Modeling Influenza Seasonality in the Tropics and Subtropics

Haokun Yuan[1*], Sarah C. Kramer[2*#], Eric H. Y. Lau[3,4], Benjamin J. Cowling[3,4], Wan Yang[1**]

[1]*Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY 10032*

[2]*Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032*

[3]*WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, Li Ka Shing Faculty of Medicine, the University of Hong Kong, Hong Kong Special Administrative Region, China*

[4]*Laboratory of Data Discovery for Health Limited, Hong Kong Science Park, New Territories, Hong Kong Special Administrative Region, China*

[*] These authors contribute equally to this work. [#]Current address: Infectious Disease Epidemiology group, Max Planck Institute for Infection Biology, Charitéplatz 1, Campus Charité Mitte, 10117 Berlin, Germany

**Correspondence to:
Wan Yang
Department of Epidemiology
Mailman School of Public Health
Columbia University
722 W 168th Street, Room 514, New York, NY 10032
Phone: (212) 305-0421
Email: wy2202@columbia.edu

Data and model code availability: Compiled daily temperature and specific humidity data along with model code are available at https://github.com/wan-yang/flu-subtropic-climate-models. Influenza data (after 2014) are available from https://www.chp.gov.hk/en/statistics/data/10/641/642/2274.html

Abstract

Climate drivers such as humidity and temperature may play a key role in influenza seasonal transmission dynamics. Such a relationship has been well defined for temperate regions. However, to date no models capable of capturing the diverse seasonal pattern in tropical and subtropical climates exist. In addition, multiple influenza viruses could cocirculate and shape epidemic dynamics. Here we construct seven mechanistic epidemic models to test the effect of two major climate drivers (humidity and temperature) and multi-strain co-circulation on influenza transmission in Hong Kong, an influenza epidemic center located in the subtropics. Based on model fit to long-term influenza surveillance data from 1998 to 2018, we found that a simple model incorporating the effect of both humidity and temperature best recreated the influenza epidemic patterns observed in Hong Kong. The model quantifies a bimodal effect of absolute humidity on influenza transmission where both low and very high humidity levels facilitate transmission quadratically; the model also quantifies the monotonic but nonlinear relationship with temperature. In addition, model results suggest that, at the population level, a shorter immunity period can approximate the co-circulation of influenza virus (sub)types. The basic reproductive number R_0 estimated by the best-fit model is also consistent with laboratory influenza survival and transmission studies under various combinations of humidity and temperature levels. Overall, our study has developed a simple mechanistic model capable of quantifying the impact of climate drivers on influenza transmission in (sub)tropical regions. This model can be applied to improve influenza forecasting in the (sub)tropics in the future.

Key words: Influenza; climate; seasonality; co-circulation; temperature; humidity; tropics; subtropics

Introduction

Influenza is a disease of considerable public health concern, causing roughly 300,000-650,000 deaths and 3-5 million cases of severe illness each year worldwide [1]. Although evidence suggests that the burden of influenza in the tropics and subtropics is not substantially less than in temperate regions [2, 3], studies on influenza in these regions are

comparatively rare. Likewise, modeling and forecasting efforts, which may promote both understanding of and preparation for outbreaks in the future, have mostly been focused on countries with temperate climates [4, 5]. A variety of factors contribute to this disparity, including lack of long-term surveillance data and competing public health interests [6-8]. We focus here on two features of influenza epidemics in these regions that particularly complicate modeling efforts, i.e., the lack of understanding of climatic drivers and cocirculation of multiple influenza types and subtypes.

In temperate regions, influenza displays a clear seasonal pattern, with epidemics occurring in the winter and very few cases observed during the summer [9, 10]. While several potential drivers of this pattern have been suggested [9], humidity appears to be particularly important in driving these "seasonal" influenza epidemics. Specifically, both survival and transmission of the influenza virus are heightened when absolute humidity (AH) is low [11], corresponding to the yearly observed peaks of influenza activity in winter [12].

In tropical or subtropical locations, this seasonal pattern is less commonly observed. Instead, influenza causes multiple epidemics each year, or else is present year-round with unpredictable variation in intensity [3, 9, 10, 13]. In addition, humidity is relatively high all year in the tropics and subtropics, and influenza epidemics tend to occur during the rainy season, when humidity is particularly high [13, 14]. Thus, the relationship observed in temperate regions and modeled by Shaman et al. [12], where influenza transmission decreases monotonically with increasing AH, is not sufficient to explain patterns in influenza transmission in the tropics and subtropics [9, 15].

While the exact impacts remain unclear, humidity, precipitation and temperature are the main contenders as climate drivers for influenza transmission in the tropics and subtropics. A few studies suggest that the impact of humidity on influenza transmission may be bimodal, rather than unimodally decreasing as suggested by Shaman at al. [11, 12]. Work by Yang et al. showed that influenza virus survival is higher at lower relative humidity (<50%), but also found increased survival at very high levels of relative humidity

(~100%) [16]. By analyzing patterns of influenza transmission in 78 locations worldwide, Tamerius et al. [14] found that influenza outbreaks in locations that experience high temperature and high AH year-round had a tendency to occur during the rainy season, when both AH and rainfall were high. This pattern has been consistently reported in several countries [6, 17-19]. Deyle et al. [20] instead suggested that influenza transmission is driven by AH, moderated by temperature. Specifically, they find that influenza activity decreases with increasing AH up to about 24°C (75 °F), then increases with increasing AH up to about 30°C (86°F), at which point high temperatures strongly restrict influenza transmission.

In addition to diverse climate drivers, co-circulation of multiple influenza viruses may contribute to the observed multiple yearly influenza epidemics in the tropics and subtropics. Circulating human influenza viruses are classified into types (A or B, based on genus) and, among influenza A viruses, subtypes (based on the genetic sequences of the hemagglutinin and neuraminidase surface proteins) [21]. Currently, circulating human influenza viruses consist of the A(H1N1) and A(H3N2) influenza subtypes, as well as influenza B viruses [1]. While influenza viruses evolve quickly, significant antigenic change only occurs over a period of one to ten years [22, 23]. However, while the extent of cross-immunity between influenza (sub)types remains unknown, current influenza vaccines are not sufficient to protect people from all (sub)types of influenza virus [21, 24]. Thus, we may expect that separate epidemics within a single year are often caused by different influenza viruses. As a result, appropriate models of influenza transmission in tropical and subtropical locations may need to take multiple types and subtypes into account, further complicating modeling efforts.

Here, we utilize influenza incidence data that have been collected since 1998 in Hong Kong, a densely populated city with a subtropical climate, to explore the impact of climate drivers on influenza epidemics. We formulate seven models, each allowing for differing roles of humidity, temperature, and influenza co-circulation. We expect that, by accounting for 1) increased transmission at both low and high values of AH, 2) decreased transmission at high temperatures (e.g., >30 °C), and 3) co-circulation of several influenza

types and subtypes, the model best representing the impact of key climate drivers and influenza co-circulation will be able to best reproduce observed influenza dynamics in Hong Kong. In addition, influenza pandemics could occur "off-season" under more extreme climate conditions (e.g., hot summer days), due to higher population susceptibility to a novel virus. Thus, we include the 2009 pandemic in our model testing and expect the best-performing model to also capture influenza dynamics during the pandemic after accounting for the increased population susceptibility. To model the impact of humidity, here we focus on AH as it is independent of temperature. Specifically, unlike relative humidity measuring the amount of water vapor (i.e. moisture) in the air relative to the total amount of vapor that can exist in the air at its current temperature, AH measures the actual amount of water vapor in the air irrespective of the air's temperature.

Methods

Influenza Data: Data on influenza-like illness (ILI) and laboratory-confirmed influenza from January 1998 through December 2018 were obtained from the Centre for Health Protection of the Hong Kong Special Administrative Region. ILI data were collected by a sentinel surveillance network consisting of 64 public out-patient clinics and roughly 50 private medical practitioners' clinics throughout Hong Kong, while laboratory testing was performed on specimens from outpatient clinics and public hospitals. Throughout our study period (21 years from January 1998 to December 2018), the same procedure for selecting specimens for viral testing was used; however, from February 10, 2014 onward, viral testing was carried out using molecular testing instead of viral culture. We multiply weekly ILI case counts by the proportion of tests positive for influenza each week and refer to the resulting, more specific measure as ILI+. Finally, data were converted to rates per 100,000 population.

