Seasonality of absolute humidity explains seasonality of influenza-like illness in Vietnam

Pham Quang Thai [a,b,*], Marc Choisy [b,c], Tran Nhu Duong [a], Vu Dinh Thiem [a], Nguyen Thu Yen [a], Nguyen Tran Hien [a], Daniel J. Weiss [d], Maciej F. Boni [e,f], Peter Horby [b,e]

[a] National Institute of Hygiene and Epidemiology, Hanoi, Viet Nam
[b] Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Hanoi, Viet Nam
[c] MIVEGEC, University of Montpellier, CNRS 5290, IRD 224, Montpellier, France
[d] Spatial Ecology & Epidemiology Group, Department of Zoology, University of Oxford, Oxford, UK
[e] Centre for Tropical Medicine, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK
[f] Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Viet Nam

ARTICLE INFO

ABSTRACT

Background: Experimental and ecological studies have shown the role of climatic factors in driving the epidemiology of influenza. In particular, low absolute humidity (AH) has been shown to increase influenza virus transmissibility and has been identified to explain the onset of epidemics in temperate regions. Here, we aim to study the potential climatic drivers of influenza-like illness (ILI) epidemiology in Vietnam, a tropical country characterized by a high diversity of climates. We specifically focus on quantifying and explaining the seasonality of ILI.

Methods: We used 18 years (1993–2010) of monthly ILI notifications aggregated by province (52) and monthly climatic variables (minimum, mean, maximum temperatures, absolute and relative humidities, rainfall and hours of sunshine) from 67 weather stations across Vietnam. Seasonalities were quantified from global wavelet spectra, using the value of the power at the period of 1 year as a measure of the intensity of seasonality. The 7 climatic time series were characterized by 534 summary statistics which were entered into a regression tree to identify factors associated with the seasonality of AH. Results were extrapolated to the global scale using simulated climatic times series from the NCEP/NCAR project.

Results: The intensity of ILI seasonality in Vietnam is best explained by the intensity of AH seasonality. We find that ILI seasonality is weak in provinces experiencing weak seasonal fluctuations in AH (annual power <17.6), whereas ILI seasonality is strongest in provinces with pronounced AH seasonality (power >17.6). In Vietnam, AH and ILI are positively correlated.

Conclusions: Our results identify a role for AH in driving the epidemiology of ILI in a tropical setting. However, in contrast to temperate regions, high rather than low AH is associated with increased ILI activity. Fluctuation in AH may be the climate factor that underlies and unifies the seasonality of ILI in both temperate and tropical regions. Alternatively, the mechanism of action of AH on disease transmission may be different in cold-dry versus hot-humid settings.

1. Introduction

Influenza remains a globally important illness which results in high mortality, morbidity and economic burden (Stephenson and Zambon, 2002). The epidemiology of influenza is diverse, with a stronger seasonality in the temperate regions than in the tropical ones, and differences in epidemic timing between different parts of the world (the two hemispheres being notably out of phase) (Viboud et al., 2006a). Characterizing and understanding the seasonality of influenza is critical for optimizing vaccination and other control policies (Bolker and Grenfell, 1996; Grenfell and Harwood, 1997; Grenfell et al., 2001) as well as sampling of viruses in order to update the vaccine from year to year (Ampofo et al., 2013).

A number of hypotheses have been explored to explain the seasonality of influenza, among which climatic factors appear to

* Corresponding author at: National Institute of Hygiene and Epidemiology, Hanoi, Viet Nam. Tel.: +84 9 13 09 27 77; fax: +84 4 38 21 04 87.
E-mail addresses: phamquangthai@gmail.com, phamthaidt@yahoo.com (P.Q. Thai).

be most important (Tamerius et al., 2011). Experimental studies on guinea pig models have shown an increase in influenza virus transmissibility as the absolute humidity (AH) decreases (Lowen et al., 2007; Shaman and Kohn, 2009). An epidemiological study of influenza in the United States has shown that a drop in AH is a good predictor of epidemic onset (Shaman et al., 2010). Other climatic factors have been proposed as drivers of influenza epidemiology, such as the number of hours of sunshine (Hope-Simpson, 1981), mediated through the effect of Vitamin D synthesis on individuals' innate immune response to infection (Cannell et al., 2006; Liu et al., 2006; Yamshchikov et al., 2009). However, this hypothesis remains unproven (Shaman et al., 2011).

The investigation of the links between influenza epidemiology and climatic factors has followed two paths. The first one has looked at predictors of epidemic onset, for example in the United States (Shaman et al., 2010). More recent studies conducted at the global scale by Bloom-Feshbach et al. (2013) and Tamerius et al. (2013) have shown that epidemics tend to start in "cold-dry" and "humid-rainy" conditions in temperate and tropical regions respectively. However, this approach of working on epidemic onset is relevant only when the epidemiology is seasonal enough for epidemic onset to be characterized. Since the intensity of the seasonality itself varies latitudinally, both at the between-country (Viboud et al., 2006a) and at the within-country (Alonso et al., 2007) levels, the second path of investigation has examined predictors of the intensity of seasonality. A recent study carried out in different localities of China quantified the seasonality of influenza from the Fourier value around the period of 1 year and related it to a number of potential climatic drivers (Yu et al., 2013). The results of this study found temperature and hours of sunshine to be good predictors of influenza seasonality. However, AH was not tested in this analysis, making comparisons with previously published results on the topic difficult (see Lowen et al., 2007; Shaman and Kohn, 2009; Shaman et al., 2010 cited above).

In this article we investigated the climatic factors that explain the intensity of the seasonality of influenza-like illness (ILI) in Vietnam. Vietnam is a tropical country characterized by a high population density (90 million inhabitants on 330,000 km^2), and a high diversity of climates resulting from Vietnam's wide latitudinal (8–24° N) and elevational (0–3200 m) ranges. These characteristics make Vietnam an ideal country to test the effects of climatic factors on ILI epidemiology. The climatic variables tested in the present study include minimum, mean, and maximum temperatures, relative and absolute humidities, rainfall, and hours of sunshine. Our results are then tentatively extrapolated to the global scale and compared with previously published reports.

2. Materials and methods

2.1. ILI data

ILI is one of 26 notifiable communicable diseases in Vietnam. Monthly data have been aggregated by province by the Ministry of Health of Vietnam since 1979. Given quality issues in the data before 1993, we focused our analysis on the 1993–2010 time period. A number of provinces split since 1993 and, for consistency, we used the province administrative borders of 1993 (i.e. 52 provinces instead of 63 as of today).

2.2. Climatic data

Meteorological data from January 1993 to December 2010 were obtained from the Vietnamese Institute of Meteorology, Hydrology and Environment in 67 climatic stations spanning all of Vietnam. The data include the monthly means of the minimum, mean, and maximum temperatures (°C), the monthly means of absolute (g/L) and relative (%) humidities, and the monthly cumulative sums of rainfall (mm) and hours of sunshine. Each province was assigned the climatic station that is closest to its population centroid. Province population centroids were calculated from commune population sizes and commune geographic centroid coordinates (≈200 communes per province).

