Computational Biology and Chemistry 33 (2009) 176–180

Contents lists available at ScienceDirect

Computational Biology and Chemistry

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

Brief Communication

A model of influenza virus spread as a function of temperature and humidity

Tomasz Żuk*, Franciszek Rakowski, Jan P. Radomski

Interdisciplinary Center for Mathematical and Computational Modeling, University of Warsaw, Pawińskiego 5A, Bldg D, 02-106 Warsaw, Poland

ARTICLE INFO

Article history:
Received 19 September 2008
Received in revised form 31 October 2008
Accepted 26 December 2008

Keywords:
Influenza
Influenza epidemiology
Influenza infectivity
Influenza transmission

ABSTRACT

The influence that atmospheric conditions might have on the efficiency of the spread of influenza virus is important for epidemiological and evolutionary research. However, it has not been satisfactorily recognized and quantified so far. Here we provide a statistical model of influenza transmission between individuals. It has been derived from the results of recent experiments, which involved infecting guinea pigs with influenza at various temperatures and relative air humidity levels. The wide range of transmission rates in those experiments reflects the ensemble-independent phenomena. The correlation between most of our simulations and the experimental results is satisfactory. For several different conditions, we obtained transmissibility values which seem to be sufficiently accurate to provide partial input for an intended large-scale epidemiological study in the near future.

1. Introduction

Among the viruses that are spread efficiently by air, the influenza A virus causes one of the highest worldwide morbidity and mortality rates. However, there are still some factors related to its spread that have not been thoroughly examined and understood. In particular, the influence of weather conditions, such as air temperature and relative humidity, on the between-host transmission of this virus is not understood clearly. Thus far, the underlying reasons for its predominantly wintertime spread, significant for the understanding of influenza epidemiology and evolution, are still unexplained. Nonetheless, the seasonality of influenza epidemics is well characterized—in temperate regions influenza epidemics recur with marked regularity: in the northern hemisphere the influenza season spans November to March, while in the southern hemisphere epidemics last from May until September. Many theories have been proposed to explain this seasonal variation (Lofgren et al., 2007).

Recently, some results have provided direct experimental evidence of the major role of weather conditions in the dynamics of influenza transmission. Lowen et al. (2007), using the guinea pig as a model host, have shown that the efficiency of airborne influenza spreading depends upon both ambient relative humidity and temperature, and that both cold and dry conditions strongly favor transmission. Aerosol-related phenomena are most likely the major contributing factors here. Some studies of the possible effect they might have on virus spread have been carried out, although they are neither numerous nor conclusive. Wang et al. (2007) studied the case of SARS; Tellier (2006) reported on influenza A aerosol spread. Lai and Cheng (2007) modeled expiratory droplets dispersion transport using Eulerian approach. Quantification of the routes of influenza transmission was attempted by Atkinson and Wein (2008).

Here we present an heuristic model of influenza transmission, which includes some environmental parameters. It is based on the experimental results of Lowen et al. (2007), involving guinea pigs.

2. Experimental Results

The underlying guinea pig experiments were conducted using eight animals per trial: four infected and four healthy but susceptible to infection. Individual pigs were placed together in a chamber and arranged pairwise on four shelves such that each shelf contained one healthy and one infected individual. Thus there were two vertical columns—one consisting of the healthy and the other, of the infected guinea pigs. The air flowed horizontally from the infected towards the healthy animals to maximize the infection rate. In order to control infectivity, the virus concentration in nasal wash of each guinea pig was measured every 2 days. The goal was to discover how many susceptible animals would become infected, and after how many days, under controlled air temperature and humidity conditions.

The experiments were performed at five different relative humidity values (20%, 35%, 50%, 65%, and 80%) and three different temperatures (30 °C, 20 °C, and 5 °C). At 30 °C no infections occurred. For both 20 °C and 5 °C, low relative humidity of 20–35% was most favorable, leading to nearly 100% transmission. At 20 °C

* Corresponding author at: ICM, University of Warsaw, Żwirki Wigury 93, 02-089 Warszawa, Poland. Tel.: +48 22 55 40 800; fax: +48 22 55 40 801.
E-mail addresses: tzuk@icm.edu.pl (T. Żuk), rakowski@icm.edu.pl (F. Rakowski), janr@icm.edu.pl (J.P. Radomski).