Climate Data: Hong Kong has a humid subtropical climate, with hot, humid, and rainy summers, and mild winters [25]. Daily mean temperature and relative humidity were taken at the Hong Kong Observatory [26]. Using these data, we used the Clausius-Clapeyron relation [27] to calculate daily mean specific humidity, a measure of AH (see Supplementary Materials).

Model Hypotheses: In this work, we formulated seven models to test several hypotheses concerning the impact of humidity and temperature on influenza transmission in Hong Kong, and how co-circulation influenza (sub)types affect this process. While specific methodology associated with each model is detailed below, we describe the hypotheses considered here:

- Null hypothesis 1 (or Null1, Constant basic reproductive number R_0): Climate conditions have no effect on influenza transmission. To represent this, we held R_0, the epidemic parameter representing the average number of new cases arising from a primary case in a fully susceptible population, constant.
- Null hypothesis 2 (or Null2, Temperate Forcing): AH affects influenza transmission in the subtropics and tropics in the same manner as in temperate regions; that is, R_0 increases monotonically with decreasing AH. Here, R_0 was modeled according to the equation first presented in [11]. This model has previously been shown to perform well when modeling influenza transmission patterns in temperate regions [12].
- AH model: The effect of AH is bimodal, with both low and high AH conditions favoring influenza transmission. Temperature is assumed to have no additional effect on transmission. R_0 was modeled according to Eqn S6, where both low and high values of AH lead to higher values of R_0.
- AH/T model: Absolute humidity has a bimodal effect on R_0, as in the AH model, but this effect is moderated by temperature as shown in Equation 3. Briefly, low temperatures promote influenza transmission, and temperatures above a certain threshold limit transmission.
- AH/T/Strain: AH and temperature impact R_0 as in the AH/T model. Additionally, two "strains" of influenza co-circulate in the population, as in [12]. Here we ignore cross-immunity between strains; in other words, infection with one strain has no effect on the potential for later infection with the other strain.
- AH/T/Short: Climate forcing is included as in the AH/T model. Additionally, we restrict the duration of immunity in the model to be about 1 year (i.e., 0.5 – 1.5 years), in order to implicitly take co-circulation into account (i.e., for an individual, multiple infections by different influenza strains could occur within a short time span).

- AH/T/Vary: Climate forcing is included as in the AH/T model. Additionally, a small proportion (<50%) of the population experiences a significantly truncated duration of immunity (<1 year) after infection. In other words, duration of immunity to influenza is heterogeneous among the population.

SIRS Model: We modeled influenza transmission in Hong Kong using a compartmental susceptible-infected-recovered-susceptible (SIRS) model with demography. While we used two distinct forms of the model for this project described below, the basic model takes the form:

$$\begin{cases} \frac{dS}{dt} = \frac{N-S-I}{L} - \frac{\beta_t I^p S}{N} - \alpha + \mu(N-S) \\ \frac{dI}{dt} = \frac{\beta_t I^p S}{N} - \frac{I}{D} + \alpha - \mu I \end{cases} \quad [1]$$

where N is the size of the model population (here set to 100,000); β_t represents the rate of transmission on day t, depending on climatic functions as described below; D is the mean infectious period; L is the average duration of immunity; and α represents random seeding from outside the model population (here we arbitrarily set it to 0.1, i.e. 1 per 10 days for all models). μ is the rate of natural birth and death (i.e., to maintain a constant population size, we assume equal birth and death rates). We set μ to be 0.00918 divided by 365, or the average daily birth rate per person in Hong Kong during 1998 – 2017 [28]. The parameter p is an exponent to introduce nonlinearity into the infection process (i.e. imperfect population mixing). The inclusion of this parameter has been shown to be helpful in modeling complicated epidemics using very simple models [29], as is the case in this work. Finally, because the 2009 pandemic was caused by a novel strain of influenza with little prior population immunity, we reset the number of susceptibles in our population in early August of 2009 to be between 60% and 80% of the total population [30-32] (vs. 40-80% for model initiation in January 1998; see Table 1). All models were run stochastically, as described in the Supplementary Materials and in [12].

SIRS Model Variations: In this work, we considered three models to account for cocirculation and heterogenous immunity period:

1) **AH/T/Strain:** To account for co-circulating influenza viruses, we ran two SIRS simulations (Eqn 1) in parallel, as done in Shaman et al. [12]. While there are three co-circulating influenza (sub)types, we combined A(H1N1) and B for simplicity, given the similar timing of the circulation of these two viruses (Fig S1). Here, we applied random seeding at each week only to the "strain" that had the greater number of positive tests that week. If both "strains" were equally common, seeding was applied to a single strain at random. We then combined the output of the two simulations and compared the resulting estimates to the overall ILI+ data. Note that per this simple model, multiple influenza epidemics due to different influenza viruses can occur at the same time, as well as co-infection of multiple influenza viruses.

2) **AH/T/Short:** This model used Equation 1 but restricted L, the duration of immunity, to be between around one year to account for multiple infections within a year due to multiple circulating strains.

3) **AH/T/Vary:** In this model, a small proportion of the population loses immunity to influenza at an accelerated rate. This was modeled by replacing the term $\frac{N-S-I}{L}$ in Equation 1 with:

$$\rho \frac{\beta_{t-L_s} I^p_{t-L_s} S_{t-L_s}}{N} + (1-\rho)\frac{N-S-I}{L} \qquad [2]$$

where ρ is the proportion of the model population that loses immunity after a short period L_s and $\beta_{t-L_s} I^p_{t-L_s} S_{t-L_s}$ represents the number of new infections L_s days ago, who would lose

their immunity on day *t*. We note that, since longer-term immunity is lost at an exponential rate, while short term immunity simply removes a set number of individuals from the recovered compartment at each time step, there is a possibility of double-counting. However, given the difference in time scales (less than one year vs. several years), we do not anticipate that this approximation will lead to severe problems.

Climate Forcing Models: Based on past work [12, 14, 16, 20], as well as the patterns observed in the influenza and climate data described above and shown in Fig 1, we modeled the impact of AH using a parabola, where transmissibility is highest at very low and very high levels of AH. In all AH/T models, this relationship was modified by temperature such that, when temperature is above some cutoff value, transmissibility is reduced. Specifically, for models AH/T, AH/T/Strain, AH/T/Short, and AH/T/Vary, the impact of AH and temperature on transmissibility was modeled as:

$$R_0(t) = [aq^2(t) + bq(t) + c][\frac{T_c}{T(t)}]^{T_{exp}} \quad [3]$$

where $q(t)$ is specific humidity (i.e. a measure of AH) at time t, and $T(t)$ is temperature at time t. The parameter β_t in the SIRS model is defined as $R_0(t)$ divided by D; thus, AH and temperature act through β in Equations 1 and 2 to influence transmission patterns. When T is below T_c, lower temperatures are able to further increase R_0, whereas temperatures above T_c inhibit influenza transmission. However, the favorable impact of low temperature may level off at relatively low temperatures. Thus, we truncated this monotonic relationship at a minimum temperature of $T_c - T_{diff}$, beyond which, the effect levels off. The strength of this relationship is further determined by the exponent T_{exp}. For the AH Only model, AH was also modeled as a parabola (i.e. the terms within the first set of squared brackets), but with no impact of temperature (Eqn. S6).

To link the coefficients a, b, and c to the AH q and R_0, we reparametrize them by solving the parabola with the nadir at (q_{mid}, R_{0min}) and maximum at both (q_{min}, R_{0max}) and (q_{max}, R_{0max}) (see Supplemental Materials).

Model optimization (parameter tuning): To assess the validity of the models, we first split the influenza data into a training set (Jan 1998 – Dec 2012; 15 years) and a testing set (Jan 2013 – Dec 2018; 6 years). That is, data from 1998 to 2012 were used to train the model and optimize parameters, the remaining data from 2013 to 2018 were holdout for testing.

