimate for the increase of the human cases of time. If r is the rate of increase then the number of cases will increase according to the function

$$N(t) = N_0 e^{rt} .$$

If N_0 is the initial number of cases at time $t = 0$ then the doubling time is

$$T_d = \frac{\ln 2}{r}$$

In the case of the CHIK epidemic the presentation of the number of cases under a semi logarithmic scale shows clearly a very good approximation of the initial increase as an exponential curve (thus the cases follow a straight line).

The best fit of the cumulative incidence function yields a value of 0.0869, which corresponds to a doubling time of the number of cases $T_d = \ln(2)/0.0869 = 7.97$ days, or about 8 days.

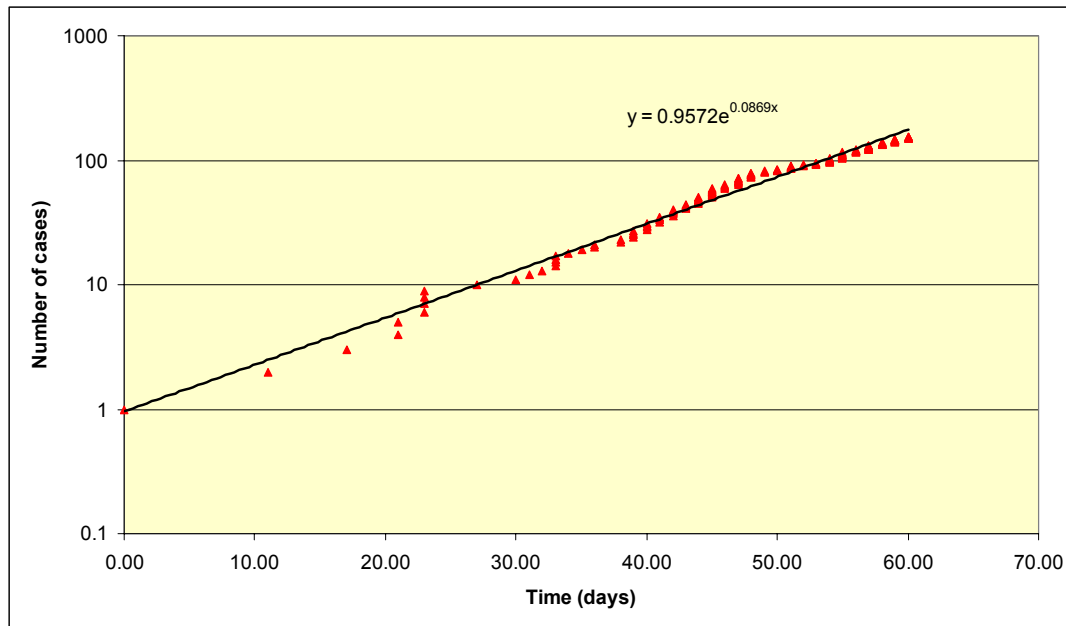


Figure 5 – Semilogarithmic axis representation of the epidemic cases

3 A simple deterministic compartmental model for chikungunya

3.1 Model assumptions

Deterministic compartmental models are a well-established means to reduce the biological complexity of epidemics to the key characteristics that are relevant to the infectious disease under consideration. For this exploratory examination of the Italian outbreak we made use of the framework developed by Bacaër (2007, p. 1082) for describing the outbreak in La Réunion, which uses the following compartments: $S(t)$, the susceptible human population; $E(t)$, the infected, but not infectious human population; $I(t)$, the infected human population; $R(t)$, the immune human population; $s(t)$, the susceptible vector population; $e(t)$, the infected, but not infectious vector population and $i(t)$, the infected vector population. The model reads:

$$\begin{aligned}\frac{ds}{dt} &= \lambda - \beta s(t) \frac{I(t)}{P} - \mu s(t) \\ \frac{de}{dt} &= \beta s(t) \frac{I(t)}{P} - (\gamma + \mu)e(t) \\ \frac{di}{dt} &= \gamma e(t) - \mu i(t) \\ \frac{dS}{dt} &= -\beta i(t) \frac{S(t)}{P} \\ \frac{dE}{dt} &= \beta i(t) \frac{S(t)}{P} - \delta E(t) \\ \frac{dI}{dt} &= \delta E(t) - \alpha I(t) \\ \frac{dR}{dt} &= \alpha I(t)\end{aligned}$$

In order to simulate an outbreak estimates for a number of parameters are needed that describe the development of the disease in humans and vectors and transition rates between the different compartments. The best and most recent information on most of these parameters relies on the outbreak in La Réunion. Therefore the following assumptions were made in the calculations:

- Incubation period in vectors, 7 days (rate $\gamma=0.14$)
- Life expectation of vectors, 30 days (rate $\mu=0.03$)
- Incubation period in humans, 4 days (rate $\delta=0.25$)
- Infectious period in humans, 7 days (rate $\alpha=0.14$)
- Period between bites, 4 days (rate $\beta'=0.25$)
- Probability of transmission from vector to human (q') 0.70 and from human to vector (q) 0.70. Note that the transmission rate $\beta=\beta'q'$ and $\beta=\beta'q$ for vector-to-human and human-to-vector transmission respectively. The values for the above probabilities are reasonable guesses.
- Initial susceptible population P , 3000 (approximately the population of the two villages)
- Assumed initial population of vectors 8000. This yields a value of $\lambda=264$ for the constant migration rate of mosquitoes before the epidemic starts and is obtained as Number of mosquitoes divided by the Life expectation ($Nm \mu$)

Since the population of mosquitoes is totally unknown we estimated its value based on the total number of symptomatic infected cases and the duration of the outbreak.

3.2 Model results

Figure 6 shows the cumulative number Y of confirmed cases (red) and the model prediction (blue), assuming that the mosquito control interventions on days 56 and 60 reduced the transmission probability 10-fold on each occasion.

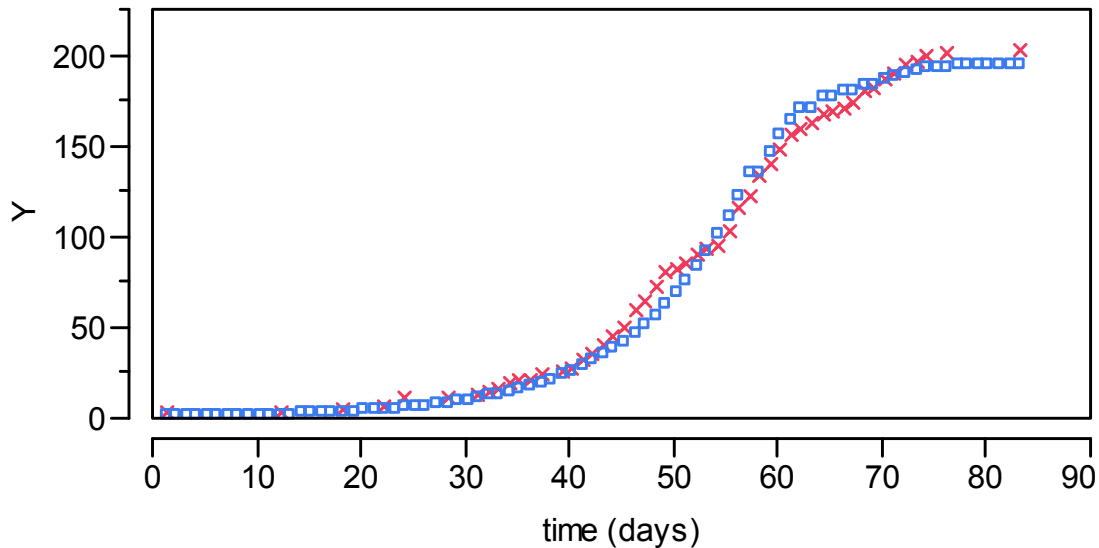


Figure 6 – Comparison of the real cases (red) with the model prediction (blue)

The simulated final number of cases (195) agrees well with the actual number (201), and the course of the epidemic. This calculation is based on an infectious period of 7 days. The doubling time of the human cases during the period before the intervention is 7.5 days, consistent with the value estimated in the previous chapter.

Decreasing the infective period leads to a much smaller number of cases.