2.3. Outlier identification

The identification of outliers was performed on stationarized and normalized versions of the time series obtained by simple differencing. On these normalized time series, we computed the 1st and 99th expected percentile values from the mean and standard deviation. Outliers were identified as data points either below the 1st percentile or above the 99th percentile. The identified points were quite far below and far above the 1st and 99th percentiles. They are likely errors created at the time of reporting and appear at a very low rate (≈1%) but with extremely high values, hence our willingness to discard them.

2.4. ILI data transformations

Outliers in the ILI time series were removed and missing data were linearly interpolated. ILI time series were then square-root transformed for variance stabilization, and then detrended, centered and reduced. The trends were estimated by lowess regressions (Crawley, 2007) with smoother spans of 0.1, which captured the trends neatly for all the provinces.

2.5. Intensity of seasonality

Wavelet decompositions using the Morlet wavelet (Torrence and Compo, 1998; Cazelles et al., 2008) were carried out on the ILI time series of all the provinces to check for stationarities. The intensities of seasonality of ILI and climatic time series were quantified by the power value at the period of 1 year in the global wavelet power spectrum (equivalent to a Fourier spectrum). Wavelet decompositions were performed using the `biwavelet` R package (Gouhier, 2013).

2.6. Regression tree

Relationships between climatic variables and the intensity of ILI seasonality were explored by a regression tree, using the `tree` R package (Ripley, 2012). In order to explore as many climatic characteristics as possible, we computed, for each climatic variable, a number of summary statistics. These statistics include the minimum, mean, and maximum values, the amplitude (difference between minimum and maximum values), the intensity of seasonality as defined by wavelet analysis, and the number of months during which the value of the variable is below a given threshold (varying this threshold over 50 values regularly spaced between the minimum and the maximum values). In total, these summary statistics for all the climatic variables generated 534 potential explanatory variables that were entered into the regression tree analysis to identify climate factors that best explained the observed intensity of ILI seasonality.

2.7. Extrapolation to the global scale

In order to extrapolate the result of the regression tree to the global scale we used monthly climatic data simulated from the NCEP/NCAR project (Kalnay et al., 1996) by pixels of 2.5° × 2.5°, over the 1993–2010 time period. Since AH was not included in
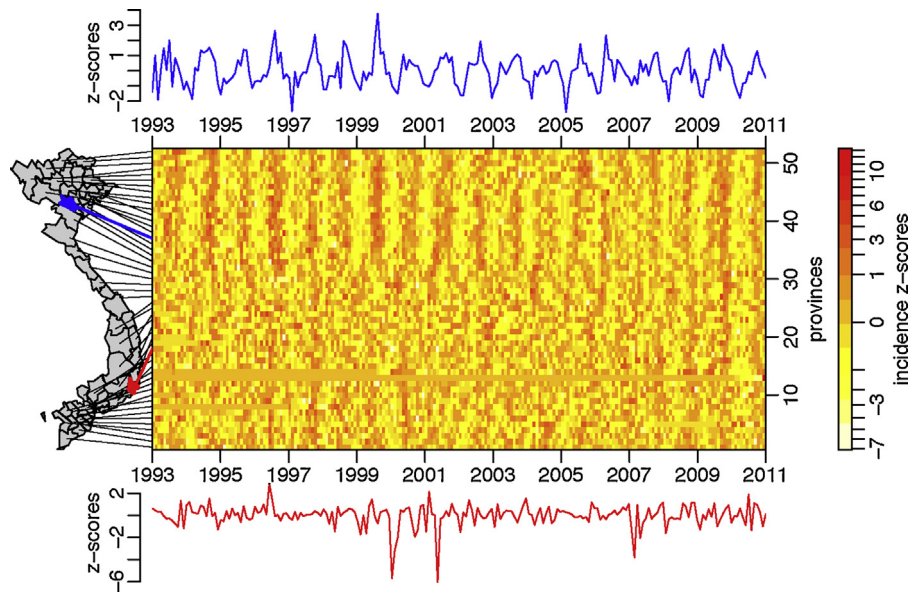
Fig. 1. Heatmap of ILI incidences. Each row of the matrix corresponds to a province (52 in total) and each column corresponds to a month from January 1993 to December 2010 (216 in total). For each province, the time series of ILI raw incidence has been detrended, centered and reduced, in order to allow comparisons between provinces and between months. The color of each cell shows the value of these incidence scores, on a square-root scale for better visibility (see the scale on the right). The rows of the matrix are arranged according to the latitude of the population centroid of the province, as can be seen from the map on the left. On this map, each line connecting to the matrix starts from the population centroid of the province. On the top and the bottom of the matrix are shown the time series of the detrended, centered and reduced time series of ILI incidence for the provinces of Hoa Binh and Ninh Thuan as examples. These provinces are colored in blue and red, respectively, on the map. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the NCEP/NCAR data base, it was calculated from temperature and relative humidity using the formula described in Bolton (1980).
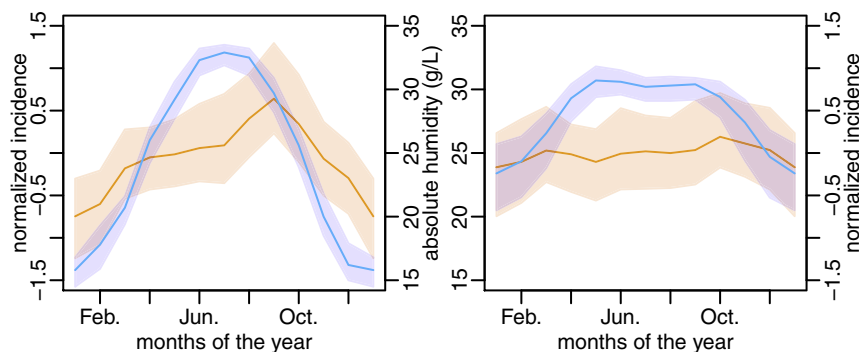
3. Results

3.1. Geographic patterns of ILI incidence

During the 18-year period studied, 26,023,574 cases of ILI were reported, ranging between 320,525 (1993) and 1,824,195 (2009) per year, due to changes in reporting volume during this period. Fig. 1 shows the heatmap of the transformed ILI time series for the 52 provinces from south to north. ILI incidence in northern Vietnam is highest in August, September, and October (see red curve in Fig. 2A), and ILI incidence exhibited little variation throughout the year in southern Vietnam (see red curve in Fig. 2B). Local wavelet power spectra (not shown) revealed stationary signals and both local and global (Fig. 3A) wavelet decompositions for the ILI time series of all the provinces showed a consistent seasonality

through time for the northern provinces. Fig. 3B shows that the non-northern provinces with high intensity of seasonality are located in mountainous areas (Fig. 3C). These latitudinal and elevational factors suggest a potential role of climatic drivers in explaining the observed diversity in the intensity of ILI seasonality.