1476-9271/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiolchem.2008.12.001

the transmission rapidly fell at 50% humidity, rose again to quite high levels at 65%, and was completely blocked at 80%. Such behavior probably reflects the virus stability in aerosols. At 5 °C the transmission was still very efficient at 50% humidity but dropped to 0.5 for higher humidity values. For infected animals housed at 5 °C, the duration of peak shedding was approximately 40 h longer than that of animals housed at 20 °C. This increased shedding probably accounts for the generally enhanced transmission seen at 5 °C.

3. Methodology

3.1. Model's Assumptions

We have formulated a discrete stochastic model based on the probability of infection. To estimate this probability, we apply a simple generalized linear model (GLM), with $f(p) = \ln(1-p)$ as the link function. This is a standard approach that has been successfully adopted in various related contexts for modeling influenza (Ferguson et al., 2005; Stegeman et al., 2004). The model starts from day 0, when all non pre-infected animals are healthy. For each consecutive day, we calculate the probabilities that each individual, among those not yet infected, will be infected on that day, and then we randomly check whether any of the possible infections actually happens. The infection history is subsequently updated, and the algorithm moves to the next day.

We assume the four following conditions:

1. The probability that the ith individual will be infected on a particular day[1] is

$$p_i = 1 - \exp\left(-\gamma(T,H)\sum_j \alpha_{ij}\beta_j\right) \quad (1)$$

where β_j is a measure of virus concentration in the respiratory tract of the jth individual (decimal logarithm of the number of PFU in one milliliter of its nasal wash). The summation goes over individuals placed on the same and on adjacent shelves. Spatial coefficients α_{ij} are related to the probabilities that an aerosol drop, shed by the jth animal, will get to the ith guinea pig's cage. The coefficient $\gamma(T,H)$ reflects the dependence of the virus transmission rate on the air temperature (T) and relative humidity (H).
2. The course of influenza (virus concentration in nasal wash on the days subsequent to infection) depends only on the ambient temperature. Therefore, for each individual that became infected via air during the course of each particular experiment, the concentrations were assigned values reflecting the average case for that temperature. Table 1 shows the mean concentrations measured on particular days after infection, separately for each temperature. Functions comprising two linear pieces appear to be a reasonable guess, as shown in Fig. 1. This may be interpreted as a two-phase infection course: first, the virus concentration grows exponentially (the scale is logarithmic), and subsequently (mostly as a result of the immune system activation) it exponentially falls. The transition between phases 1 and 2 is always sharp.
3. A detectable virus presence in the respiratory tract begins 2 days after the infection. Such an assumption is drawn from the fact that the virus was not detectable in nasal wash earlier than the third day of experiment. A lack of infections during the first day, when the virus concentrations in pre-infected individuals is supposed to be high, is of little probability. This is consistent with the major biological and clinical experience, which shows that the influenza proper begins rapidly after 2–3 days of a rather asymptomatic and non-infectious incubation period (Carrat et al., 2008; Collier and Oxford, 1993). Accordingly, we shifted the assumed average infection course such that $\beta = 0$ on the first day. The final concentration values are shown in Table 2.
4. The infection course in pre-infected individuals also consists of two linear (on a logarithmic scale) phases. Concentration values for the descending phase come from the measurement[2] (Lowen et al., 2007). Due to the lack of the respective experimental data, for the ascending phase we assumed the slope to be the same as for animals infected via air. The virus concentration values for those individuals are also shown in Table 2.