The total number of mosquitoes that have humans as the preferred host is an important parameter and one of our major unknowns. The results in Figure 6 are based on the assumptions that the ratio of mosquitoes to humans is 2.8:1. Figure 7 show the cumulative number of cases if the mosquito population is halved.

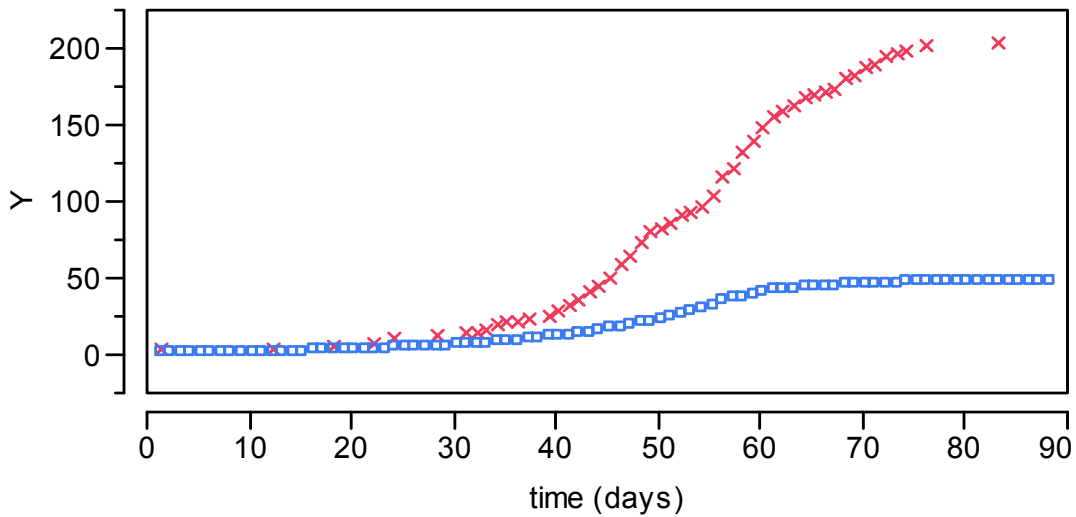


Figure 7 – Results of the simulation with mosquitoes population halved

The total number of cases is now only 48.

The following figure illustrates the effects of control measures. This simulation assumes that already on day 36 control measures were implemented. The final number of cases is then only 31.

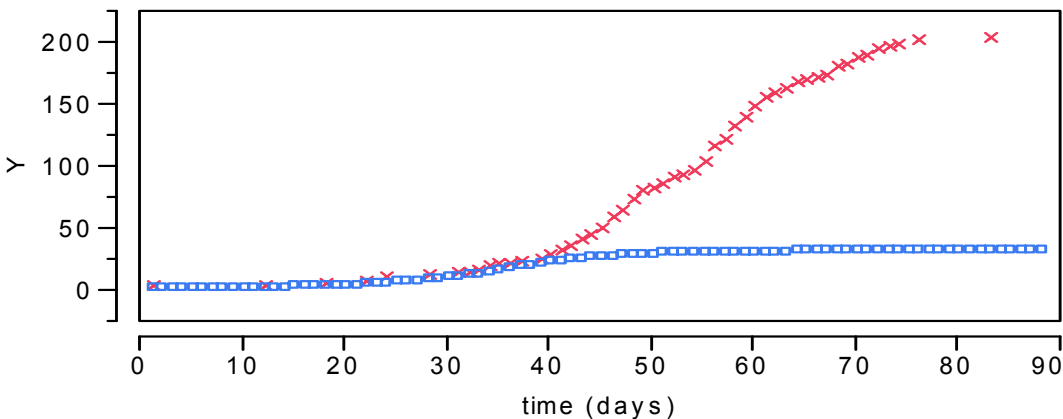


Figure 8 – Results of the simulation showing the effect when control measures are applied at day 36

These simulations show the importance of some essential parameters of which we have extremely limited information and of which improvement in relevant data generation should be pursued.

In the case that no intervention is applied (Fig. 6) the outbreak would have continued for more than a month with a dramatic increase in cases. In fact almost everybody would have been infected. This is unlikely though and shows the weakness of more reasonable predictions due to lack of data for crucial parameter values.

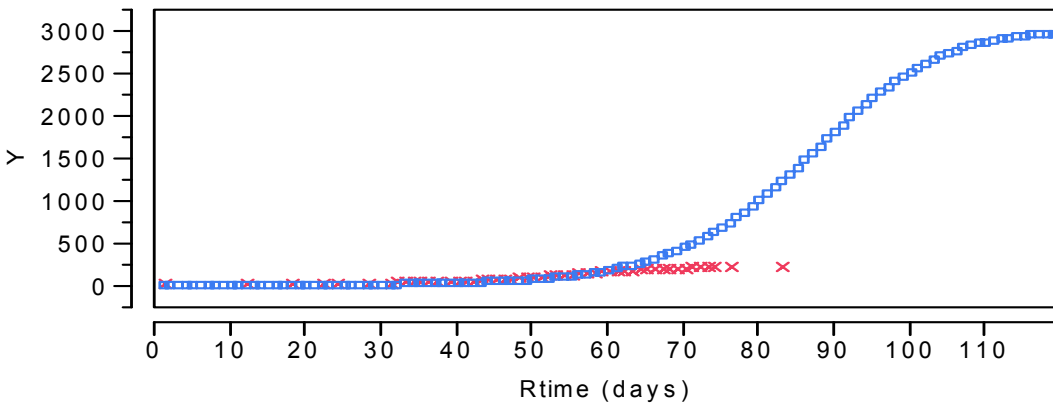


Figure 9 - Theoretical course of the outbreak without any intervention

These simulations show the importance of some essential parameters of which we have extremely limited information and of which improvement in relevant data generation should be pursued.

3.3 Concluding remarks

This exercise using a simple compartmental model shows that using parameter values derived from the La Réunion-outbreak does lead to an extremely good fit between the model output and the Italian outbreak data (Figure 6).

It should be pointed out though that in that case crucial unknown parameters such as the number of mosquitoes and the probability of transmission (and only those) were chosen in such a way that the model could fit the data based on the duration and the number of cases showing symptoms.

Thus the fit should be taken with caution. It is possible that the behaviour and physiological ecology of the vector differs between Italy and La Réunion which implies differences in the fixed parameters too.. Differences in temperature may, for example, affect the incubation period of the virus in the mosquito and other ecological factors may influence the life expectancy of the vector *Aedes albopictus*. Quantitative predictions also depend on the preference of the vector for outdoor vs. indoor feeding and its use of alternative hosts (dogs, cats, and cows). This shows that significant effort needs to be spent on collecting fundamental entomological information on the vector. Even if this information is still not available, modelling can help to identify sensitive parameters and thus perhaps assist in focusing monitoring efforts.

4 An agent-based approach

4.1 Introduction to Agent Based models

The contact structure in a population is important for the spread of an infectious disease. For diseases like influenza or smallpox, transmission is most likely when persons are at the same place at the same time. Homogeneous-mixing models, such as the compartmental model presented above, assume that all hosts have identical rates of disease-causing contacts. Agent-based models (ABM) are a type of model which has the ability to represent complex systems through the local interactions between individual entities in a spatially explicit environment. The behaviour of agents is determined by a set of individual and collective decision rules. This type of model can thus more easily represent heterogeneities in the mechanisms of disease propagation.

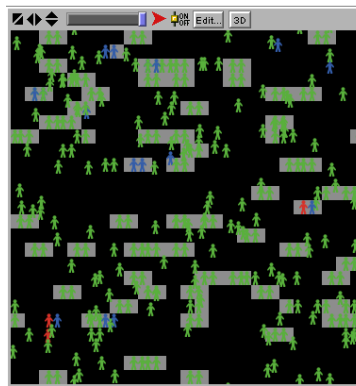


Figure 10. Agents move interact an environment and can have individual behaviour.

Agent-based approaches have been used in epidemiology to model the spread and control of infectious diseases in a spatially-explicit manner. Ferguson *et al.* (2006) present a large-scale, individual based model to estimate the consequences of control measures such as border and travel restrictions, school closures, case isolation, household quarantine and anti-viral treatment on an influenza pandemic in the US and the UK. For the case of smallpox, Halloran *et al.* (2002) present an individual-based, stochastic, spatially explicit and socially heterogeneous model to calculate the consequences of mass vs. targeted vaccination.