3.2. Geographic patterns of climatic variables

Fig. 4 shows the annual variations of the climatic variables in the 67 stations in Vietnam. It shows a gradient of seasonality of temperature and AH from south to north where both temperature and AH peak in the summer. In contrast, relative humidity appears to be much more variable in the south than in the north where it drops during the winter. Rainfall is seasonal throughout Vietnam, with a shift of the rainy season from the summer in both the north and the south to the autumn months in central Vietnam. In central Vietnam, the amplitude of the rainfall time series is also much higher than in the north and the south of the country. Finally,

Fig. 2. Seasonalities of ILI and AH. Normalized ILI (orange) and AH (blue) are shown for the north (left) and the south (right). The limit between north and south is here arbitrarily defined at the latitude of 19° N, but the results are robust respective to the definition of this limit. The lines represent the medians over the 18 years of the studied period and the shaded areas represent the inter-quartile ranges. The numbers of provinces are 21 in the north (left panel) and 31 in the south (right panel), and the numbers of climatic stations are 32 and 35 for the north and the south respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
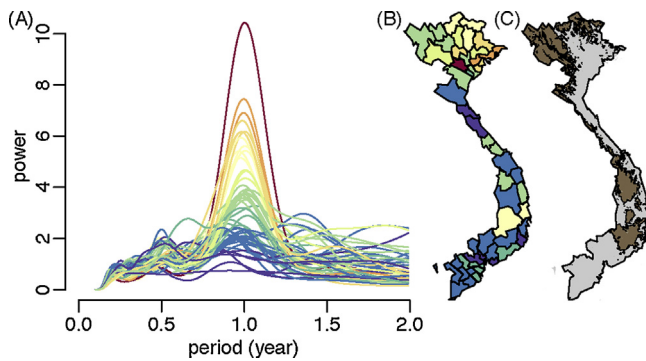
Fig. 3. Intensities of ILI seasonality. Global wavelet power spectra for the detrended, centered and reduced time series of ILI incidences (A, one curve per province, 52 in total). The colors of the curves code the values of the maximum power within the 0.9-to-1.1-year period range. The same color code is used to plot the intensity of seasonality on the map of Vietnam (B). The second map (C) shows elevations higher than 500 m in brown. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

sunshine is strongly seasonal throughout the country with the peak shifting from the winter time in the south to the summer time in the north. The mean sunshine is also higher in the south, with the minimum values in the south equal to the maximum values in the north. A principal components analysis (PCA) of the climatic variables reveals that temperature and AH are positively correlated and contribute to explain 59% of the total climatic variance (first PCA axis, Fig. 5). Rainfall and relative humidity are positively correlated and both, as expected, negatively correlated to sunshine. These variables contribute to explain 24% of the total climatic variance (second PCA axis). We can further see that the first PCA component has a variance that increases from south to north. No such trend is observed for the second PCA component.

3.3. Association between ILI seasonality and climatic variables

The regression tree analysis showed that the intensity of AH seasonality has more explanatory power than any other climatic summary statistic in explaining the intensity of ILI seasonality (Fig. 6A). This result was robust to the number of explanatory variables included into the regression tree analysis. Temperature plays a role also, both alone (node d) and in combination with absolute humidity (i.e. relative humidity, nodes a and c). However, the effect of temperature is negligible compared to the effect of the absolute humidity as can be judged from the lengths of the branches from nodes a, c and d. The regression tree analysis shows that provinces with low intensity of AH seasonality (annual power below a threshold of 17.60) have ILI seasonality of low intensity, and provinces with high intensity of AH seasonality (annual power above the threshold of 17.60) have ILI seasonality of high intensity. Fig. 6B shows that the relationship between the intensities of the two seasonalities (ILI and AH) is non-linear. Finally, note that the tree is well balanced, with comparable numbers of data points on each side of each node. This suggests that there is no outlier that drives the entire signal.

In order to characterize the link between AH and ILI seasonalities in more detail, Fig. 7 investigates the phase difference between the seasonal components of AH and ILI. As expected from the fact that both AH and ILI seasonalities are much more apparent in the north than in the south, the phase differences between the two variables is much more consistent both in time and space in the north (left panel of Fig. 7) than in the south (right panel of Fig. 7). Despite some temporal and spatial variability in the phase difference in the north, we observe an average lag of one month between AH and ILI incidence, with changes in AH preceding changes in ILI.

To facilitate comparison with previously published results on the onset of epidemics, Fig. 2 shows the annual variations of ILI (orange) and AH (blue) for the 21 northern (left) and 31 southern provinces (right), and Fig. 8 shows, for each of the provinces ordered by latitude (y-axis), the values of the Pearson correlation coefficients between AH and ILI (x-axis). Both these figures show that AH and ILI tend to vary in the same direction, with high rates of ILI notification correlated with high AH, even though the correlations appear stronger in the north of the country where the seasonalities of both ILI and AH are the strongest.

3.4. Extrapolation to the global scale

Results from our analysis on the Vietnamese data show that the intensity of AH seasonality is the variable that best explains the intensity of ILI seasonality. Specifically, intensities of AH seasonalities below and above a threshold power of 17.60 predict ILI seasonalities of low and high intensities respectively. In order to extrapolate this result to the global scale, we calculated, from the NCEP/NCAR project (Kalnay et al., 1996), monthly time series of AH over the 1993–2010 time period, from all around the globe, with a spatial resolution of $2.5° \times 2.5°$. This yielded 3509 time series on terrestrial pixels. For each of these time series, we computed the intensity of AH seasonality (Fig. 9A) from which we predicted the intensity of ILI seasonality (Fig. 9B, high in green, low in red). Fig. 9B also plots the pattern of seasonality (annual versus biannual) as reported by Tamerius et al. (2013). Assuming that biannual signals are less seasonal that annual ones, Fig. 9B shows some agreement between our model predictions and documented influenza seasonality ($p = 0.04$, Fisher exact test). It gives a sensitivity and a specificity of 85 and 39% for predicting seasonality.

4. Discussion

Our analysis shows that in Vietnam the intensity of ILI seasonality is best explained by the intensity of AH seasonality. Extrapolation of this result to the global scale seems to be consistent with the majority of seasonality patterns reported by Tamerius et al. (2013). From our quantification of the intensity of seasonality based on the power value at the period of 1 year in the global power spectrum, we can argue that biannual signals will be classified as less seasonal than annual ones.

It is interesting to note that, in the general latitudinal gradient of seasonality, locations with higher elevation tend to have stronger seasonality (Fig. 3). This suggests that it is the local environment that controls the ILI epidemiology as already observed on influenza in Brazil (Alonso et al., 2007), rather than the ILI epidemiology being remotely driven by the seasonal environment of some driver location (e.g. a large city) according to a gravity model (Xia et al., 2004) as observed on measles in England & Wales (Grenfell et al., 2001), pertussis in the US (Choisy and Rohani, 2012) and influenza in the US (Viboud et al., 2006b).