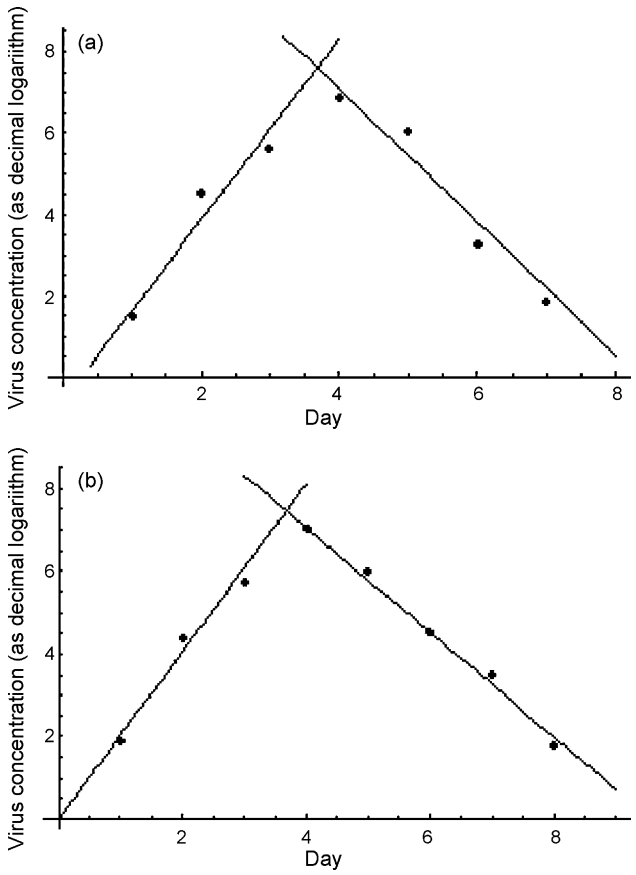

Fig. 1. The two piecewise linear approximations of the infection course. For 20 °C the best fit is $\beta(t) = 2.23t - 0.61$ for the growing part and $-1.63t + 13.62$ for the falling; for 5 °C it is $2.03t - 0.01$ and $-1.27t + 12.13$ respectively. We assume $\beta = 0$ if the amount of virus is undetectable or zero.

3.2. Propagation of Infection

Assuming values for $\beta(t)$, α_{ij}, and γ, it is possible to compute the probability of each particular healthy individual becoming infected on the first day of the experiment (Eq. (1)). And from the probabilities determined for days 1 to t, we can calculate the value for day $t+1$ for each guinea pig in the following manner: for each sequence of days $\tau = (\tau_1, \ldots, \tau_n)$, where $\tau_k \in \{0, 1, \ldots, t\}$, and $n = 4$ is the number of individuals to infect, let us take:

[1] Provided that it had not been infected previously.

[2] We fitted a linear function for these values. However, they were originally of a nearly linear nature.

Table 1
The average virus concentration in nasal wash of individual guinea pigs infected via air.

	Day								
	1	2	3	4	5	6	7	8	9
Concentration, $T = 20\,^{\circ}C$	1.52	4.52	5.63	6.89	6.06	3.26	1.88	–	–
Std. deviation square	0.15	0.48	0.63	0.37	0.67	1.92	2.72	–	–
Concentration, $T = 5\,^{\circ}C$	1.91	4.40	5.72	7.02	6.01	4.54	3.52	1.77	–
Std. deviation square	0.40	0.59	1.26	0.32	2.38	1.44	3.18	2.34	–

For each individual, day 1 is the day when a measurable virus quantity first appeared in its respiratory tract.

(a) the probability p_τ that, for each j from 1 to n, the jth individual became infected on day τ_j (or has not yet been infected if $\tau_j = 0$) and
(b) the sum S_τ that occurs in the exponent in Eq. (1) if (a) is the case.

The probability $p_i(t+1)$ that the ith individual will become infected on day $t+1$ may then be considered as the expected value of a random variable which takes values

$$\varphi_\tau = P_i(t)(1 - \mathrm{e}^{-\gamma S_\tau})$$

with probabilities p_τ. Hence it equals

$$p_i(t+1) = P_i(t)\left(1 - \sum_\tau p_\tau\, \mathrm{e}^{-\gamma S_\tau}\right). \tag{2}$$

$P_i(t)$ is the probability that the ith individual has not been infected previously, i.e., up to day t (if it has, infection cannot happen anymore). Of course,

$$P_i(t+1) = (1 - p_i(t+1))P_i(t).$$

The above procedure establishes an iterative method for finding the infection probabilities on each consecutive day. The results are consistent with what was obtained from a number of Monte Carlo simulations, run with the same parameter values.