We optimized each model by testing a wide range of parameter values (Table 1) based on the literature [12, 22, 23, 33-39], and tuning parameter range against the influenza data observed during Jan 1998 – Dec 2012 (i.e. the training period). Specifically, for each model, we draw 1 million parameter combinations from ranges listed in Table 1 using Latin hypercube sampling [12]. Using each parameter combination, we ran each model stochastically from Jan 1998 to Dec 2012 with a daily time step and aggregated the simulated daily ILI+ to weekly intervals for model assessment. To account for model stochasticity, we repeated the simulation 500 times for each parameter set. We then calculated, for each model run, the corresponding root mean square error (RMSE) and correlation against the full training dataset and weekly averaged dataset (i.e., for each of the 52 weeks of the year, averaged the ILI+ over 15 training years), separately; We refer to these metrics as *full.RMSE*, *avg.RMSE*, *full.Correlation* and *avg.Correlation,* respectively, hereafter. We further averaged across the 500 model runs for each parameter combination to obtain a single set of metrics for the corresponding parameter combination. To combine the metrics and simplify the process of parameter selection, we averaged the two RMSE metrics (RMSES) and the two correlation metrics (CORR), separately, as:

$$RMSES = 0.5\ full.RMSE + 0.5\ avg.RMSE$$

and:

$$CORR = 0.5\ full.Correlation + 0.5\ avg.Correlation$$

Based on these two final metrics, we selected the top 1000 parameter combinations out of the one million for each model: the 1000 with the highest *CORR* among those with RMSES lower than the 0.5 percentile. We then computed the 95% highest density interval (HDI) for each parameter using those top 1000 to generate new parameter ranges for subsequent round of optimization. That is, we updated the range of a parameter with the HDI, if the upper (or lower) bound of HDI is 10% smaller (or larger) than the corresponding bound of the original range. This tuning process was repeated until the parameter range no longer shrinks substantially (i.e. by 10%). In this study, it took two rounds of such parameter selection; at each round, we checked to ensure the model fit was improving using the new parameter ranges.

After obtaining the final parameter ranges, we drew 10,000 random parameter combinations from these final ranges and ran the models from Jan 1998 – Dec 2018 using each parameter combination. In this last batch, we obtained the top 1000 parameter combinations as described above for further assessment and model comparison.

Model assessment: We evaluated performance of the seven models based on the four metrics described above (i.e., *full.RMSE, avg.RMSE, full.Correlation, avg.Correlation*) during both the training and testing period. To quantify the differences in performance among models, we applied the Kruskal-Wallis test to test if the mean rank of different models were similar for each metric. Since significant differences were found between models, we further performed a pairwise comparison using the Nemenyi test with an adjusted p-value of 0.007 (i.e., 0.05 /7). We summarized the model ranking in Table 2 based on model performances including the pandemic period. In addition, as a sensitivity analysis, we ranked the models based on their performances excluding the 2009 pandemic in Table S1.

Results

General influenza transmission dynamics in Hong Kong. Influenza activity is highly diverse in Hong Kong. Unlike the common wintertime epidemics in temperate regions, during our study period (Jan 1998– Dec 2018), Hong Kong experienced two epidemics per year in most years (17 out of 21; the four exceptions were: Year 2001, 2011, 2012 and

2018; Fig 1A). For years with two epidemics, one epidemic typically occurred during the winter, and the other, usually smaller in magnitude, occurred during the summer (Fig 1A). Usually, there was a single influenza (sub)type predominating during each epidemic, although co-circulation of multiple strains with comparable magnitudes could occur (Fig 1A). While any (sub)type can dominate winter epidemics, summer epidemics were almost always caused by A(H3N2) (15 out of 17 years with summer epidemics; Fig 1A).

Climate conditions in Hong Kong overall and during influenza epidemics. Both mean daily temperature and AH in Hong Kong displayed strong seasonality, where temperature and AH were highest in summer and lowest in winter (Fig 1B). In the hottest days, the mean temperature in Hong Kong can exceed 30°C (around six days per year), but most of the time, mean temperature was between 15°C to 30°C. Hong Kong was also very humid. Compared to the relatively drier and cooler climates in temperate regions, the humid climate in Hong Kong may provide favorable conditions for influenza transmission during the summer despite the high temperatures. In addition, temperature and AH were highly correlated (Pearson's $\rho = 0.94$, $p < 0.001$). Such a high correlation could confound the potential effect on influenza activity due to either climate variable and suggests that simple linear models will not be able to separate such effects.

Models incorporating the impact of humidity and temperature best replicate the influenza dynamic in Hong Kong. In this work, we designed seven models and performed two rounds of parameter selection to obtain the optimized HDI for each parameter in every model, and eventually compared the models using the selected parameter combinations. Among the seven models constructed, the Null1 model served as a “control” model since it had a constant R_0 that is independent of any climate factors. In other words, if a model failed to outperform the Null1 model, then the proposed relationship between climate factors and the influenza transmission dynamics would not be supported. In addition, we note, again, that we included the 2009 influenza pandemic in our main analysis, because the off-season occurrence of pandemic influenza provided a unique opportunity to test the climate forcing models under more extreme climate conditions. Table 2 shows the model

ranking including the pandemic and Table S1 excluding the pandemic. Results are consistent across the two analyses.

Overall, all models incorporating seasonality outperformed the Null1 model, for which R_0 is set to a constant (mean testing rank: 6.5 out of seven models; Table 2 and Fig 2), suggesting the seasonal variation of influenza activity. In addition, while the AH model (i.e. without including temperature as a variable) outperformed the Null1 model, it was consistently inferior when compared to other models (mean testing rank: 5.5 out of 7 models; Table 2, Fig 2). The model failed to capture much of the dynamics observed in Hong Kong, especially in colder months (Fig S2C). This suggests that, in addition to AH, temperate can modulate influenza transmission in Hong Kong.

To test if the monotonic increase in transmission with decreasing AH as observed in temperate regions also applies to the subtropics, we also tested the model by Shaman et al. (i.e. the Null2 model here). The Null2 model was able to better recreate the observed transmission dynamics in Hong Kong than the Null1 and AH model. However, its performance during the training period ranked lower than models additionally incorporating increasing transmission under high AH conditions (i.e. a bimodal relation) (Table 2).

Models incorporating the impact of both AH and temperature improved the model performance significantly (Fig 2). The AH/T models (AH/T, AH/T/Short, AH/T/Vary, and AH/T/Strain) had the best performance among all. Among these four models, the AH/T and AH/T/Short models assumed the same model construct but with different initial ranges for the immunity duration (L). After rounds of parameter selection, the HDI of L eventually converged for the two models (Table 1), as did other parameters. Therefore, we considered these two models to be the same. Indeed, either model ranked among the top two during the training period. However, the AH/T model appeared to be more stable and remained as a top performing model during the testing period (Table 2).

Two models (i.e., AH/T/Vary and AH/T/Strain) additionally accounted for long/short-term immunity and/or co-circulation of different influenza viruses. Both models performed comparably to the AH/T model but not better (Fig 3). In addition, for the common parameters, the HDIs of the three models converged to similar ranges (Table 1). The AH/T/Vary model, although formulated differently, had similar HDI and performances as the AH/T model. The AH/T/Strain model was the only model that specifically accounted for co-circulation of different influenza viruses. Although the model was able to simulate ILI+ with high correlation with the observed data and recreate large epidemics, it did not fit smaller epidemics as well and thus had a high RMSE (Fig 3).

Taken together, the AH/T model had the most consistent and best performance over the entire study period. It is also the most parsimonious model among all top performing ones.

Impact of humidity and temperature on influenza transmission, as estimated by the top models. To account for the impact of humidity on influenza transmission, we included three parameters in the models to describe a bimodal relationship. That is, below a threshold AH (q_{mid}), transmission would increase as AH decreases and level off when AH is $\leq q_{min}$, whereas above that threshold, transmission would increase as AH increases and level off when AH is $\geq q_{max}$. Tamerius et al 2013 [14] defined a similar bimodal relationship and found a threshold of 11-12 g/kg for regions with single-peak vs bimodal epidemics. In addition, they found that most influenza peak activities occurred in "cold-dry" conditions (i.e., AH < 8 g/kg) or "humid-rainy" conditions (AH > 14 g/kg). Consistent with Tamerius et al., we estimated the threshold q_{mid} to be 10 – 12.6 g/kg, where R_0 troughs (Fig 4). In addition, we further quantified how influenza transmission (i.e. as indicated by R_0 estimates) changes with AH in the two regimes (Eqn 3). We estimated that when AH is below 10-12.6 g/kg, R_0 would increase with decreasing AH quadratically up to a minimum of 2 – 4 g/kg; when AH is above 10-12.6 g/kg, R_0 would increase with increasing AH quadratically up to a maximum of 17-20 g/kg (Table 1).