When looking at the actual time series, ILI is significantly and positively correlated with AH in 19 ($p < 0.01$) or 28 ($p < 0.05$) out of 52 provinces in Vietnam (mostly in northern Vietnam). This positive association is in contrast to the association between influenza transmission and low AH reported by Lowen et al. (2007) and Shaman and Kohn (2009) in guinea pig transmission experiments and by Shaman et al. (2010) from human epidemiological data in the United States. However, it is in agreement with the observation that influenza in tropical regions peaks in the "humid-rainy" season Tamerius et al. (2013). Only one province in Vietnam (Phu Yen) had a significant negative correlation between AH and ILI. Our results confirm the view of Tamerius et al. (2011, 2013) that the exact links between the different climatic variables and influenza transmission
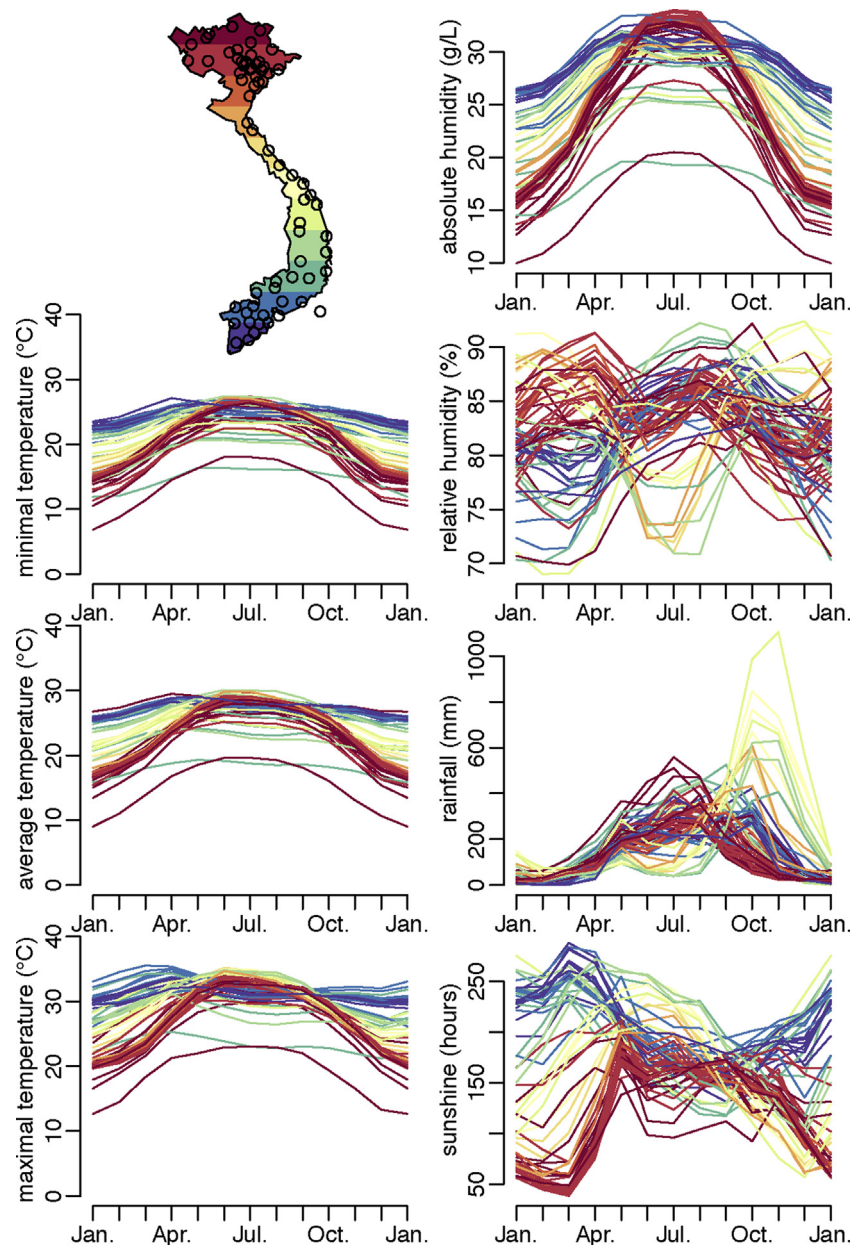
Fig. 4. Climatic variables. Annual variations of climatic variables in Vietnam: means (over the 18 years of data) of minimum, mean, and maximum temperatures (°C), absolute and relative humidities (g/L and % respectively), and cumulative sums of rainfall (mm) and number of hours of sunshine. Each line corresponds to one climatic station (67 in total). The colors of the lines vary as the latitude of the station, as coded on the map. The map also shows the locations of the climatic stations (open circles).

probably vary in the different parts of the world, and they clearly call for further detailed studies on the direction of correlation and the mechanistic link between AH and ILI.

The climatic data of Vietnam show, on an area of moderate size (330,000 km²), a great diversity of seasonality for the different climatic variables. Temperatures and AH peak in the summer time with a gradient of amplitude from the south to the north. Relative humidity also peaks in the summer but its gradient of amplitude is opposite to the one of AH: it increases from north to south. The peak in hours of sunshine shifts from the winter time in the south to the summer time in the north; the amplitude of these seasonalities are similar in both the north and the south but the mean is higher in the south than in the north (the peak value in the north is similar to the trough value in the south). Rainfall seasonality displays a more complicated pattern with a peak shift from the summer time in the south to the autumn in the center and back to the summer

time in the north. Vietnam thus offers an exceptional diversity of both amplitude and timing of the seasonalities of the different climatic variables, making itself an ideal candidate country to test the relationships between climatic variables and infectious diseases transmission. For example, we can see that the range of absolute humidity values observed in Vietnam (Fig. 6B) covers a large proportion of the global range of absolute humidity variation (Fig. 9A, same color code). Such a diversity of climates in such a small and densely populated area is an asset compared to global-level comparative studies which aggregated disease epidemiology over large areas and for which other factors (demography, behavior, etc.) may vary substantially and act as important confounders.

The extrapolation of our model to the global scale is not perfect but it still gives surprisingly good results. Of note, a number of mismatches correspond to boundary regions of our model prediction. Given that our model predictions are themselves based on

simulated climatic variables, it is likely that the boundary regions of our model predictions are imprecise. Moreover, the size of the pixel used in the NCEP/NCAR data base (2.5° × 2.5°) is much larger than the area over which the epidemiological data have been collected (most of them correspond to one city). In consequence, local climatic conditions that affect the epidemiology may be poorly rendered by the pixel averages of the NCEP/NCAR data base. This is probably the case, in particular, on the coastal regions of Australia and south China. Given these potential limitations, it is actually surprising to observe that a model fitted to Vietnamese data gives such a good extrapolation to the global scale. We think this may be explained by the high diversity of climates observed in Vietnam.