ABMs have so far not been used extensively to model the propagation of vector-borne diseases, so that their strength and weaknesses concerning this type of diseases are still not clear.

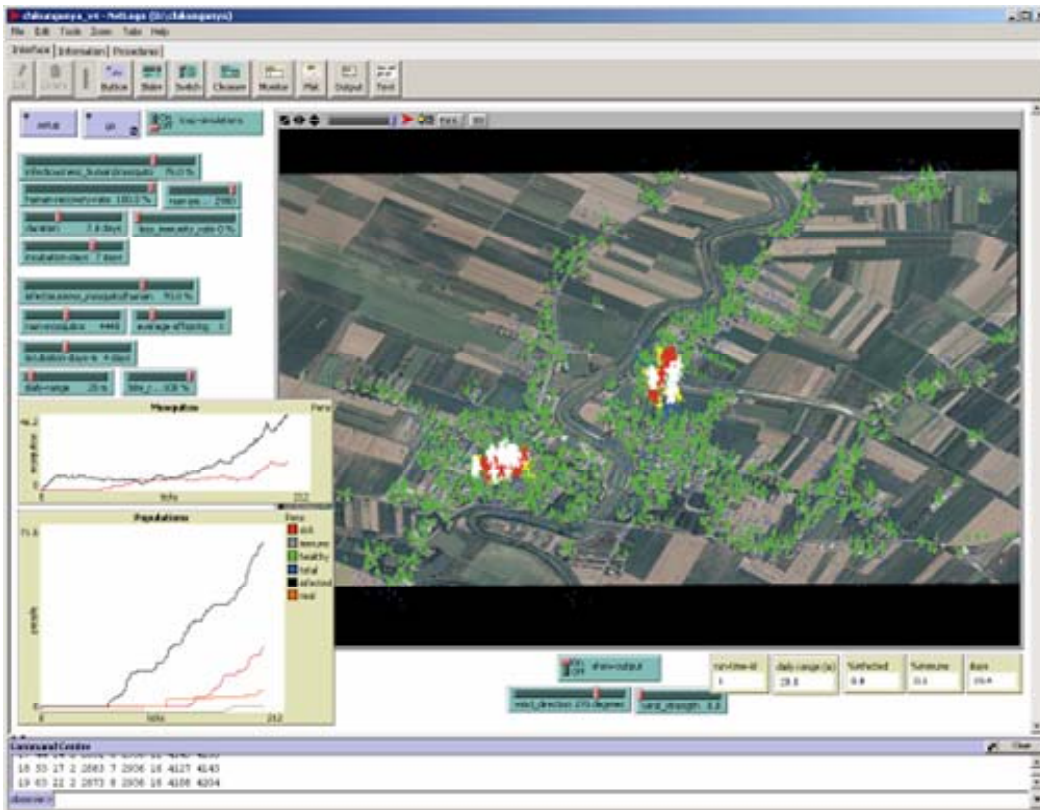


Figure 11. Spatially explicit modelling: for instance human agents are located in houses, while mosquitoes are more randomly located and follow Brownian motion, with a expected daily range. The initial infection can be put in the centre of a populated area or on the outskirts, with different behaviour of the disease propagation.

One important feature of ABMs is that they are stochastic. Every model run has different results because agents interact in a stochastic manner with each other. Unlike deterministic models, many model runs are necessary to understand the macro-behaviour of the modelled system.

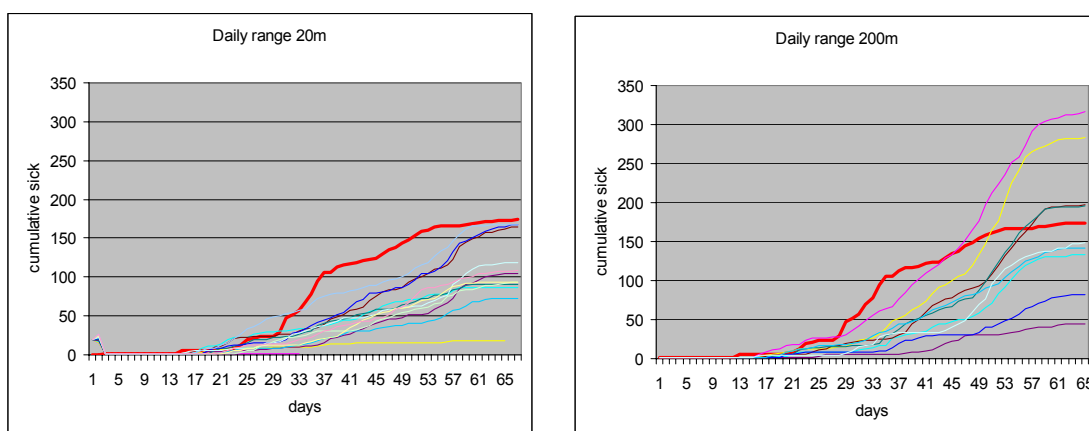


Figure 12. Influence of daily travel range of mosquitoes. While some runs with daily range of 200m yield less infected people than runs with daily range of 20m, this is not so on average.

Since many runs are needed to understand the macro-behaviour, this can also be used to determine parameter values following the Monte Carlo method. For instance, in each run one parameter value

can be varied. Comparing the model outcomes with known values, the best parameter value can be estimated.

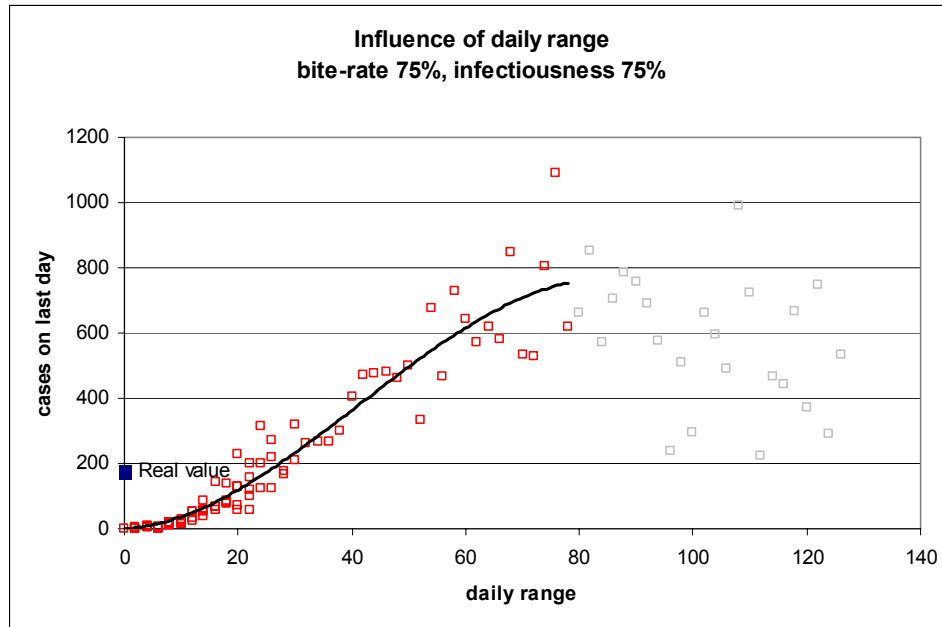


Figure 13. Optimal value for daily range is between 25 and 30 m. Note that when the daily range exceeds a given value (80m in this case) the outcome is influenced strongly by the size of the world and associated boundary effects.

The outcome of ABM models is determined by the contact model between agents.

Different ABMs model contact in different ways. However, since an agent can only be infected by having contact with another agent, the contact model is critical. Typically, ABMs first consider an arbitrary area (called contact area) within which agents *can* interact in each time step. Further, interaction of agents within the contact area is governed by statistical rules (such as “all mosquitoes bite all humans with 75% probability”). Finally, at each time step agents move whereby some agents move into another (neighbouring) contact area.

The contact rates, the bite probability and the movement variables need to be set-up properly because the same behaviour may be obtained by increasing one and decreasing the other parameter.