In addition to AH, we included three parameters to model the impact of temperature on influenza transmission. Foremost, as described in Eqn 3, the impact of temperature is modeled as a multiplicative adjustment to the impact of AH. Below a threshold temperature T_c, decreasing temperature increases transmission up to a minimum of $T_c - T_{diff}$; above T_c, increasing temperature reduces transmission. Deyle et al. [20] suggested temperature around 24°C (75°F) as a threshold dividing the negative and positive effect of AH on transmission. Similarly, we estimated T_c to have an HDI between 20-24°C (Table 1). As shown in Fig 4, when temperature is below T_c, low temperature further facilitates transmission, in addition to the favorable transmission conditions due to concurrent, low AH in cold months. However, when temperature exceeds T_c, high temperature suppresses transmission and lowers the overall R_0 despite the favorable transmission conditions due to concurrent, very high AH in the summer. In addition, we estimated T_{diff} to have an HDI of 0.4-5.3°C, which implies temperature below 16-23°C does not afford additional increases in transmission. Similar transmission behavior was also found by Brown et al. [33] with avian influenza viruses, whose infectiveness stabilized after the temperature dropped below 17°C. The estimated value of T_{exp}, the exponent of the T_c/T ratio (Eqn 3), had an HDI of 0.95-1.54, suggesting the moderation of temperature is slightly super-linear.

We also estimated the duration of immunity (L) with our models to better understand transmission dynamics. The immunity period L directly affects the frequency of influenza epidemics – i.e., shorter L would lead to more frequent epidemics. Here we estimated L to be around 1-1.3 years for the AH/T model, and 1.5-2.3 year for AH/T/Strain. For the AH/T/Vary model, considering population heterogeneity, we included another parameter L_s to represent short-term immunity (<1 year). We found L_s was around 3-11 months and L around 1-1.6 years; combing these two parameters (i.e., L, Ls), the average is very similar to the L estimated by the AH/T model. Overall, these L estimates are consistent with the frequent influenza epidemics observed in Hong Kong; however, we note that these estimates are on the lower end of the reported 1-10 year range [22].

Validate model findings using laboratory results. The reproductive number represents the epidemic potential of an infection. To examine if the identified relationship with AH and temperature is consistent with data from laboratory studies, we further compared the basic reproductive number R_0 and effective reproductive number R_e calculated by our model (Eq 3) with the influenza virus survival rate in aerosols [35] as well as the influenza transmission rate observed in guinea pigs [39, 40]. At the population level, epidemic can occur when R_e is above 1 and will subside when R_e drops below 1. Indeed, no transmission occurred in the guinea pig studies [40] under conditions where the estimated R_e per our model (Eqn 3) was below 1; in contrast, transmission occurred when estimated R_e was above 1, ranging 1.21-2.25 (corresponding R_0 range: 1.68-3.11; Fig. 5). In addition, fitting a linear regression of estimated R_0 (or R_e) against the log survival rate and log transmission probability, we found that estimated R_0 (or R_e) positively correlated with both laboratory-observed survival rate and transmission rate of the influenza virus (Fig 5). In particular, the correlation between R_0 (or R_e) estimated by our AH and temperature model (Eqn 3) and the observed transmission rate was 0.82 (r^2 = 0.67; Fig 5), suggesting it was able to explain around 70% of the variances of the transmission rate. These findings indicate that our model is able to represent both the impact of AH and temperature on influenza epidemic patterns at the population level as well as influenza transmission and survival observed in laboratory settings.

Discussion

Despite influenza's profound impact on public health, little is known about how the virus transmits from person to person and what environmental and climate conditions make this process more likely. To date, there have been only a few studies on the influenza burden in tropical and subtropical regions. Of those studies, a handful analyzed the effect of climate drivers on influenza incidence using time-series or logistic regression models. Although those statistical models were able to identify significant climate covariates (including AH, temperature, and precipitation) and estimate the effect size, they did not provide much information to the underlying mechanism of how those climate covariates affect transmission. In contrast, infectious disease models (e.g., the SIRS model) provide a means to model and test the relationship between climate factors and influenza transmission

observed in laboratories. In this work, by building seven SIRS models under different hypotheses, we explored how temperature, AH, and influenza co-circulation affect the transmission of influenza. We showed that models that included both AH and temperature as covariates consistently outperformed those that did not in recreating the observed influenza epidemic patterns in Hong Kong. These results support that climate variables play a critical role in influenza transmission, where temperature is particularly influential in moderating the impact of AH. Model results also indicate that the effect of climate drivers in tropical and sub-tropical regions is different from those observed in the temperate regions, and models built for temperate areas will not be sufficient to reproduce the transmission patterns in the tropics or the subtropics.

Previous studies had suggested a bimodal effect of AH on influenza transmission [20]. However, such an effect has not been quantified nor incorporated in influenza transmission models as done for temperate regions [12]. Here, we have developed a model that can effectively represent this U-shaped relationship with AH, moderated further by temperature. When incorporated into an SIRS model, the combined model was able to recreate the diverse influenza epidemic dynamics observed in Hong Kong over a 21-year period. As described earlier, the breaking point between the "cold-dry" and "hot-humid" environment is an AH of approximately 10-12.6 g/kg, which is consistent with Tamerius et al 2013 [14]. This finding helps to explain the biannual epidemics observed in Hong Kong. In the winter when weather is "cold-dry," the influenza transmission rate increases as the environment gets drier. On the other hand, summer in Hong Kong is hot and humid – the very high AH also promotes influenza transmission, despite high temperatures.

The combined relationship of influenza transmission with both AH and temperature we delineated in Fig 4 also consistently explains observations in temperate regions where influenza surges predominantly in winter and increases monotonically with decreasing AH. During summers in temperate areas, absolute AH is typically lower than 15 g/kg (vs. as high as 23g/kg in subtropics/tropics) [12], and temperature is relatively high. Under such environmental conditions, R_0 would stay in the trough during the summer (Fig 4), and only increases in the winter when AH becomes low (Fig 4). As such, it appears that influenza

transmission rate increases monotonically with decreasing AH, leading to a single epidemic each year in the winter in temperate regions.

When there are multiple strains co-circulating, individuals who recover from infection gain specific immunity against the infecting strain and may remain susceptible to other influenza strains in co-circulation either during the same or subsequent epidemics. In this study, we combined data from all three co-circulating influenza viruses and, as such, our estimated duration of immunity, without distinguishing specific strains, was relatively short (~1–1.5 years). In comparison, the AH/T/Strain model which modeled H3N2 and H1N1/B epidemics separately, estimated a longer immunity period (1.5-2.5 years). Nevertheless, either estimate is quite low compared to the duration an influenza clade in circulation (1-10 years). Therefore, we note that the short immunity duration estimated here is more to capture the frequent influenza epidemics in Hong Kong. For locations with less frequent influenza epidemics, re-estimating the immunity duration per local epidemic data is warranted.

We recognize several limitations of our study. First, although our models fit the observations in Hong Kong well and provide support to the role of AH and temperature on influenza transmission, it is important to note that further laboratory and epidemiologic works are needed to establish a causal relationship. Second, while we accounted for multi-strain co-circulation to some extent (e.g., the AH/T/Short and the AH/T/Strain model), our models are highly simplified and do not distinguish the different epidemiological characteristics for H3N2, H1N1, and B; as such, our estimate of the immunity period may not reflect the true value. Third, for model simplicity, we did not model the interactions between (sub)types, which can modify the association with climate variables [41]. Future research is necessary to explore the impact of co-circulation, as a reasonable inclusion of multiple influenza (sub)types may be more appropriate than combining all strains. Fourth, as a first step, here we combined all influenza viruses without separating influenza by subtype or type. Previous studies have reported differential responses of different influenza strains to climate variables [41]. Future work could also model the impact of

536 climate variables for each influenza virus separately to examine such potential differences
537 and the impact on influenza epidemic dynamics at the population level.
538
539 Our study also sidestepped several key factors that may shape influenza transmission
540 dynamics, in particular, age, vaccination, and seasonal changes in contact patterns. Due to
541 a lack of age-specific incidence data, we did not include age structure in our models.
542 Vaccination rate in Hong Kong has increased in recent years, particularly among children
543 aged <12 years and adults 65 years or older, due to vaccination subsidy provided to these
544 age groups from 2008 onwards [42-44]. For instance, vaccination rate among children aged
545 <12 years increased from 9.7% in 2011/12 to 55.4% in 2018/19; and vaccination rate
546 among adults 65+ increased from 31.7% in 2011/12 to 43.6% in 2018/19 [45, 46]. These
547 increases in vaccination coverage may in part explain the apparent decreases in ILI+ in
548 recent years; a one-sided *t*-test indicated that the yearly ILI+ during 2011-2018 were
549 significantly lower than years before 2009 (i.e., excluding the 2009 pandemic period;
550 p=0.009). However, we do not have detailed data on vaccination rate nor vaccine efficacy
551 over the entire study period in order to account for the impact of vaccination. Further,
552 seasonal changes in contact patterns could also contribute to the observed seasonal
553 influenza epidemic dynamics in Hong Kong. For instance, during a hot, humid summer,
554 people may spend more time indoors and thus create additional opportunities for onward
555 transmission to occur. Similar increases of indoor contact and transmission could occur
556 during colder days in the winter. As such, the increased transmission under highly humid
557 conditions in the summer or colder-drier conditions in the winter may be in part due to
558 increased contact indoors, in addition to the higher survival rate of influenza viruses under
559 these environmental conditions as shown in laboratory studies [15, 16, 39]. Our modeling
560 here is not able to tease this apart. Future work is warranted to incorporate these
561 additional factors that may further improve understanding of influenza transmission
562 dynamics in tropical and subtropical climates. Nevertheless, we note that, all climate
563 forcing models tested here are subject to the same limitations discussed above. Thus, the
564 relative performance of the models and our general findings regarding key climate drivers
565 still hold.
566