Investigation of the relationship between climatic variables and influenza transmission is usually done by testing each potential climatic variable in turn and comparing the fits. Using such methods, Shaman and Kohn (2009) reanalyzed the data of Lowen et al. (2007) and were the first to propose that AH was better than relative humidity as a predictor of influenza transmission. Such investigations involve different statistical tools such as linear regressions (e.g. Tamerius et al., 2013), generalized linear models (e.g. Yu et al., 2013), and mathematical modeling (e.g. Shaman et al., 2010). However, whatever the framework, interactions and colinearities between potential explanatory variables are rarely accounted for. This poses a problem given that the different climatic variables are most of the time highly colinear (see for example the PCA in Fig. 5). This issue questions the validity of comparing different studies with different sets of potential explanatory variables, and different ways of dealing with their interactions and colinearities. As an example, our analysis in Vietnam is very similar in approach to the one recently published by Yu et al. (2013) in China. Using generalized linear modeling they show that their observed latitudinal gradient of influenza seasonality is best explained by temperature and hours of sunshine. However, their analysis did not test the effect of AH, which makes their result difficult to compare with our study, and also with previous work reporting the influence of AH (e.g. Shaman et al. (2010) in the United States). This calls for the combined analysis of data from different sources as initiated by Tamerius et al. (2013).

The regression tree that we used to investigate the relationships between ILI seasonality and the climatic variables is a simple binary recursive partitioning method and thus allows us to detect links
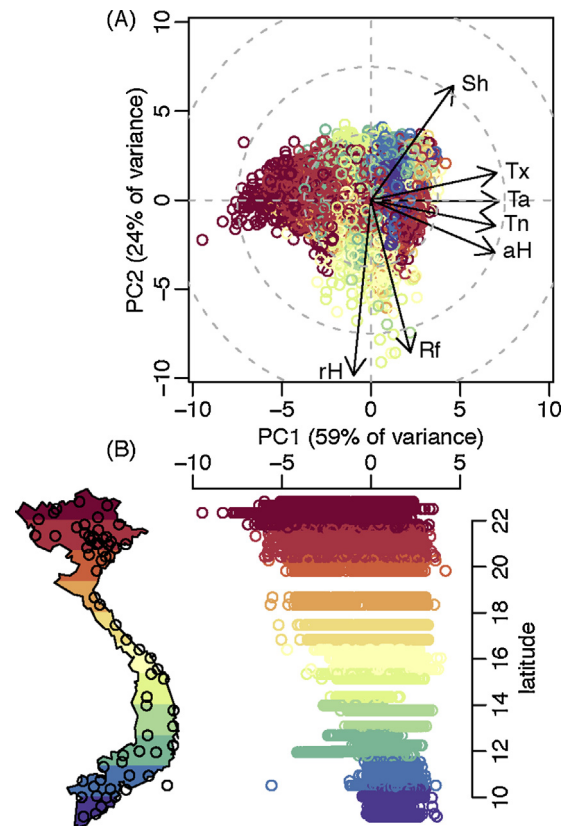

Fig. 5. Latitudinal gradient of climatic seasonality. (A): principal component analysis (PCA) of the seven climatic variables of Fig. 4: minimum (Tn), mean (Ta), and maximum (Tx) temperatures, absolute (aH) and relative (rH) humidities, amount of rainfall (Rf), and hours of sunshine (Sh). The first two components explain more than 83% of the total variance. Each dot corresponds to a given month, for a given climatic station. The colors of the dots vary according to the latitude of the climatic stations, as shown on the map showing the locations of the climatic stations. (B): relationship between the first component of the PCA (A) and the latitude of the climatic stations. The latitude axis of the plot also corresponds to the map on the left.
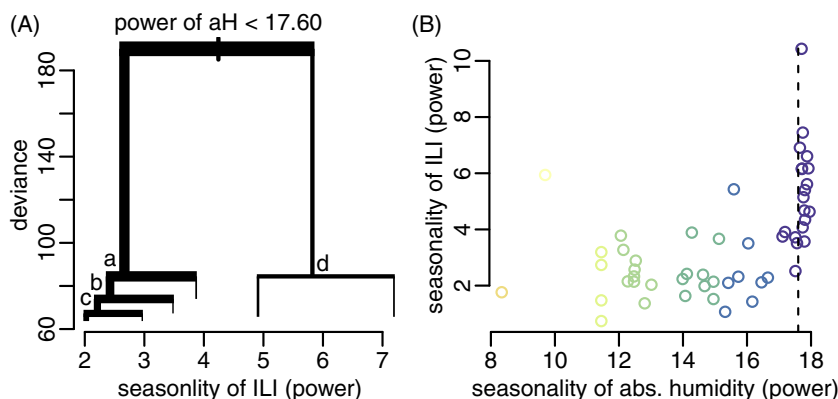
Fig. 6. Relationship between ILI seasonality and climatic characteristics. (A): regression tree of the intensity of ILI seasonality using, for the explanatory variables, the 534 summary statistics of the climatic variables computed as described in Section 2. The widths of the horizontal segments are proportional to the number of provinces. The first node discriminates the provinces for which the intensity of AH seasonality is below (left, 36 provinces) and above (right, 16 provinces) a threshold value of 17.60. Node (a) discriminates the provinces for which the total number of months (over the 18-year period) with a relative humidity above a threshold value of 77% is above (left) and below (right) 169. Node (b) discriminates the provinces for which the range (maximal minus minimal values) of AH is below (left) and above (right) a threshold value of 14.46 g/L. Node (c) discriminates the provinces for which the intensity of the relative humidity seasonality is below (left) and above (right) a threshold value of 11.63. Node (d) discriminates the provinces for which the total number of months (over the 18-year period) with a maximum temperature above 26 °C is above (left) and below (right) a threshold value of 130. (B): relationship between the seasonality of ILI and the seasonality of absolute humidity (corresponding to the first node in A). Each dot is a province (52 in total) and the vertical dashed line shows the threshold of the first node in the regression tree (A). The color code reflects the value of AH seasonality and is the same as on Fig. 9A.
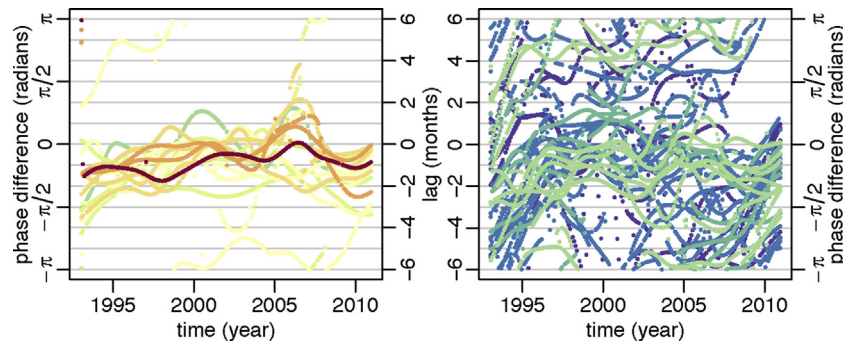
Fig. 7. Lag between ILI and AH. Differences of phase angles between the AH and ILI time series, filtered around the 1-year period (a negative difference thus meaning that AH is ahead of ILI). For better visibility, the left and right panels show the 16 and 36 provinces for which the intensities of AH seasonalities are respectively above and below the threshold value of 17.60 as identified in Fig. 6A. The color code is the same as in Fig. 3 and reflects the intensity of ILI seasonality. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
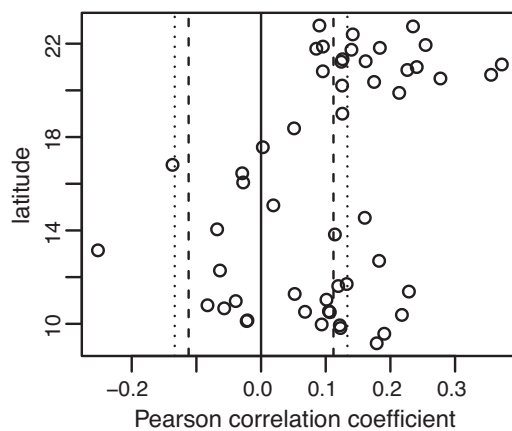
Fig. 8. Correlations between AH and ILI time series. Pearson correlation coefficients between AH and ILI (x-axis) for each of the 52 provinces ordered by latitude (y-axis). The dashed and dotted vertical lines show the 0.05 and 0.01 significance levels respectively.