In fact, within each contact area, a deterministic model holds, similar to what was previously described by

$$R_0 = \frac{\beta^2 q q' p}{\alpha \mu P}$$

However, the vector and people populations in each cell change at each step by movement, resulting in an open world model instead of a closed world model (at the scale of a contact area). If contact areas increase in size, more agents are in the contact area and the number of possible interactions changes with l^2 , where l is the dimension of the contact area (radius of circle or side of square). In order to keep a constant number of interactions, the bite rate β needs to be modified. Therefore, equivalent models must have: $\beta \cdot l^2 = \text{constant}$. Similarly for models with different time steps, equivalent models have: $\beta / \Delta t = \text{constant}$.

ABMs are able to model a very important aspect of disease propagation: when a single infected individual (mosquito or person) is introduced in a healthy population, the susceptible population is

restricted to the neighbourhood of the infected individual. Indeed, initially only the agents in the same contact area as the infected agent are at risk.

Only when infected agents move outside of the initial contact area, other agents in neighbouring contact areas become susceptible. In other words, the susceptible population grows with time, based on the movement of mosquitoes and people. For instance, if an infected person lives in the centre of a village with radius 300m and mosquitoes have an average daily range of 30m it takes around 10 days for a person living at the edge of the village to be at risk.

4.2 Application to chikungunya

We use the chikungunya outbreak in the two villages of Castiglione di Cervia and Castiglione di Romagna for the following reasons:

- the greater cluster of cases was in these two locations
- the initiating case (index case) was here

The information we could use is as follows

- Population data from “Comuni Italiani”:
 - 1284 Castiglione di Cervia
 - 1660 Castiglione di Romagna
- High Resolution Image from Virtual Earth

In order to position the human agents we decided to use the satellite image and a technique developed at JRC to obtain the built-up area. Census data would have certainly be more accurate but they are not publicly available. In addition this procedure may be used also for areas where census data are not present at all.

In order to define the built-area in the affected area, a procedure for the calculation of a “built-up presence index” has been used (Pesaresi, 2007).

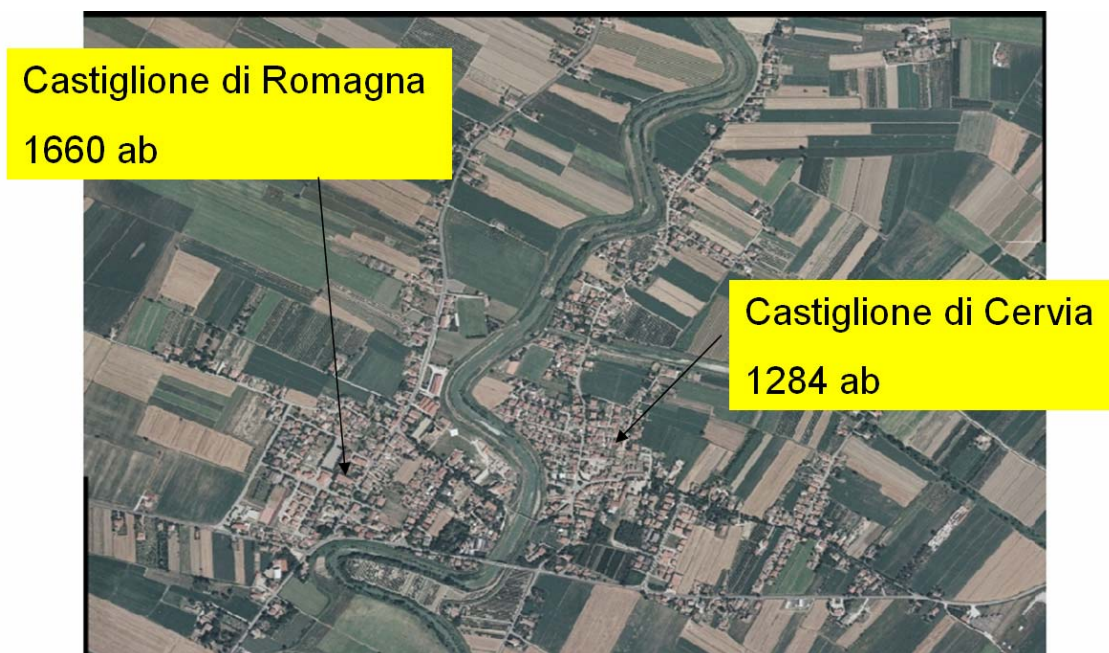


Figure 14. – The two villages in an aerial image from Virtual Earth

The methodology allowed deriving a dataset containing the geographical coordinates of each inhabitant of the two villages and which was respecting the actual population density and the actual overall population. This information was used as a basis for the agent based modelling. Details about the procedure are described in Appendix A.



Figure 15 – Population distribution in the two villages

Two agents have then to be defined: the mosquitoes and the humans.

4.2.1 Mosquito model

An arbitrary number of mosquitoes has been defined (3000); the mosquitoes are characterized by the following behaviour:

- Their initial position coincides, initially, with one of the humans, on a random basis.
- They can move in any direction within a maximum radius (we tested 200 m and 500 m)
- The probability of contact between a mosquito and a human is defined as follows: when the distance between the mosquitoes and one human is lower than 10 m, there is a probability of 10% of having a contact
- 3 states are possible for mosquitoes:
 - Susceptible
 - Infected (if a mosquito bites an infectious human can become infected in 75% of the cases)
 - Infectious, after 4 days of incubation period
- The average lifespan is 30 days²

² A lifespan of 30 days is based on laboratory studies. It is safe to assume that the life expectancy in the wild is much shorter

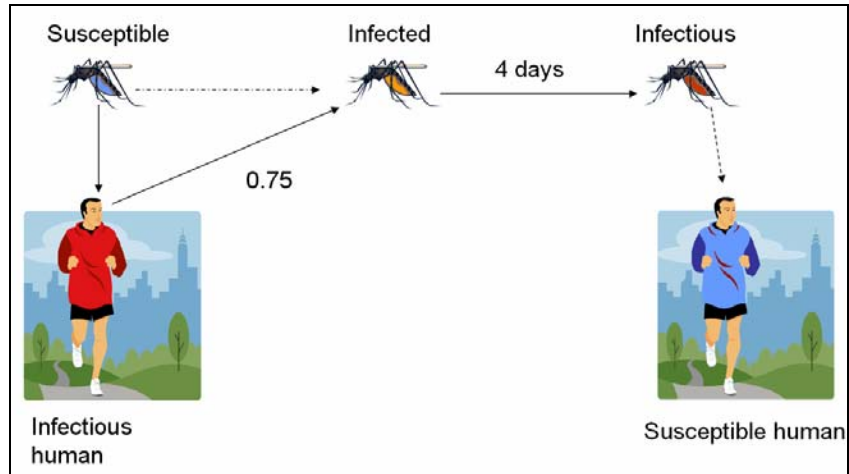


Figure 16 – Scheme for mosquitoes states

4.2.2 Human model

A number of humans that respect the population of the two villages is distributed over the built-up area (2944 individuals). They are characterized by the following behaviour:

- Humans do not move. It is intended that since most of the bites are occurring in the night, it is “as if” they don’t move. In reality the literature on the biology of *Aedes albopictus* states that the mosquito is mostly active during the day. However the human movements would cause a change in the macro-spreading (transfer to other cities) but not at village level.
- The probability of contact between a mosquito and a human is defined as follows: when the distance between the mosquitoes and one human is lower than 10 mt, there is a probability of 10% of having a contact
- 4 states are possible for humans:
 - Susceptible
 - Infected (if an infected mosquito bites a susceptible human, the human can become infected in 75% of the cases)
 - Infectious, after 7 days of incubation period
 - Immune, after 14 days from the bite

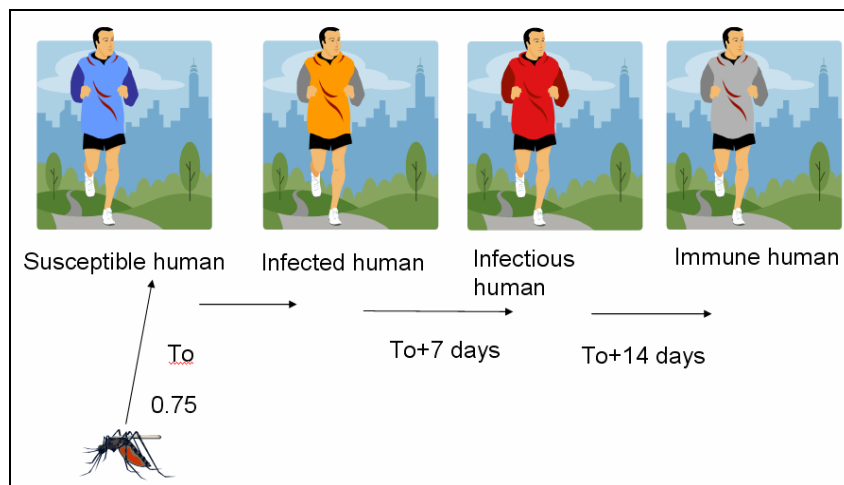


Figure 17 – Scheme for human states

4.2.3 Initial condition

The initial condition is modelled as one infected human (at the beginning of his infectious period) located in a certain position in the village. We tested 4 initial locations in order to check which is the effect of the index case location, when located in areas at different population density.