In summary, we have developed a simple mechanistic model incorporating the impact of AH and temperature on influenza transmission that is able to recreate the long-term influenza epidemic dynamics in Hong Kong, a subtropical city with highly diverse epidemic patterns. Past work has demonstrated that, incorporating information on AH into mechanistic influenza models significantly improved model fit and forecast accuracy in temperate regions [11]. Given that forecasts in the tropics and subtropics tend to be substantially less accurate than forecasts in temperate regions [36], the climate forcing model developed here could support a better understanding of climatic drivers of influenza transmission in these regions. Future work could incorporate this model into a SIRS model-forecast system to improve forecast accuracy in the tropics and subtropics to aid public health and medical workers in better anticipating influenza transmission in forthcoming weeks.

Acknowledgements:

We thank the Hong Kong Center for Health Protection for providing data on sentinel surveillance and laboratory surveillance for influenza. We also thank Columbia University Mailman School of Public Health for access to high performance computing. **Funding:** HY and WY were supported by the NIH (AI135926 and ES009089). SCK was supported by the NIH (T32ES023770 and F31AI138410). EHYL and BJC were supported by the Theme-based Research Scheme project no. T11-712/19-N, the University Grants Committee of the Hong Kong Government.

Conflict of Interest: BJC reports receipt of honoraria from Roche and Sanofi. The authors declare that they have no other potential competing interests.

References:

1. World Health Organization. *Influenza (Seasonal)*. 2018; Available from: https://www.who.int/en/news-room/fact-sheets/detail/influenza-(seasonal).
2. Wong, C.M., et al., *Influenza-Associated Hospitalization in a Subtropical City.* PLoS Medicine, 2006. **3**(4): p. e121-e121.

3. Ng, S. and A. Gordon, *Influenza Burden and Transmission in the Tropics.* Current Epidemiology Reports, 2015. **2**(2): p. 89-100.
4. Chretien, J.-P., et al., *Influenza Forecasting in Human Populations: A Scoping Review.* PLoS ONE, 2014. **9**(4): p. e94130-e94130.
5. Nsoesie, E.O., et al., *A systematic review of studies on forecasting the dynamics of influenza outbreaks*. 2014, Blackwell Publishing. p. 309-316.
6. Cummings, M.J., et al., *Epidemiologic and spatiotemporal Characterization of influenza and severe acute respiratory infection in Uganda, 2010-2015.* Annals of the American Thoracic Society, 2016. **13**(12): p. 2159-2168.
7. Soebiyanto, R.P., et al., *The Role of Temperature and Humidity on Seasonal Influenza in Tropical Areas: Guatemala, El Salvador and Panama, 2008–2013.* PLoS ONE, 2014. **9**(6): p. e100659-e100659.
8. Yang, L., et al., *Influenza associated mortality in the subtropics and tropics: Results from three Asian cities.* Vaccine, 2011. **29**(48): p. 8909-8914.
9. Tamerius, J., et al., *Global influenza seasonality: Reconciling patterns across temperate and tropical regions.* Environmental Health Perspectives, 2011. **119**(4): p. 439-445.
10. Bloom-Feshbach, K., et al., *Latitudinal Variations in Seasonal Activity of Influenza and Respiratory Syncytial Virus (RSV): A Global Comparative Review.* PLoS ONE, 2013. **8**(2): p. e54445-e54445.
11. Shaman, J. and M. Kohn, *Absolute humidity modulates influenza survival, transmission, and seasonality.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(9): p. 3243-3248.
12. Shaman, J., et al., *Absolute humidity and the seasonal onset of influenza in the continental United States.* PLoS Biology, 2010. **8**(2).
13. Shek, L.P.-C. and B.-W. Lee, *Epidemiology and seasonality of respiratory tract virus infections in the tropics.* Paediatric Respiratory Reviews, 2003. **4**(2): p. 105-111.
14. Tamerius, J.D., et al., *Environmental Predictors of Seasonal Influenza Epidemics across Temperate and Tropical Climates.* PLoS Pathogens, 2013. **9**(3).
15. Lowen, A.C. and J. Steel, *Roles of Humidity and Temperature in Shaping Influenza Seasonality.* Journal of Virology, 2014. **88**(14): p. 7692-7695.

16. Yang, W., S. Elankumaran, and L.C. Marr, *Relationship between Humidity and Influenza A Viability in Droplets and Implications for Influenza's Seasonality.* PLoS ONE, 2012. **7**(10): p. 1-8.
17. Moura, F.E.A., et al., *Seasonality of Influenza in the Tropics: A Distinct Pattern in Northeastern Brazil*. 2009. p. 180-183.
18. Lutwama, J.J., et al., *Clinic-and hospital-based sentinel influenza surveillance, Uganda 2007-2010.* Journal of Infectious Diseases, 2012. **206**(SUPPL.1).
19. Mahamat, A., et al., *Climatic drivers of seasonal influenza epidemics in French Guiana, 2006-2010.* Journal of Infection, 2013. **67**(2): p. 141-147.
20. Deyle, E.R., et al., *Global environmental drivers of influenza.* Proceedings of the National Academy of Sciences of the United States of America, 2016. **113**(46): p. 13081-13086.
21. Carrat, F. and A. Flahault, *Influenza vaccine: The challenge of antigenic drift*. 2007. p. 6852-6862.
22. Liu, M., et al., *Antigenic Patterns and Evolution of the Human Influenza A (H1N1) Virus.* Scientific Reports, 2015. **5**.
23. Smith, D.J., et al., *Mapping the antigenic and genetic evolution of influenza virus.* Science, 2004. **305**(5682): p. 371-376.
24. Couch, R.B. and J.A. Kasel, *Immunity to Influenza in Man.* Annual Review of Microbiology, 1983. **37**(1): p. 529-549.
25. Greene, S.K., E.L. Ionides, and M.L. Wilson, *Patterns of Influenza-associated Mortality among US Elderly by Geographic Region and Virus Subtype, 1968–1998.* American Journal of Epidemiology, 2006. **163**(4): p. 316-326.
26. *Climatological Information Services ｜Hong Kong Observatory(HKO) ｜Climate*. [cited 2018; Available from: https://www.hko.gov.hk/en/cis/climat.htm.
27. Wallace, J. and P. Hobbs, *Atmospheric Science: An Introductory survey*. 2nd Edition ed. 2006, New York: Academic Press. 504.
28. Census and Statistics Department HKSAR, *Fertility Trend in Hong Kong, 1981 to 2017.* Hong Kong Monthly Digest of Statistics, 2018.