The coefficients α_{ij} indicate the fraction of the virus shed by the individual j that contributes to infecting the individual i. We assume the following:

(a) $\alpha_{ij} = 1 - 2\alpha_1$ if the individual j is located on the same shelf, in the adjacent column;
(b) $\alpha_{ij} = \alpha_1$ if it is located on an adjacent shelf, in the adjacent column;
(c) $\alpha_{ij} = \alpha_2$ if it is located on an adjacent shelf, in the same column (so it is not pre-infected).

This is due to the fact that the virus particles shed by a particular animal from the pre-infected column will move mainly to the other cage located on the same shelf. However, a certain fraction of them, α_1, may diffuse to each of the two adjacent shelves (or out of the system if the source shelf is at the edge). In addition, a certain fraction α_2 of the virus shed by individuals infected via air may also cross the shelf border and reach a neighboring animal in the same column. It seems reasonable to expect that $\alpha_2 < \alpha_1$.

We have already assumed that the virus concentration is always the same for all pre-infected individuals. The part of S_τ that comes from them is then equal to this concentration (multiplied by $1 - \alpha_1$ if the respective individual is located on the very top or bottom shelf).

3.3. An Infection Space

The coefficients α_1, α_2, and $\gamma(T, H)$ must be determined experimentally. In order to do so, we introduce a three-dimensional vector space V. The result of a particular experiment will be represented by a vector $v = [x, t, \sigma] \in V$. The coordinates of this vector, in a certain base, have the following meaning:

x = the ratio of the individuals that were infected via air during the whole experiment to all the previously healthy individuals.
t = the average time (number of days) elapsed before infection onset.
σ = standard deviation for the distribution of infection onset days.

We define a metric tensor on V as

$$g = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Thus, we are able to compute an abstract distance between two post-experimental states v_1, v_2 by simply taking the square root of the scalar product of $v_1 - v_2$ with itself:

$$d(v_1, v_2) := \sqrt{10(x_1 - x_2)^2 + 5(t_1 - t_2)^2 + 2(\sigma_1 - \sigma_2)^2}. \tag{3}$$

This will allow comparison of the results of different experiments and simulations quantitatively. The diagonal terms of g have been chosen according to the range of variety of appropriate coordinates. We have set a large coefficient (equal to 10) for the total infection rate x as it is a value between 0 and 1, and differs only slightly. On the other hand, the standard deviation of infection day, σ, is taken with a small coefficient (equal to 2) because it may differ considerably, even for simulations performed with similar parameter values. Note that this choice, although it seems reasonable, is arbitrary. Using another distance definition, it is possible to obtain quite different results.

Table 2
The assumed virus concentration values for the guinea pigs infected via air (A) and pre-infected (P).

	Day									
	1	2	3	4	5	6	7	8	9	10
A										
20 °C	0.00	1.62	3.85	6.08	7.08	5.45	3.82	2.18	0.55	0.00
5 °C	0.00	2.02	4.05	6.08	7.06	5.79	4.52	3.25	1.98	0.72
P										
20 °C	5.07	7.30	6.70	5.42	4.14	2.87	1.60	0.32	0.00	0.00
5 °C	4.97	7.00	8.54	7.32	6.10	4.87	3.64	2.42	1.20	0.00

For the latter we took $\beta(t) = -1.28t + 10.52$ at 20 °C and $-1.23t + 12.22$ at 5 °C for the descending phase.

Table 3
Abstract distances between the experimental data points.

	20%	35%	50%	65%	80%
20%	–	1.20	3.75	1.39	×
35%	×	–	3.82	1.23	×
50%	×	1.88	–	2.60	×
65%	×	2.61	1.21	–	×
80%	×	4.15	2.33	1.77	–

The values in the right-upper corner correspond to temperature 20 °C, and in the left-lower corner, to 5 °C. We have taken the average *V*-coordinate values wherever there were two experiments performed under the same conditions.