Positions 1 and 2 (see Figure 18) are characterised by a local medium population density in Castiglione di Cervia (1010 and 1169 persons per km²) and the difference is the distance from the river and thus from Castiglione di Romagna (position 1 closer to Castiglione di Romagna). Position 3 is in a low density area (358 persons per km²) and position 4 in a high density area (2800 persons per km²). It should be noted that the position (1) was the one closest to the real case.

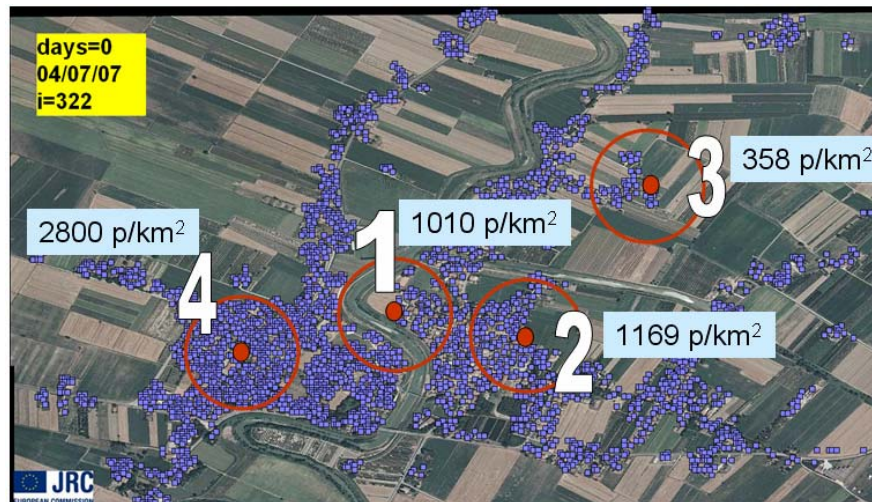


Figure 18 – Initial position of the index case

The calculations were performed by letting time run and having the mosquitoes move and interact according to their rules. The results are presented in terms of location of affected humans (pink if infected, red if infectious, gray if immune) and time behaviours. For all the cases 4 calculations have been done because each one, due to the random nature of the process, behaves differently.

Time 0 is represented by 23/6/2007, when the index case visited his cousin in Castiglione di Cervia, during his second fever episode. The cousin was the second case and became ill after about 12 days.

4.2.4 Actual cases

4.3 Calculations results

4.3.1 Initial case in position (1)

Starting from position 1 the trend is the one indicated in Figure 19. It should be noted that the results from an agent based model are not depicted by a single curve. In the figure the red curve represents the actual cases while all the other curves have been obtained starting from the same initial condition. Due to the random process in some cases (the upper curve) the infection reaches the village on the other side of the river and the number of cases is very high, in other cases (the lower curve) the infections remain mostly in the village of Castiglione di Cervia. The actual case is in the middle and is well represented by a couple of our calculations. It is interesting to note the same type of start/stop of the cases in coincidence with the various generations of mosquitoes.

It should be noted that the cases can be compared only until the 56th day, when the disinfection starts. The disinfection has not been accounted for in these calculations. It could be possible to repeat the calculations considering the disinfection, as done for the global case, but it would be necessary to guess an efficiency rate, which at the moment is not well known.

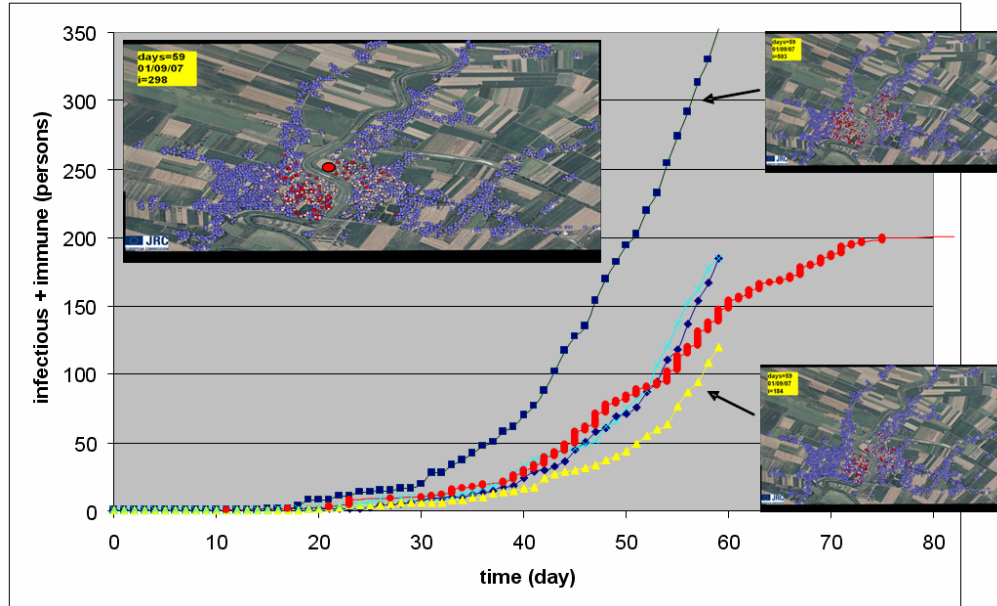


Figure 19 – Results of the agent base model starting from position (1)

4.3.2 Initial case in position (2)

The results of this case are very close to the previous case with lower values for all the curves because the initial position is further from the river and thus from the other village at higher population density.

4.3.3 Initial case in position (3)

The results of the calculations starting from position (3), where a very low population density is present show that the number of cases is much lower than in the previous simulations. It should be noted that in some of the simulations no secondary case have been created because the infection did not propagate and extend.

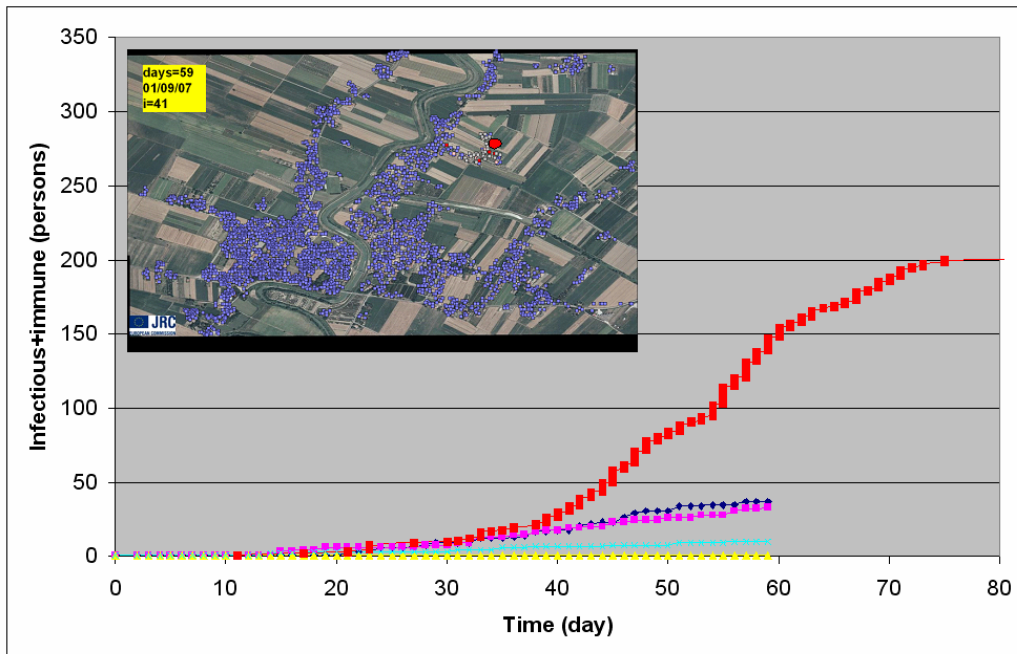


Figure 20 – Calculations results starting from the north position

4.3.4 Initial case in position (4)

The results of the calculations starting from position (4), a location with a higher population density, show that the number of cases is much higher than in the previous ones. In this case it is even more evident the 3 generations of mosquitoes which cause quite evident periods of cases increase along the time.