between a response variable and a number of potential explanatory variables when these latter can be numerous and when the shapes of the relationships can be non-linear and complex. These properties allowed us to test multiple summary statistics of the climate time series and to identify those that have the most explanatory power. For these reasons, we believe that the regression tree technique is an appropriate one to detect the most relevant explanatory variables for such studies. The exact relationship between the best explanatory variable(s) and the response variable can then be refined with other techniques such as generalized linear models, non-linear models, or generalized additive models.

Our analysis revealed a non-linear relationship between the intensities of AH and ILI seasonalities that is almost a threshold effect. Whether this threshold effect emerges from the population-level non-linearity of the transmission process (as in Dushoff et al., 2004) or rather directly from the non-linearity of the effect of AH on individual virus transmission (as shown in Shaman and Kohn, 2009) would require the development and analysis of a mathematical model of influenza transmission, in which the effect of the seasonality of the force of infection on the seasonality of the incidence could be investigated by controlling for all other factors (such as the amplitude or variance of the force of infection). This will be the focus of future work.

A major limitation of our study is the fact that we are working with ILI data rather than data on confirmed influenza infections. This is particularly a limitation in the tropical countries where ILI

activity does not always correlate well with influenza activity (Leo et al., 2010; Nguyen et al., 2009, 2013). Thus our result here should be interpreted more in terms of respiratory disease than influenza specifically. The next appropriate steps from here to understand the potential climatic drivers of influenza transmission would be (i) analyses of confirmed influenza cases as well as cases confirmed to be caused by other respiratory viruses, (ii) an analysis of these different viruses' epidemic dynamics to determine if immunological interference is partially responsible for this pattern, and (iii) the inclusion of climatic variables and school-term into this combined system in order to determine how strongly these extrinsic factors influence respiratory disease dynamics.

For southern Vietnam and the majority of central Vietnam, climate-ILI associations could not be detected because of the low intensity of ILI seasonality. In this sense, the ILI pattern in southern and central Vietnam is similar to that of other tropical regions which exhibit either multiple peaks per year or unpredictable disease patterns (Viboud et al., 2006a; Tamerius et al., 2013). It is these regions whose influenza dynamics have become of interest over the past decade due to the possibility that low-level but long-term influenza persistence in tropical countries may create optimal conditions for generating immune-escape variants that can spread worldwide (Boni et al., 2006; Rambaut et al., 2008; Russell et al., 2008; Adams and McHardy, 2011), and Vietnam may in fact have the right conditions for longer term persistence of influenza viruses (Le et al., 2013). Despite the importance of understanding the tropics and their role in global influenza dynamics, high-quality influenza reporting time series in tropical Asia remain rare with the exceptions of Hong Kong (Cheng et al., 2009, 2012), Singapore (Leo et al., 2010), and recently Vietnam (Nguyen et al., 2009, 2013).

The main finding from our time series analysis on subtropical and tropical Vietnam is that ILI seasonality is most closely associated with AH seasonality. The intensity of AH seasonality seems to correlate well with ILI seasonality in other parts of the world where both climate and influenza data are available, but the most surprising fact about this association, as can be seen in northern Vietnam, is that AH and ILI can correlate either positively (this study) or negatively (other studies). Positive AH-ILI associations are compatible with the increased influenza transmission observed in the 'humid-rainy' season in other tropical climates. We have studied the strength of seasonal variation of AH and it is possible that fluctuation in AH may be the climate factor that underlies and unifies the seasonality of influenza in both temperate and tropical regions. Alternatively, the mechanism of action of AH on virus transmission may be different in cold-dry versus hot-humid settings. This underscores the need for continued future studies on seasonal patterns of influenza transmission in different regions of the world.
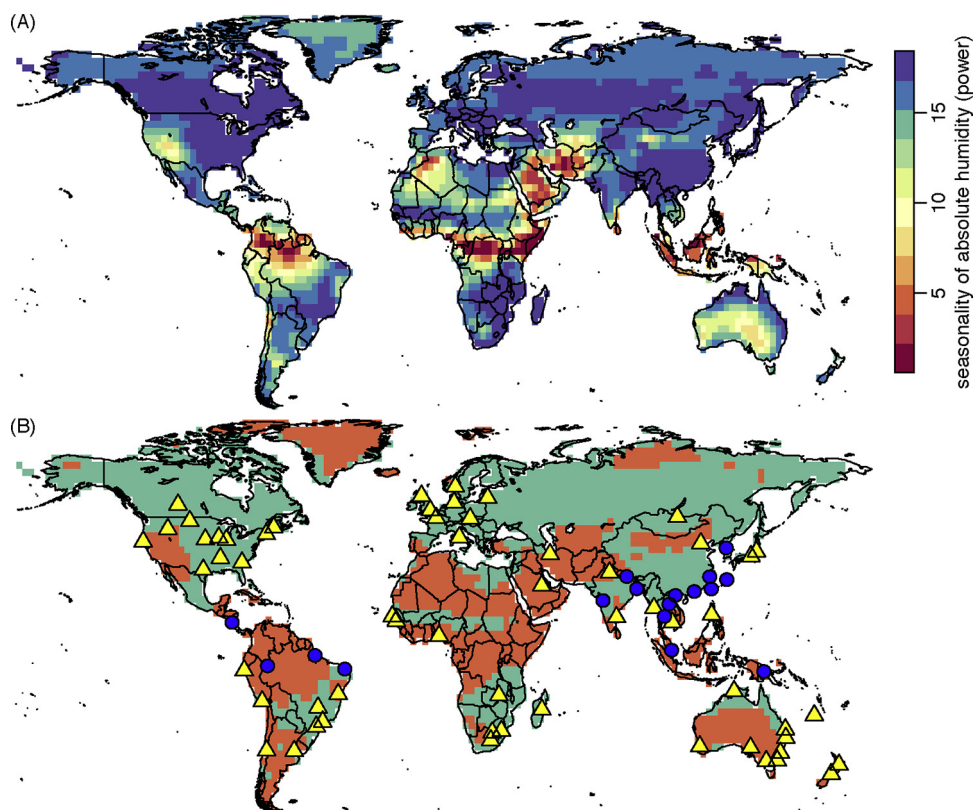
Fig. 9. Global seasonality of AH. (A): seasonality of monthly time series of absolute humidity calculated with a 2.5° × 2.5° spatial resolution (see Section 2). (B): discrimination of the pixels with a intensity of AH seasonality above (green) and below (orange) the threshold value of 17.60 identified in Fig. 6. From Fig. 6, green color predicts high intensity of ILI seasonality and orange color predicts low intensity of ILI seasonality. The yellow triangles and the blue circles show locations where influenza seasonality has been reported as annual or biannual, respectively, by Tamerius et al. (2013).