Distances between points representing the experiments of Lowen et al. (2007) are collected in Table 3.

3.4. Extraction of Unknown Parameters

Using the few experimental points in *V*, it is necessary to determine the unknown parameters α_1, α_2, and $\gamma(T, H)$. This is done by using the iterative algorithm described in Section 3.2 to generate a number of points corresponding to various reasonable parameter sets. Each of these sets will then be classified, separately for each temperature, as an approximation of parameter values for one particular humidity level *H* corresponding to the available experimental data, namely, the one that is represented by the experimental point of least distance to the point considered. After that, a search is performed for the most relevant values of α_1 and α_2, that cover the maximal number of experimental cases for both temperatures and with the smallest possible distances. The last step involves determination of γ values for each *T* and *H* separately.

For the simulations here we set 17 irregularly spaced checkpoints for α_1 and α_2 from 0 to 0.5, and 62 checkpoints for γ from 0.001 to 0.5. They are distributed more densely for lower parameter values.

4. Results and Discussion

The obtained classification of experimental points in the parameter space is shown graphically in Fig. 2. The most striking feature is that neither of the parameter sets has been classified to 35% humidity[3] for 20 °C. Also the points belonging to 20% humidity lie only in a quite narrow range of high, and rather unexpected, α_2 values. We consider the occurrence of late infections in the corresponding experiments, together with the lack of infections on intermediate days, to be a possible cause of this behavior. Indeed, in our model such a situation is most similar to the distributions provided by surprisingly high values of α_2. But this could never be typical because the infection probability (Eq. (2)) is always a decreasing function of time, at least from the 3rd day. However, during Monte Carlo simulations, it is possible sometimes to obtain results that display this feature, even for quite reasonable parameters. Due to the experimental data set limitations, involving a rather small number of animals, it is not easy to determine whether those late infections are simply fluctuations, or rather a symptom of a certain phenomenon the model does not yet take into consideration.

Another striking feature of the classification for 20 °C is the fact that the 50% humidity area is placed uppermost, above the areas representing 20% and 65% humidity. This is not consistent with the above-mentioned pattern of transmission dependency on relative humidity that was observed in the experimental results. Indeed, there was a local minimum of transmission observed at this humidity value, and therefore we would rather expect γ values in corresponding parameter sets to be smaller, not larger than the rest. This divergence is probably due to the extreme limitation of available data—as the virus transmission under these conditions was at a low level, only two infection cases occurred. This resulted in a very small, unrealistic standard deviation for the distribution of infection days. Thus, the distance (Eq. (3)) between the simulation results for parameters expected to be relevant and the corresponding experimental result was much greater than it should be and exceeded the distance to the closest improper condition set. On the other hand, a small σ value corresponds to the situation that occurs in the model when γ is very high—nearly all pigs become infected on the 1st day. This might be a possible explanation why the parameter sets with high γ values have been classified to this humidity level conditions.

For 5 °C the least distances (with all experimental points included) appear for the four sets of α_1 and α_2 values shown in