29. Finkenstadt, B.F. and B.T. Grenfell, *Time Series Modelling of Childhood Diseases: A Dynamical Systems Approach*. 2000. p. 187-205.
30. Roper, W.L., et al., *Serum Cross-Reactive Antibody Response to a Novel Influenza A (H1N1) Virus After Vaccination with Seasonal Influenza Vaccine*. 2009.
31. Miller, E., et al., *Incidence of 2009 pandemic influenza A H1N1 infection in England: a cross-sectional serological study.* The Lancet, 2010. **375**(9720): p. 1100-1108.
32. Yang, W., E.H.Y. Lau, and B.J. Cowling, *Dynamic interactions of influenza viruses in Hong Kong during 1998-2018.* PLoS Comput Biol, 2020. **16**(6): p. e1007989.
33. Brown, J.D., et al., *Avian influenza virus in water: Infectivity is dependent on pH, salinity and temperature.* Veterinary Microbiology, 2009. **136**(1-2): p. 20-26.
34. Chowell, G., M.A. Miller, and A.N.D.C.V. Ib, *Seasonal influenza in the United States, France, and Australia: transmission and prospects for control.*
35. Harper, G.J., *Airborne micro-organisms: survival tests with four viruses.* Epidemiology and Infection, 1961. **59**(4): p. 479-486.
36. Kramer, S.C. and J. Shaman, *Development and validation of influenza forecasting for 64 temperate and tropical countries.* PLoS Computational Biology, 2019. **15**(2): p. e1006742-e1006742.
37. Mills, C.E., J.M. Robins, and M. Lipsitch, *Transmissibility of 1918 pandemic influenza.* Nature, 2004. **432**(7019): p. 904-906.
38. Yang, W., et al., *Forecasting Influenza Epidemics in Hong Kong.* PLoS Computational Biology, 2015. **11**(7): p. 1-17.
39. Lowen, A.C., et al., *Influenza Virus Transmission Is Dependent on Relative Humidity and Temperature.* PLoS Pathogens, 2007. **3**(10): p. e151-e151.
40. Lowen, A.C., et al., *High Temperature (30°C) Blocks Aerosol but Not Contact Transmission of Influenza Virus.* Journal of virology, 2008. **82**(11): p. 5650-5652.
41. Yang, W., et al., *Dynamics of influenza in tropical Africa: Temperature, humidity, and co-circulating (sub)types.* Influenza and other Respiratory Viruses, 2018. **12**(4): p. 446-456.
42. Hong Kong's Department of Health, *2008-2009 Annual Report*.
43. Hong Kong's Department of Health, *Vaccination subsidy schemes launched*. 2009.

686 44. Hong Kong's Department of Health, *Subsidised vaccination for young children and*
687 *elderly.* 2009.
688 45. Legistrative Council of the Hong Kong Special Administrative Region of the People's
689 Republic of China. *Seasonal Influenza Vaccination*. 2018 [cited 2021; Available from:
690 https://www.legco.gov.hk/research-publications/english/essentials-1718ise06-
691 seasonal-influenza-vaccination.htm.
692 46. Centre for Health Protection. *Statistics on Vaccination Programmes in the Past 3*
693 *years*. 2020 [cited 2021; Available from:
694 https://www.chp.gov.hk/en/features/102226.html.
695

Table Captions:

Table 1. Descriptions and estimates of the parameter ranges. The range estimates are the 95% highest density intervals for parameters, obtained after two rounds of parameter selection from the initial parameter range (shown in parentheses).

696 **Table 2.** The model performance ranks for each comparison metric using training and
697 testing data. The rankings are determined by the model's absolute mean rank differences
698 with the best-ranked model. Different rankings between models indicate significantly
699 different absolute mean ranks (i.e., p-value<0.007). The mean value of the corresponding
700 metric is shown in the parentheses.

Figure Captions:

Fig 1. Influenza epidemics observed in Hong Kong during 1998-2018 and the corresponding mean daily temperature and AH. Upper panel: stacked barplot of the weekly ILI+ incidence time series. Segments of the bar represent different virus (sub)types circulating during the week. Lower panel: daily mean temperature and specific humidity (a measure of AH) observed during 1998-2018.

Fig 2. Model performance. Boxes and whiskers show the median (thick horizontal lines), interquartile range and 95% CI of RMSE (1st row), average RMSE (2nd row), correlation (3rd row) and average correlation (4th row) of the top 1000 parameter combinations for each model, during the training (red) and testing (green) period, separately.

Fig 3. Top10 model fits for three climate forcing models: AH/T (A), AH/T/Vary (B), and AH/T/Strain (C). Black crosses show observed ILI+; the colored lines run through the crosses are the top10 model estimates. The vertical dash line indicates a pandemic (2009). The shaded region represents testing years (2013-2018), while the rest are the training years.

Fig 4. Estimated relationship between influenza transmission with AH and temperature. We use the basic reproductive number (R_0) to represent the level of influenza transmission. Each point shows the estimated R_0 at different specific humidity, a measure of AH, (and temperature if included) calculated per the AH/T model (left) or the AH model (right) using the top 10 parameter combinations for the corresponding model. For the AH/T model (left panel), the color of the point shows the concurrent temperature included in the model to moderate the relationship between R_0 and specific humidity.

Fig 5. Comparison of the reproductive numbers estimated by the AH/T model with laboratory observed virus survival rate and transmission rate in guinea pigs. Left panel plots the viral survival rate (A) and transmission rate (C) against R_0 calculated using Eqn 3 and best-fit parameters for the AH/T model. Right panel plots the viral survival rate (B) and transmission rate (D) against R_e, where R_e is calculated as R_0 multiplied by the estimated mean population susceptibility during the study period. The grey vertical line indicates where R_e =1. The viral survival data came from Harper 1961 [35] and transmission rate data came from Lowen et al. 2007 and 2008 [39, 40].

Supplementary Table Captions

Table S1. The model performance ranks for each comparison metric using training and testing data (excluding the 2009 pandemic). The rankings are determined by the model's

absolute mean rank differences with the best-ranked model. Different rankings between models indicate significantly different absolute mean ranks (i.e., p-value<0.007). The mean value of the corresponding metric is shown in the parentheses.

Supplementary Figure Captions

Fig S1. The change of mean ILI+ of each influenza (sub)type in circulation within a year. The mean ILI+ took average of the ILI+ observed in Hong Kong over 21 years (1998-2018).

Fig S2. Top10 model fits for Null1 (A), Null2 (B), and AH (C) model. Black crosses show observed ILI+; the colored lines run through the crosses show the top 10 model estimates. The vertical dash line indicates the onset of the 2009 pandemic. The shaded regions indicate testing years (2013-2018); and the rest are the training years.

Fig S3. Top10 model fits for the observed seasonality (averaged over training or testing years) for the seven models: Null1 (A), Null2 (B), AH (C), AH/T (D), AH/T/Short(E) AH/T/Vary (F) and AH/T/Strain (G). Black crosses show observed averaged ILI+ over training or testing years; colored lines run through the crosses show the top10 model estimates. Left panels show model fits for the training period and the right panels show model fits for the testing period.

Tables:

Table 1. Descriptions and estimates of the parameter ranges. The estimated parameter ranges are the 95% highest density intervals, obtained after two rounds of parameter selection from the initial ranges (shown in parentheses).

Parameter and sources of initial ranges	Parameter Description	Null1	Null2	AH	AH/T	AH/T/ Strain	AH/T/ Short	AH/T/ Vary
R0 [12, 34, 36-38]	The basic reproductive number (i.e., the average number of cases caused by a primary case in a fully susceptible population).	2.06-2.11 (1.0-3.0)	NA	NA	NA	NA	NA	NA
R_{0mx} [12, 34, 36-38]	The theoretical value of R_0 at $q = q_{min}$ and $q = q_{max}$, when $T = T_c$	NA	2.52-2.69 (1.5-3.0)	1.95-2.50 (1.5-3.0)	2.34-2.93 (1.5-3.0)	2.38-2.88 (1.5-3.0)	2.13-2.67 (1.5-3.0)	2.27-2.66 (1.5-3.0)
R_{0diff}	The difference between R_{0max}, and R_0 at $q = q_{mid}$.	NA	1.10-1.20 (0.6-1.2)	0.60-1.18 (0.6-1.2)	0.86-1.18 (0.6-1.2)	0.70-1.17 (0.6-1.2)	0.65-1.09 (0.6-1.2)	0.72-1.12 (0.6-1.2)
q_{min} (g/kg)	The absolute humidity value at which $R_0 = R_{0max}$ when $T = T_c$; the minimum value of absolute humidity permitted.	NA	NA	7.7-8.0 (2.0-8.0)	2.2-4.0 (2.0-8.0)	2.0-3.6 (2.0-8.0)	2.1-3.4 (2.0-8.0)	2.1-3.7 (2.0-8.0)
q_{max} (g/kg)	The absolute humidity value at which $R_0 = R_{0max}$ when $T = T_c$; the maximum value of absolute humidity permitted.	NA	NA	22.6-23.0 (16.0-23.0)	17.0-20.0 (16.0-23.0)	17.0-19.0 (16.0-23.0)	18.0-19.0 (16.0-23.0)	17.0-19.0 (16.0-23.0)
q_{mid} (g/kg)	The absolute humidity value at which $R_0 = R_{0max} - R_{0diff}$.	NA	NA	12.6-13.0 (10.0-13.0)	10.2-11.3 (10.0-13.0)	10.2-12.5 (10.0-13.0)	10-12.2 (10.0-13.0)	10.2-12.6 (10.0-13.0)