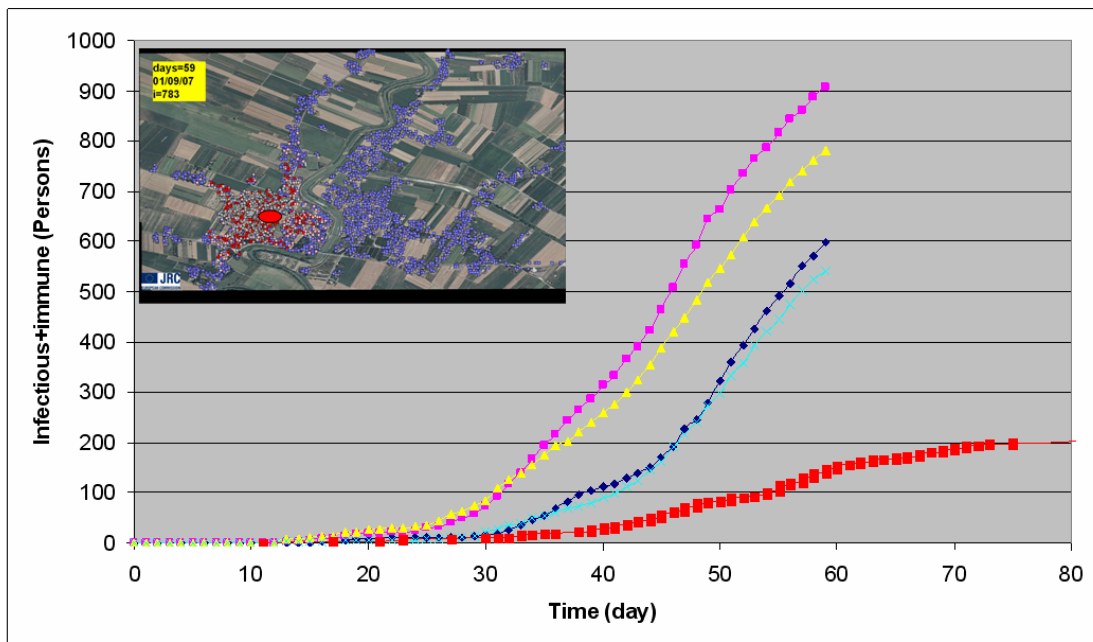


Figure 21 – Calculations results starting from the higher density position (4)

4.3.5 Effective population concept

When the epidemic starts (i.e. the index case in a certain position), the number of susceptible persons are not the whole village but only the persons that are in the area of activity of the mosquitoes (i.e. 200 m circle in the assumption of constant radius migration).

As the infection progresses the number of susceptible persons increase because the radius of the more distant infected person increase and allows other mosquitoes to infect other persons.

Fig. 16 shows the number of persons involved in the maximum radius which is starting from 200 m and ends with about 600 m from the initial case. At 60 days the persons inside this radius is only 1200 and not 2850 as the total number of persons. At the beginning only 153 persons are susceptible which are the ones included in the 200 m initial radius from the index case.

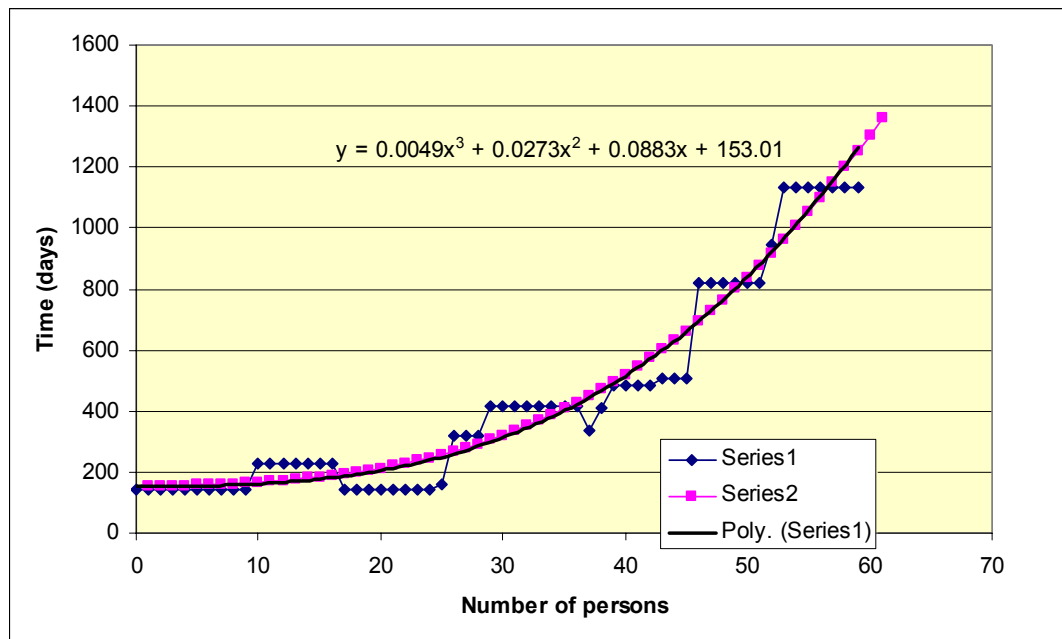


Figure 22 – Number of persons involved in a case

This could be reflected also in the global model and consider a variable number of susceptible persons instead of a fixed number. In such a way the global model could also take into account the spatial propagation.

4.4 Discussion of the results

The results of the agent based simulations lead to the following conclusions:

- Although the number of cases tested in this simulation is rather limited, they show a great variability even starting from the same initial condition (20-30 % of standard deviation)
 - This result may be due to the low assumed number of mosquitoes (3000). Maybe a much higher number of mosquitoes (and a corresponding lower bite rate) would have led to a more reduced variability.
 - It could however that the variability is inherent in the stochastic nature of the process
- In several cases no secondary case is generated, in particular when population density is very low or when the mosquitoes density is very low
 - It would be interesting to show what happens in the other locations where there were some limited clusters, such as Cervia
- The calculated trend show a clear relationship with the population density
 - This relation is valid for the assumed mosquitoes density; an increase or a decrease of the mosquitoes population would have a similar effect on the results, as already shown in 3.3.

	Cast Rav	EAST	WEST	NORTH
Pop. Density (n/km ²)	2800	1010	1169	358
Case 1	597	185	198	37
Case 2	907	353	228	33
Case 3	781	120	162	1
Case 4	532	377	294	10
Average	704	259	221	20
Standard Dev	148	126	56	18

Table 4 – Agents based modeling results

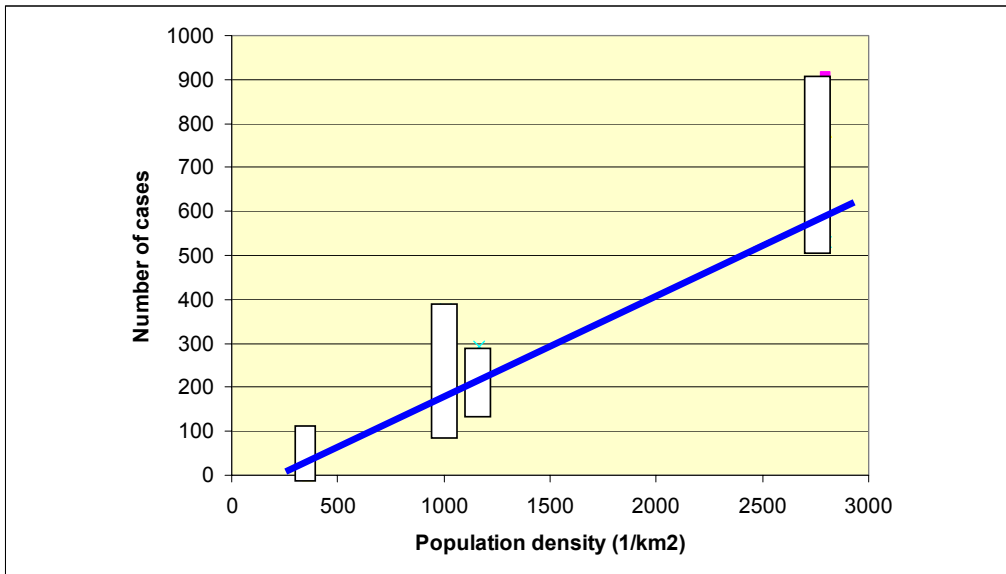


Figure 23 – Relation between population density and Number of cases

5 Discussion on the two approaches

5.1 *Complementarity of the two approaches*

(May I note some of the problems I see with ABMs?)