Authors contribution

PQT, MC, MFB and PH designed the study, performed the analyses and wrote the manuscript. DJW generated the climatic data from the NCEP/NCAR project. TND, VDT, NTY, NTH supervised the collection and entering of the epidemiological and climatic data.

Acknowledgements

References

Adams, B., McHardy, A.C., 2011. The impact of seasonal and year-round transmission regimes on the evolution of influenza a virus. Proc. Biol. Sci. 278, 2249–2256, http://dx.doi.org/10.1098/rspb.2010.2191.

Alonso, W.J., Viboud, C., Simonsen, L., Hirano, E.W., Daufenbach, L.Z., Miller, M.A., 2007. Seasonality of influenza in Brazil: a traveling wave from the Amazon to the subtropics. Am. J. Epidemiol. 165, 1434–1442, http://dx.doi.org/10.1093/aje/kwm012.

Ampofo, W.K., Busaidy, S.A., Cox, N.J., Giovanni, M., Hay, A., Huang, S., Inglis, S., Katz, J., Mokhtari-Azad, T., Peiris, M., Savy, V., Sawanpanyalert, P., Venter, M., Waddell, A.L., Wickramasinghe, G., Zhang, W., Ziegler, T., 2013. Strengthening the influenza vaccine virus selection and development process: outcome of the 2nd WHO. In: Informal Consultation for Improving Influenza Vaccine Virus Selection held at the Centre International de Conférences (CICG) Geneva, Switzerland, 7–9 December 2011, Vaccine, 31, pp. 3209–3221.

Bloom-Feshbach, K., Alonso, W.J., Charu, V., Tamerius, J., Simonsen, L., Miller, M.A., Viboud, C., 2013. Latitudinal variations in seasonal activity of influenza and respiratory syncytial virus (RSV): a global comparative review. PLOS ONE 8, e54445, http://dx.doi.org/10.1371/journal.pone.0054445.

Bolker, B.M., Grenfell, B.T., 1996. Impact of vaccination on the spatial correlation and persistence of measles dynamics. Proc. Natl. Acad. Sci. U. S. A. 93, 12648–12653.

Bolton, D., 1980. The computation of equivalent potential temperature. Monthly Weather Rev. 108, 1046–1053, http://dx.doi.org/10.1175/1520-0493(1980)108<1046:TCOEPT>2.0.CO;2.

Boni, M.F., Gog, J.R., Andreasen, V., Feldman, M.W., 2006. Epidemic dynamics and antigenic evolution in a single season of influenza A. Proc. Biol. Sci. 273, 1307–1316, http://dx.doi.org/10.1098/rspb.2006.3466.

Cannell, J.J., Vieth, R., Umhau, J.C., Holick, M.F., Grant, W.B., Madronich, S., Garland, C.F., Giovannucci, E., 2006. Epidemic influenza and vitamin D. Epidemiol. Infect. 134, 1129–1140, http://dx.doi.org/10.1017/S0950268806007175.

Cazelles, B., Chavez, M., Berteaux, D., Ménard, F., Vik, J., Jenouvrier, S., Stenseth, N., 2008. Wavelet analysis of ecological time series. Oecologia 156, 287–304, http://dx.doi.org/10.1007/s00442-008-0993-2.

Cheng, C.K., Lau, E.H., Ip, D.K., Yeung, A.S., Ho, L.M., Cowling, B.J., 2009. A profile of the online dissemination of national influenza surveillance data. BMC Public Health 9, 339, http://dx.doi.org/10.1186/1471-2458-9-339.

Cheng, C.K.Y., Cowling, B.J., Lau, E.H.Y., Ho, L.M., Leung, G.M., Ip, D.K.M., 2012. Electronic school absenteeism monitoring and influenza surveillance, Hong Kong. Emerg. Infect. Dis. 18, 885–887, http://dx.doi.org/10.3201/eid1805.111796.

Choisy, M., Rohani, P., 2012. Changing spatial epidemiology of pertussis in continental USA. Proc. Biol. Sci. 279, 4574–4581, http://dx.doi.org/10.1098/rspb.2012.1761.

Crawley, M.J., 2007. The R Book. John Wiley & Son, Ltd.

Dushoff, J., Plotkin, J.B., Levin, S.A., Earn, D.J.D., 2004. Dynamical resonance can account for seasonality of influenza epidemics. Proc. Natl. Acad. Sci. U. S. A. 101, 16915–16916, http://dx.doi.org/10.1073/pnas.0407293101.

Gouhier, T., 2013. biwavelet: Conduct Univariate and Bivariate Wavelet Analyses (Version 0.14).

Grenfell, B.T., Bjørnstad, O.N., Kappey, J., 2001. Travelling waves and spatial hierarchies in measles epidemics. Nature 414, 716–723, http://dx.doi.org/10.1038/414716a.

Grenfell, B.T., Harwood, J., 1997. (meta)population dynamics of infectious diseases. Trends Ecol. Evol. 12, 395–399.