T_c (°C) [33, 35]	The cutoff temperature above which temperature negatively impacts R_0.	NA	NA	NA	20.24-24.0 (20-25)	20.04-24.46 (20-25)	20.16-24.47 (20-25)	20.52-23.87 (20-25)
T_{diff} (°C)	A parameter that, when subtracted from T_c, yields the minimum temperature permitted as a model input.	NA	NA	NA	0.40-5.09 (0-15)	1.28-7.05 (0-15)	1.34-5.31 (0-15)	0.77-4.46 (0-15)
T_{exp}	An exponent determining the strength of the impact of temperature on R_0.	NA	NA	NA	0.95-1.54 (0.5-2.0)	0.67-1.49 (0.5-2.0)	0.78-1.33 (0.5-2.0)	0.78-1.34 (0.5-2.0)
D (days) [12, 30, 31, 38]	The duration of influenza infection.	3.35-3.71 (2-5)	4.68-4.98 (2-5)	2.56-4.99 (2-5)	4.05-4.99 (2-5)	3.50-4.70 (2-5)	3.87-4.98 (2-5)	3.67-4.90 (2-5)
L (days) [12, 22, 23, 36, 38]	The duration of influenza immunity; in the AH/T/Vary model, the duration of immunity among those who do not lose immunity at an accelerated rate.	369-391 (365-3650)	421-607 (365-3650)	1298-2860 (365-3650)	376-489 (365-3650)	536-838 (365-3650)	310-451 (183-548)	374-567 (365-3650)
L_S (days)	The duration of influenza immunity among those with accelerated immunity loss in the AH/T/Vary model.	NA	NA	NA	NA	NA	NA	82- 336 (30-365)
ρ	The proportion of the population with accelerated immunity loss in the AH/T/Vary model.	NA	NA	NA	NA	NA	NA	0.0006-0.29 (0-0.5)
S_0	The number of people susceptible to influenza at the beginning of the model run.	48.78% - 76.16%	78.77%-79.98% (40%-80%)	71.58% - 79.71%	59.27% - 76.79%	57.03% - 69.86%	67.03% - 77.61%	66.02%-79.36 (40%-80%)

		(40%-80%)		(40%-80%)	(40%-80%)	(40%-80%)	(40%-80%)	
I_0	The number of people infected at the beginning of the model run.	520-1489 (500-1500)	534-851 (500-1500)	611-1161 (500-1500)	515-1052 (500-1500)	512-1105 (500-1500)	515-900 (500-1500)	519-1256 (500-1500)
p	An exponent to allow for imperfect mixing. Homogenous mixing occurs when p = 1.0.	0.97	0.97	0.97	0.97	0.97	0.97	0.97

740

741 **Table 2.** The model performance ranks for each comparison metric using training and
742 testing data. The rankings are determined by the model's absolute mean rank differences
743 with the best-ranked model. Different rankings between models indicate significantly
744 different absolute mean ranks (i.e., p-value<0.007). The mean value of the corresponding
745 metric is shown in the parenthesis.

	Models	Null1	Null2	AH	AH/T	AH/T/ Strain	AH/T/ Short	AH/T/ Vary
Train	full.RMSE	6 (772)	4 (675)	7 (804)	1 (611)	6 (720)	1 (604)	1 (611)
	avg.RMSE	6 (422)	5 (349)	7 (541)	1 (243)	4 (310)	1 (240)	1 (249)
	full.Correlation	7 (0.12)	5 (0.43)	6 (0.32)	2 (0.53)	4 (0.50)	1 (0.55)	2 (0.52)
	avg.Correlation	7 (-0.16)	5 (0.73)	6 (0.62)	3 (0.87)	1 (0.88)	1 (0.89)	3 (0.87)
	Average rank	6.5	4.75	6.5	1.75	3.75	1	1.75
Test	full.RMSE	6 (620)	3 (478)	5 (613)	2 (479)	7 (676)	3 (492)	1 (470)
	avg.RMSE	6 (434)	2 (237)	4 (422)	2 (249)	6 (476)	4 (274)	1 (228)
	full.Correlation	7 (-0.0005)	5 (0.49)	6 (0.10)	3 (0.54)	1 (0.55)	1 (0.54)	4 (0.52)
	avg.Correlation	7 (-0.0007)	5 (0.76)	6 (0.11)	1 (0.78)	1 (0.78)	1 (0.79)	4 (0.78)
	Average rank	6.5	3.75	5.25	2	3.75	2.25	2.5

746

Figures:

Fig 1. Influenza epidemics observed in Hong Kong during 1998-2018 and the corresponding mean daily temperature and AH. Upper panel: stacked barplot of the weekly ILI+ incidence time series. Segments of the bar represent different virus (sub)types circulating during the week. Lower panel: daily mean temperature and specific humidity (a measure of AH) observed during 1998-2018.

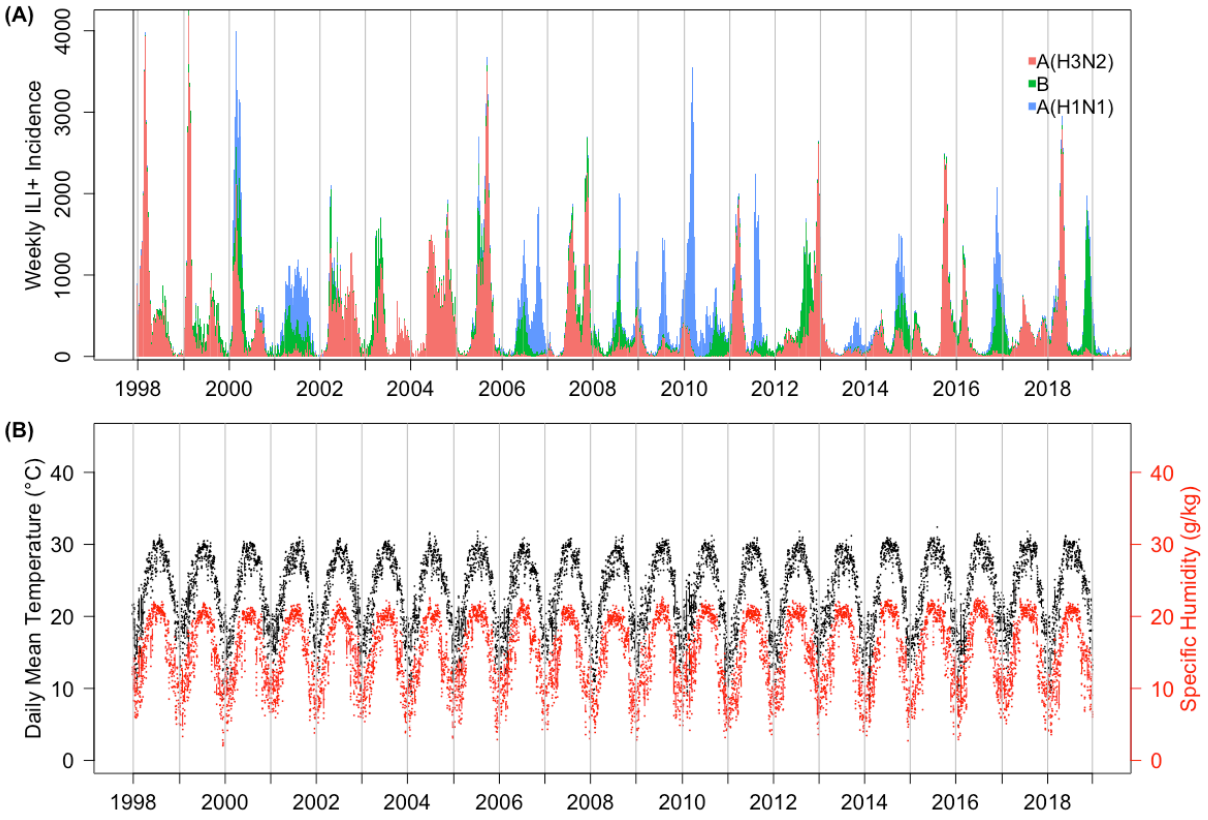


Fig 2. Model performance. Boxes and whiskers show the median (thick horizontal lines), interquartile range and 95% CI of RMSE (1st row), average RMSE (2nd row), correlation (3rd row) and average correlation (4th row) of the top 1000 parameter combinations for each model, during the training (red) and testing (green) period, separately.