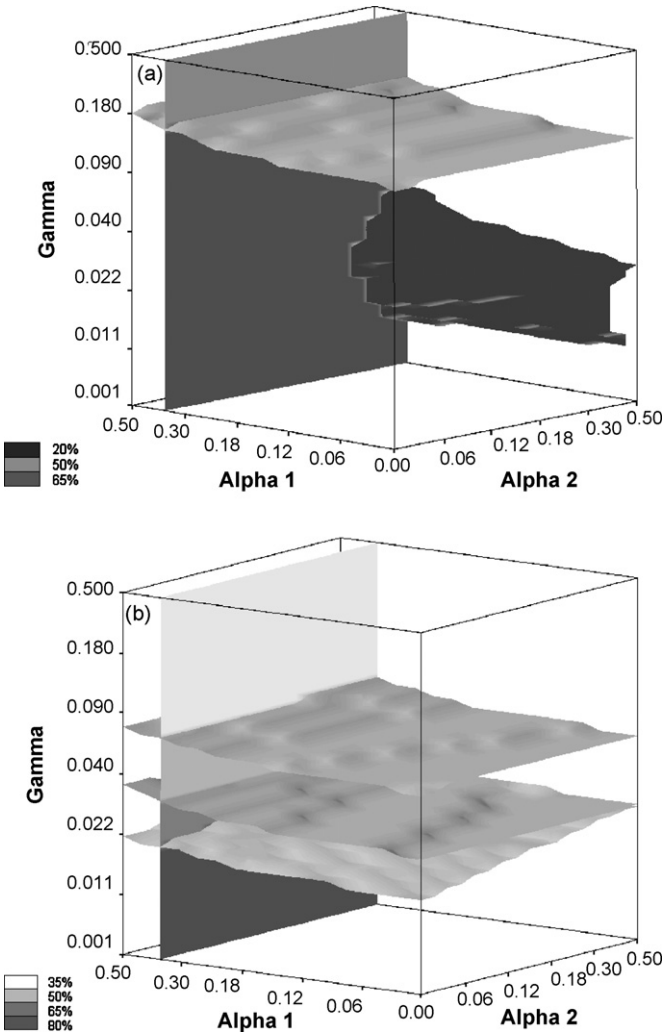

Fig. 2. Classification of the parameter sets for (a) 20 °C and (b) 5 °C. For both temperatures, each humidity value sets a single compact area in the parameter space. The curved surfaces are boundaries between those areas; the respective humidity values are indicated by the color of the appropriate region of the vertical slice surface. The dark surface on (a) is drawn between the 20% and 65% areas.

Table 4
The four α_1 and α_2 value sets for which minimal distances at 5 °C appear, with all humidity values being represented.

α_1	0.35	0.40	0.45	0.50
α_2	0.00	0.00	0.02	0.02

[3] The lack of 80% points for 20 °C is due to lack of the corresponding virus transmission, and of 20% points for 5 °C, of an appropriate data, respectively.