- ABMs rely on assumptions on individual behaviour that are not always easy to measure or validate; this applies to humans and mosquitoes
 - it has to be decided which behaviours are included and which are left out
 - there are computational and memory limitations to simulate outbreaks with more susceptible persons
- The ABM are however able to give information related to the time and special distribution, whereas the compartmental models assume an homogeneous distribution of the cases.

Compartmental models are more austere in their assumptions and give precise, numerical answers. The low number of assumptions and the uncertainty of parameter estimates, however, only allow the exploration of fundamental, abstract processes. In order to better capture the complexity of acute epidemic threats, additional structure needs to be included, for instance stochasticity, heterogeneity and spatial structure.

5.2 *Simulation of a large city not possible with an ABM*

The simulation of a huge amount of humans (or vectors) makes the system difficult to manage and time consuming, thus for this type of evaluations global models are more easy to use. However there are examples and studies (University of Trento) of simulation of the whole Italian population with ABM; also in USA they prepared a large computational model using several parallel processors. Alternatively it could be possible to reduce the granularity: agents could be defined differently, with different behaviour. (e.g as blocks or groups of people)

5.3 *How is it possible to use the global method in real time modelling*

The real time modelling of an epidemic outbreak uses a 2-step method: fast data collection and estimation of parameters and then the use of these parameter estimates in predictive models. The deterministic compartmental model presented here could be used for such a purpose. However, real-time modelling of a vector-borne disease faces some additional challenges. Any predictive model (if compartmental or agent-based) relies on parameters that describe the biology of the vector and these parameters have to be known beforehand. The incubation period of the virus in the vector or the lifespan of the vector cannot be measured in real-time.

6 Conclusions

An analysis of the chikungunya outbreak which affected the region of Emilia Romagna, in Italy has been performed. The analysis was conducted using 3 approaches: a descriptive statistical method trying to identify common clinical characteristics among the cases; a deterministic compartmental modelling method for an almost closed population case; an agents based model to simulate the geographical progression of the epidemic.

The main results are here reported for each of the approaches.

The descriptive statistics of disease symptoms shows that the cases in Italy are similar to the symptoms reported for earlier outbreaks (e.g. Borgherini et al. 2007). The clinical features of chikungunya fever are still poorly documented and such a documentation is important in order to facilitate rapid diagnosis. It was possible to identify and evaluate the doubling time (about 8 days) for the initial part of the epidemic before intervention. Unfortunately the data we received aren't final. New confirmed positive and negative case appeared later on. Thus the numbers of cases we used refer to a period of 90 days after the initial case appeared and should be regarded as preliminary. The calculation of R_0 cannot be done through the classic formula but has to be provided through other methods in that case statistical methods (work in progress).

The deterministic compartmental modelling approach shows that based on the estimates for the chikungunya outbreak in La Réunion a description of the outbreak in Italy is feasible, but some unknowns still remain. Some selected numerical calculations show how the results depend on different parameter values as well as how earlier intervention would change the dynamics of transmission.

The Agents Base Model calculations show that it is possible to represent the spatial component distribution of the cases. However also in this case some important assumptions have been done in the model parameters which are not well known (i.e. biting rate or contact area). It is however possible to find a parameter combination which describes extremely well both the temporal and the spatial components. In addition the Agents Base Model results are stochastic in nature, thus a single calculation could not be significative of the real average behaviour and the analysis with ABM models can be very much time consuming.

Acknowledgments

We thank the Department of Prevention (Off. V) of the Italian Ministry of Health and in particular Dott. MG Pompa, the National Institute for Public Health (Istituto Superiore di Sanità, Dip. Mal. Parassitarie e Virologiche), MIPI, CNESPS and the Region of Emilia Romagna, and in particular Dott. C. Po, for providing us their cooperation and the provision of the outbreak related data..

References

- Anderson, R.M. & May, R.M. (1991) *Infectious Diseases of Humans. Dynamics and Control*. Oxford: Oxford University Press
- Bacaër, N. (2007) Approximation of the basic reproduction number R_0 for vector-borne diseases with a periodic vector population. *Bull. Math. Biol.* **69**: 1067-1091
- Borgherini, G, Poubeau, P, Staikowsky, F, Lory, M, Le Moullec, N, Becquart, JP et al. (2007) Outbreak of chikungunya on Reunion Island: Early clinical and laboratory features in 157 adult patients. *Clin. Infect. Dis.* **44**: 1401-1407
- Diallo, M., Thonnon, J., Traore-Lamizana, M. & Fontenille, N. (1999) Vectors of Chikungunya virus in Senegal: current data and transmission cycles. *Am. J. Trop. Med. Hyg.* **60**: 281-286
- Ferguson, N.M., Cummings, D.A.T., Fraser, C., Cajka, J.C., Cooley, P.C. & Burke, D.S. (2006) Strategies for mitigating an influenza pandemic. *Nature* **442**: 448-452
- Fontenille, N., Failloux, A.B. & Romi, R. (2007) Should we expect chikungunya and dengue in Southern Europe? Pp. 169-184 in: *Emerging Pests and Vector-Borne Diseases in Europe*. (W. Takken & B.G.J. Knols, eds.), Wageningen Academic Publishers
- Halloran, M.E., Longini, I.M., Nizam, A. & Yang, Y. (2002) Containing bioterrorist smallpox. *Science* **298**: 1428-1432
- MacDonald, G. (1952) The analysis of equilibrium of malaria. *Trop. Dis. Bull.* **49**: 813-829
- Pesaresi, M., Gerhardinger, A. & Kayitakire, F. (2007) Monitoring settlement dynamics by anisotropic textural analysis by panchromatic VHR data. Proceeding of the Urban Remote Sensing Joint Event 2007, Paris, France, 11-13 April, 2007
- Pialoux, G., Gauzere, G.A., Jauregulberry, S. & Strobel, M. (2007) Chikungunya, an epidemic arbovirus. *Lancet Infect. Dis.* **7**: 319-327
- Ross, R.W. (1956) The Newala epidemic, III. The virus: isolation, pathogenic properties and relationship to the epidemic. *J. Hyg.* **54**: 177-191

Appendix A – Geolocalization procedure

The list of 300 cases, as provided by the Ministry of Health, was geo-coded using an address matching algorithm into a GIS environment (Esri ArcMap) followed by a manual check (fig. 1); at the end of the procedure two new attributes was added to each geolocated record: the latitude and longitude. The reference data used to match the addresses was the Teleatlas MultiNet dataset.

At the end of the procedure 298 out of 300 cases has been geolocated (fig. 2).