Hope-Simpson, R.E., 1981. The role of season in the epidemiology of influenza. J. Hyg. (Lond.) 86, 35–47.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Roy, J., Dennis, J., 1996. The NCEP/NCAR 40-year reanalysis project. Bull. Am. Meteorol. Soc. 77, 437–471, http://dx.doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Le, M.Q., Lam, H.M., Cuong, V.D., Lam, T.T.-Y., Halpin, R.A., Wentworth, D.E., Hien, N.T., Thanh, L.T., Phuong, H.V.M., Horby, P., Boni, M.F., 2013. Migration and persistence of human influenza A viruses, Vietnam, 2001–2008. Emerg. Infect. Dis. 19, 1756–1765, http://dx.doi.org/10.3201/eid1911.130349.

Leo, Y.S., Lye, D.C., Barkham, T., Krishnan, P., Seow, E., Chow, A., 2010. Pandemic (H1N1) 2009 surveillance and prevalence of seasonal influenza, Singapore. Emerg. Infect. Dis. 16, 103–105, http://dx.doi.org/10.3201/eid1601.091164.

Liu, P.T., Stenger, S., Li, H., Wenzel, L., Tan, B.H., Krutzik, S.R., Ochoa, M.T., Schauber, J., Wu, K., Meinken, C., Kamen, D.L., Wagner, M., Bals, R., Steinmeyer, A., Zgel, U., Gallo, R.L., Eisenberg, D., Hewison, M., Hollis, B.W., Adams, J.S., Bloom, B.R., Modlin, R.L., 2006. Toll-like receptor triggering of a vitamin D-mediated human antimicrobial response. Science 311, 1770–1773, http://dx.doi.org/10.1126/science.1123933.

Lowen, A.C., Mubareka, S., Steel, J., Palese, P., 2007. Influenza virus transmission is dependent on relative humidity and temperature. PLoS Pathog. 3, 1470–1476, http://dx.doi.org/10.1371/journal.ppat.0030151.

Nguyen, H.T., Dharan, N.J., Le, M.T.Q., Nguyen, N.B., Nguyen, C.T., Hoang, D.V., Tran, H.N., Bui, C.T., Dang, D.T., Pham, D.N., Nguyen, H.T., Phan, T.V., Dennis, D.T., Uyeki, T.M., Mott, J., Nguyen, Y.T., the Vietnam National Influenza Surveillance and Evaluation Team, 2009. National influenza surveillance in Vietnam, 2006–2007. Vaccine 28, 398–402, http://dx.doi.org/10.1016/j.vaccine.2009.09.139.

Nguyen, Y.T., Graitcer, S.B., Nguyen, T.H., Tran, D.N., Pham, T.D., Le, M.T.Q., Tran, H.N., Bui, C.T., Dang, D.T., Nguyen, L.T., Uyeki, T.M., Dennis, D., Kile, J.C., Kapella, B.K., Iuliano, A.D., Widdowson, M.-A., Nguyen, H.T., 2013. National surveillance for influenza and influenza-like illness in Vietnam, 2006–2010. Vaccine 31, 4368–4374, http://dx.doi.org/10.1016/j.vaccine.2013.07.018.

Rambaut, A., Pybus, O.G., Nelson, M.I., Viboud, C., Taubenberger, J.K., Holmes, E.C., 2008. The genomic and epidemiological dynamics of human influenza A virus. Nature 453, 615–619, http://dx.doi.org/10.1038/nature06945.

Ripley, B., 2012. tree: Classification and Regression Trees. R Package Version 1.0-33.

Russell, C.A., Jones, T.C., Barr, I.G., Cox, N.J., Garten, R.J., Gregory, V., Gust, I.D., Hampson, A.W., Hay, A.J., Hurt, A.C., de Jong, J.C., Kelso, A., Klimov, A.I., Kageyama, T., Komadina, N., Lapedes, A.S., Lin, Y.P., Mosterin, A., Obuchi, M., Odagiri, T.,

Osterhaus, A.D.M.E., Rimmelzwaan, G.F., Shaw, M.W., Skepner, E., Stohr, K., Tashiro, M., Fouchier, R.A.M., Smith, D.J., 2008. The global circulation of seasonal influenza A (H3N2) viruses. Science 320, 340–346, http://dx.doi.org/10.1126/science.1154137.

Shaman, J., Jeon, C.Y., Giovannucci, E., Lipsitch, M., 2011. Shortcomings of vitamin D-based model simulations of seasonal influenza. PLoS ONE 6, e20743, http://dx.doi.org/10.1371/journal.pone.0020743.

Shaman, J., Kohn, M., 2009. Absolute humidity modulates influenza survival, transmission, and seasonality. Proc. Natl. Acad. Sci. U. S. A. 106, 3243–3248, http://dx.doi.org/10.1073/pnas.0806852106.

Shaman, J., Pitzer, V.E., Viboud, C., Grenfell, B.T., Lipsitch, M., 2010. Absolute humidity and the seasonal onset of influenza in the continental United States. PLoS Biol. 8, e1000316, http://dx.doi.org/10.1371/journal.pbio.1000316.

Stephenson, I., Zambon, M., 2002. The epidemiology of influenza. Occup. Med. 52, 241–247.

Tamerius, J., Nelson, M.I., Zhou, S.Z., Viboud, C., Miller, M.A., Alonso, W.J., 2011. Global influenza seasonality: reconciling patterns across temperate and tropical regions. Environ. Health Perspect. 119, 439–445, http://dx.doi.org/10.1289/ehp.1002383.

Tamerius, J.D., Shaman, J., Alonso, W.J., Bloom-Feshbach, K., Uejio, C.K., Comrie, A., Viboud, C., 2013. Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. PLoS Pathog. 9, e1003194, http://dx.doi.org/10.1371/journal.ppat.1003194.

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79, 61–78.

Viboud, C., Alonso, W.J., Simonsen, L., 2006a. Influenza in tropical regions. PLoS Med. 3, e89, http://dx.doi.org/10.1371/journal.pmed.0030089.

Viboud, C., Bjørnstad, O.N., Smith, D.L., Simonsen, L., Miller, M.A., Grenfell, B.T., 2006b. Synchrony, waves, and spatial hierarchies in the spread of influenza. Science 312, 447–451, http://dx.doi.org/10.1126/science.1125237.

Xia, Y., rnstad, O.N.B., Grenfell, B.T., 2004. Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. Am. Natur. 164, 267–281, http://dx.doi.org/10.1086/422341.

Yamshchikov, A.V., Desai, N.S., Blumberg, H.M., Ziegler, T.R., Tangpricha, V., 2009. Vitamin D for treatment and prevention of infectious diseases: a systematic review of randomized controlled trials. Endocr. Pract. 15, 438–449, http://dx.doi.org/10.4158/EP09101.ORR.

Yu, H., Alonso, W.J., Feng, L., Tan, Y., Shu, Y., Yang, W., Viboud, C., 2013. Characterization of regional influenza seasonality patterns in China and implications for vaccination strategies: spatio-temporal modeling of surveillance data. PLoS Med. 10, e1001552, http://dx.doi.org/10.1371/journal.pmed.1001552.