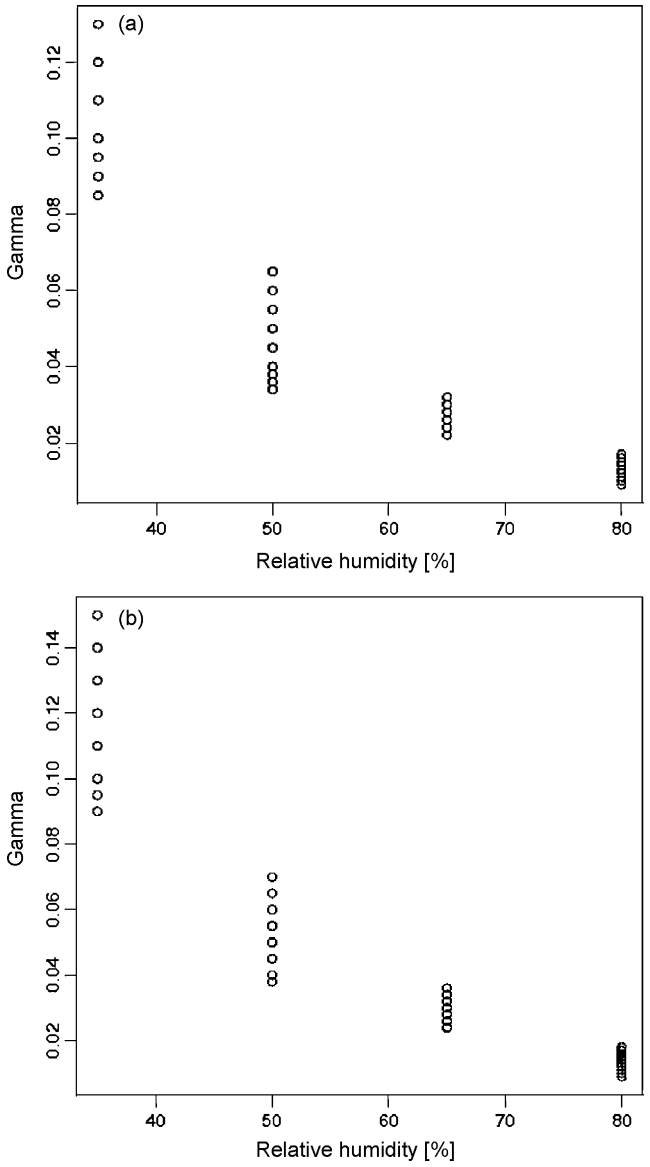

Fig. 3. Gamma vs. humidity for 5 °C and (α_1, α_2) equal (0.35, 0) and (0.45, 0.2), respectively. Only the points for which the distance is not greater than 0.8 are shown.

Table 4. Possible values of γ for particular humidity levels at this temperature, provided those values of alphas, are shown in Fig. 3. They seem to be consistent with the experimental results from (Lowen et al., 2007).

For 20 °C the minimal distances (of about 0.2) to the 65% humidity experimental result occur, with the above alphas, for γ values 0.04–0.045. This is the only conclusion that appears reasonable since the experimental data for other humidity values are probably distorted by statistical fluctuations, arising from the insufficient number of cases. The *V*-space distances between any point classified to 20% or 50% humidity and the appropriate experimental result are greater than most of the distances between particular experimental results, indicating that the classification for this temperature is not satisfactory. Due to the sparseness of experimental data, it would be interesting to search for another metric that does not rely on the distribution of the days on which infection occurred. Or, possibly, to calibrate our simulation model with a significantly larger experimental data set. Nevertheless, we consider the correlation between experimental and simulation results to be sufficient for the purpose of using transmissibility values obtained from this simple model as one of the inputs for intended large-scale epidemiological studies in the near future.

Acknowledgements

The authors were partially supported by the EU project SSPE-CT-2006-44405, and also partially supported from the 352/6.PR-UE/2007/7 grant. We also appreciate the contribution of Dr. Krzysztof Nowiński, from whom we received much help with generating plots, and of Andrew James Churchard, who corrected our English in this article.

References

Atkinson, M.P., Wein, L.M., 2008. Quantifying the routes of transmission for pandemic influenza. Bull. Math. Biol. 70, 820–867.
Carrat, F., Vergu, E., Ferguson, N.M., Lemaitre, M., Cauchemez, S., Leach, S., Valleron, A.J., 2008. Time lines of infection and disease in human influenza: a review of volunteer challenge studies. Am. J. Epidemiol. 167 (7), 775–785.
Collier, L., Oxford, J., 1993. Human Virology. Oxford University Press.
Ferguson, N.M., Cummings, D.A.T., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., Iamsirithaworn, S., Burke, D.S., 2005. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature 437, 209–214.
Lai, A.C.K., Cheng, Y.C., 2007. Study of expiratory droplet dispersion and transport using a new Eulerian modeling approach. Atmos. Environ. 41 (35).
Lofgren, E., Fefferman, N., Naumov, Y.N., Gorski, J., Naumova, E.N., 2007. Influenza Seasonality: Underlying Causes and Modeling Theories. J. Virol. 81 (11), 5429–5436.
Lowen, A.C., Mubareka, S., Steel, J., Palese, P., 2007. Influenza virus transmission is dependent on relative humidity and temperature. PLoS Pathogens 3 (10), e151, doi:10.1371/journal.ppat.0030151.
Stegeman, A., Bouma, A., Elbers, A.R.W., de Jong, M.C.M., Nodelijk, G., de Klerk, F., Koch, G., van Boven, M., 2004. Avian influenza A virus (H7N7) epidemic in The Netherlands in 2003: course of the epidemic and effectiveness of control measures. J. Infect. Dis. 190.
Tellier, R., 2006. Review of aerosol transmission of influenza A virus. Emerg. Infect. Dis. 12 (11).
Wang, B., Zhang, A., Sun, J.L., Liu, H., Hu, J., Xu, L.X., 2007. Study of SARS transmission via liquid droplets in air. J. Biomech. Eng. 127 (1), 32.