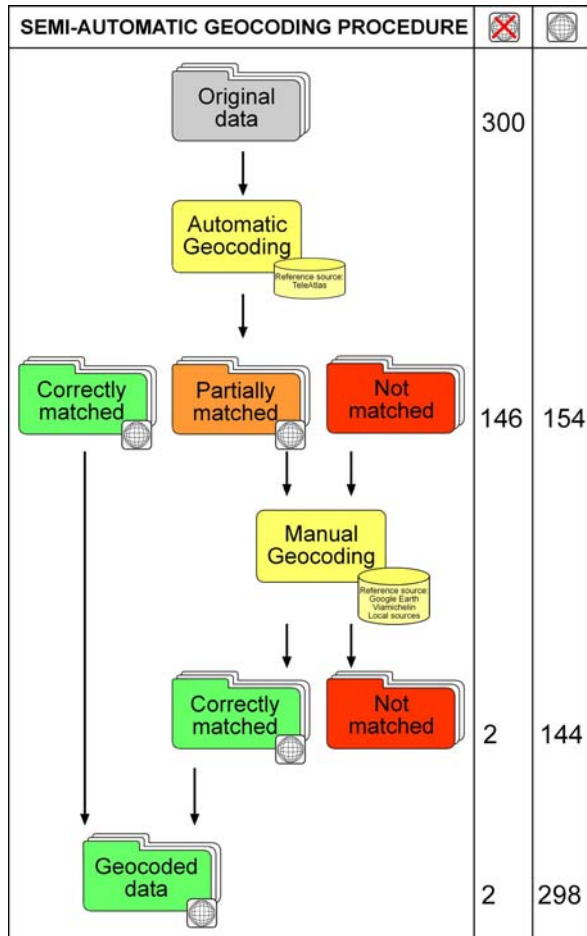


Figure 24 - Semi-automatic geocoding procedure flow-chart

Starting from the original address list a preliminary homogenization control was performed in order to correct eventual spelling errors and to rewrite all the addresses to match the adopted address style.

After this, the data were automatically geo-coded using an address matching algorithm. This procedure located each address comparing it with the reference data, in our case the TeleAtlas Multinet dataset, and added to the original table four new fields:

- the coordinates (lat/lon in decimal degrees),
- the resulting code of the matching operation (EXACT, CANDIDATES, NOMATCH),
- and an estimation of the accuracy (0-100%).

This process produced 154 exactly matched records.

The remaining 146 records underwent a secondary manual process aimed to:

- check the results for records that were matched but still having a candidates status;

- geocode the records not matched or the candidates wrongly matched using different data sources and services.
-

Out of the 146 record not exactly matched during the first automatic process, 74 of the candidates were correctly matched while 70 were matched manually using different data sources and services such as Google Earth³, ViaMichelin⁴ maps and navigation services and local sources.

At the end of the process only 2 records were not matched because of the lack of address information in one case and the impossibility to find the name street of the second one.

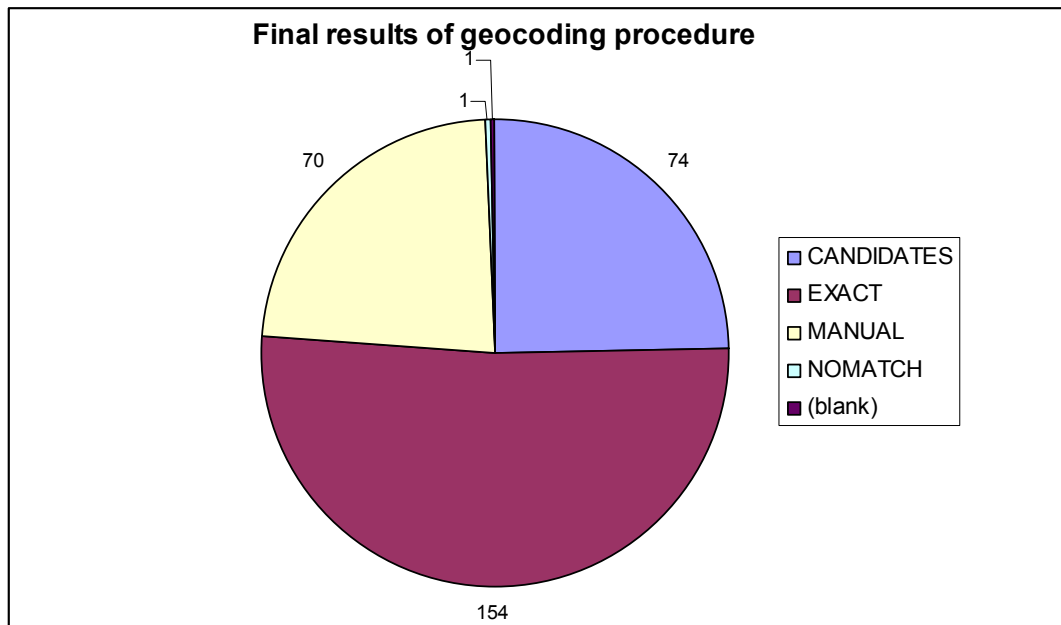


Figure 25 – Final results of the geo-coding process.

³ <http://earth.google.com/intl/it/>

⁴ <http://www.viamichelin.it>

Appendix B – Description of the methodology used for retrieve population density in the affected area

In order to define the built-area in the affected area, a procedure for the calculation of a “built-up presence index” has been used (Pesaresi, 2007⁵). The index is based on fuzzy rule-based composition of anisotropic textural measures derived from the satellite data by the gray-level co-occurrence matrix (GLCM). Even if others methodologies could be used in order to retrieve built-up areas from the area in exam (digitalization, color pattern technique), the textural analysis is a very generic technique that has proved to work excellent in different conditions.

As input, a snapshot from MS Virtual Earth, that covers the whole Italy with ortofoto from airborne (2002), has been done. The places cover an area of about 3x2 kilometers with a total surface of about 6 square kilometers.

As a result of the “built-up presence index” procedure, a polygon vector representing the built-up area has been obtained.

The input image as well the polygon vector of built-up areas has been georeferenced.



Figure 26: Raw image from MS Virtual Earth



Figure 27: The georeferenced image with the TeleAtlas street network

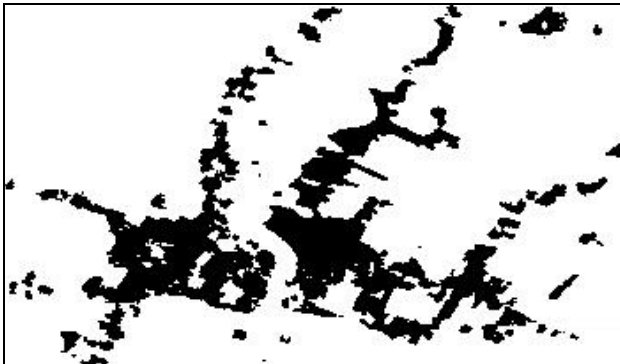


Figure 28: The built-up area obtained from the “built-up index”

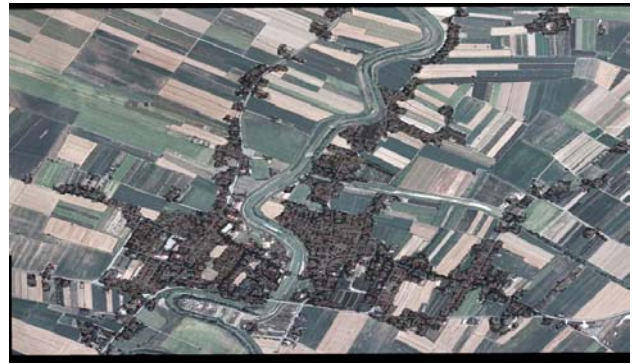


Figure 29: The airborne image of the affected area with the built-up area obtained from the “built-up index”

In order to calculate the population density in the area, the vector “built-up” layer has been converted to two raster of 10 meters of resolution, one for each place. Then for each output raster the total population of each place was divided by the total built-up area converted in 10 meters, and the value obtained was assigned to each pixel as population density.

⁵ M. Pesaresi A. Gerhardinger. F. Kayitakire, 2007: Monitoring settlement dynamics by anisotropic textural analysis by panchromatic VHR data. Proceeding of the Urban Remote Sensing Joint Event 2007, Paris, France, 11-13 April, 2007.

European Commission

EUR 23065 EN – Joint Research Centre – Institute for the Protection and Security of the Citizen

Title: Modelling the 2007 Chikungunya Outbreak in Italy

Author(s): A. Annunziato, N.I. Stilianakis, T. de Groeve, L. Vernaccini, S. Gadenz

Luxembourg: Office for Official Publications of the European Communities

2007 – 36 pp. – 21 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1018-5593

Abstract

The report describes the modelling analysis performed by JRC on request of the DG-SANCO related to the Chikungunya Outbreak occurred in Italy in the summer of 2007. The analysis reports the calculations performed using global models and agent based model

How to obtain EU publications

Our priced publications are available from EU Bookshop (<http://bookshop.europa.eu>), where you can place an order with the sales agent of your choice.

The Publications Office has a worldwide network of sales agents. You can obtain their contact details by sending a fax to (352) 29 29-42758.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