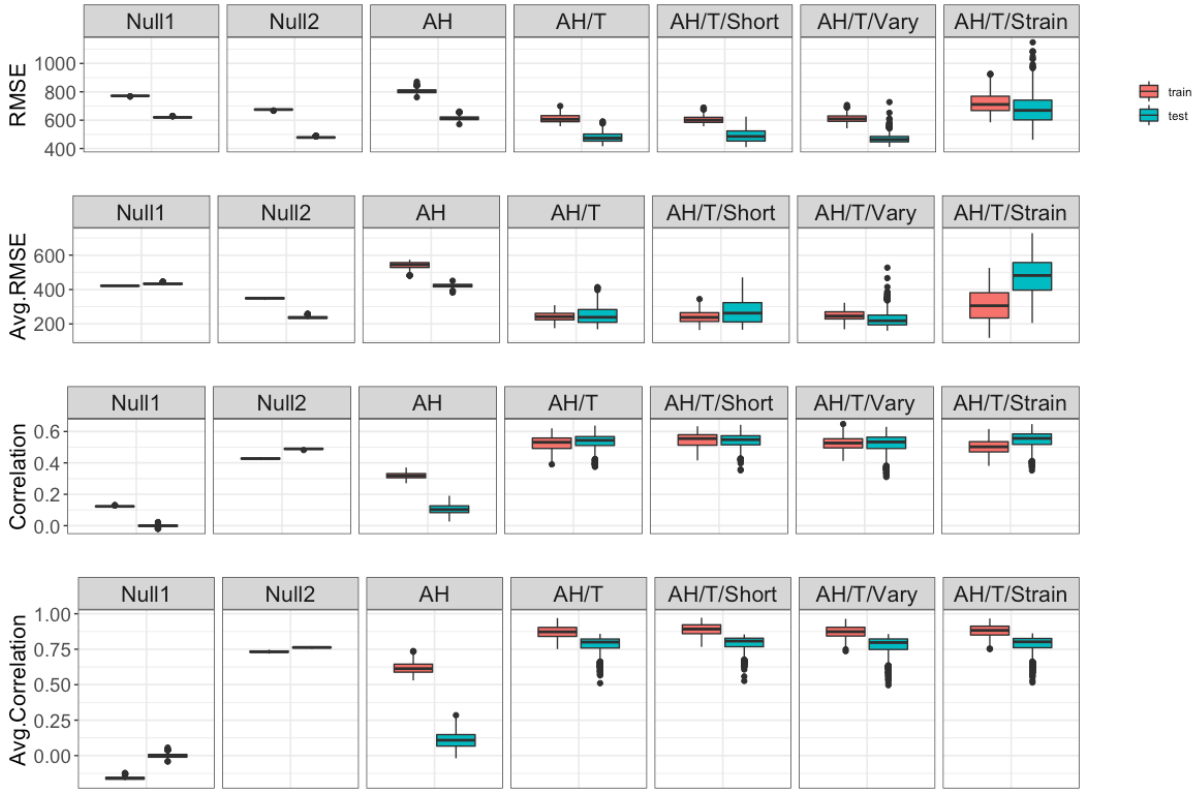


Fig 3. Top10 model fits for three climate forcing models: AH/T (A), AH/T/Vary (B), and AH/T/Strain (C). Black crosses show observed ILI+; the colored lines run through the crosses are the top10 model estimates. The vertical dash line indicates a pandemic (2009). The shaded region represents testing years (2013-2018), while the rest are the training years.

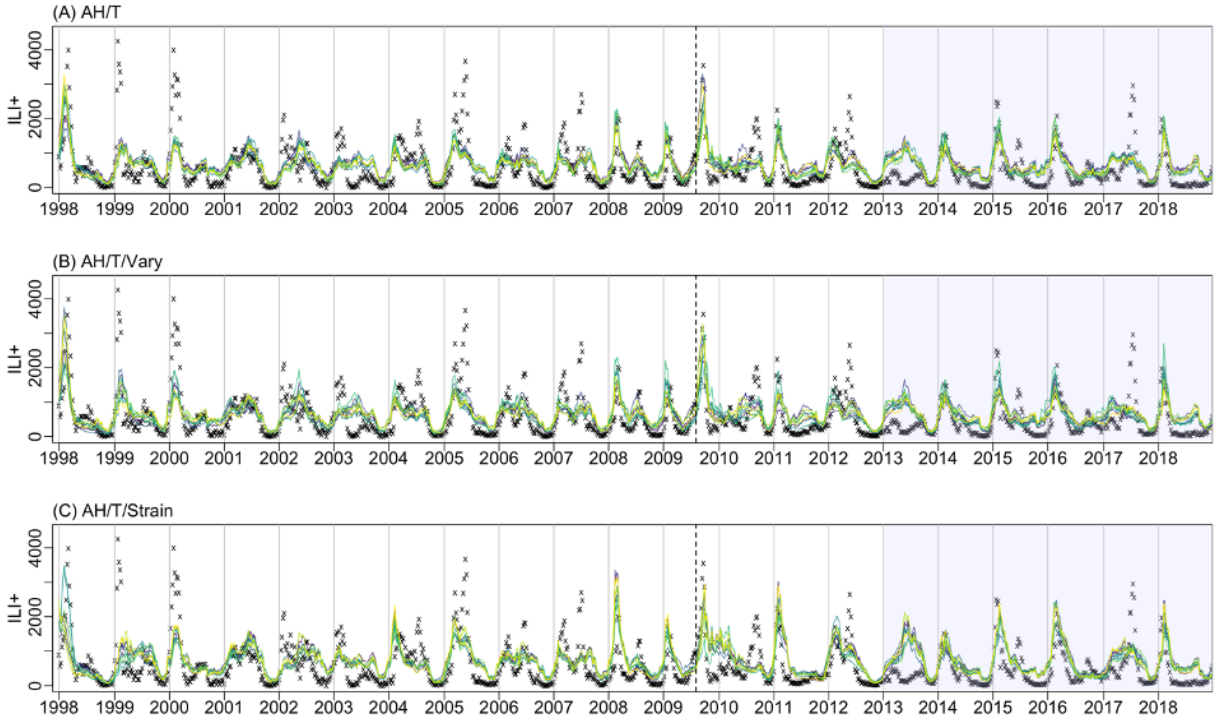


Fig 4. Estimated relationship between influenza transmission with AH and temperature. We use the basic reproductive number (R_0) to represent the level of influenza transmission. Each point shows the estimated R_0 at different specific humidity, a measure of AH, (and temperature if included) calculated per the AH/T model (left) or the AH model (right) using the top 10 parameter combinations for the corresponding model. For the AH/T model (left panel), the color of the point shows the concurrent temperature included in the model to moderate the relationship between R_0 and specific humidity.

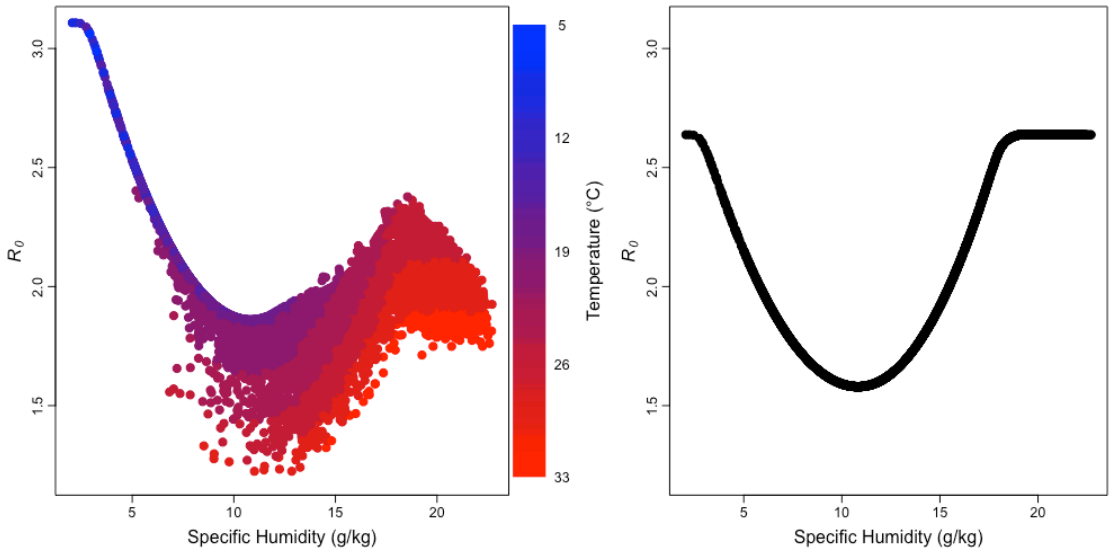
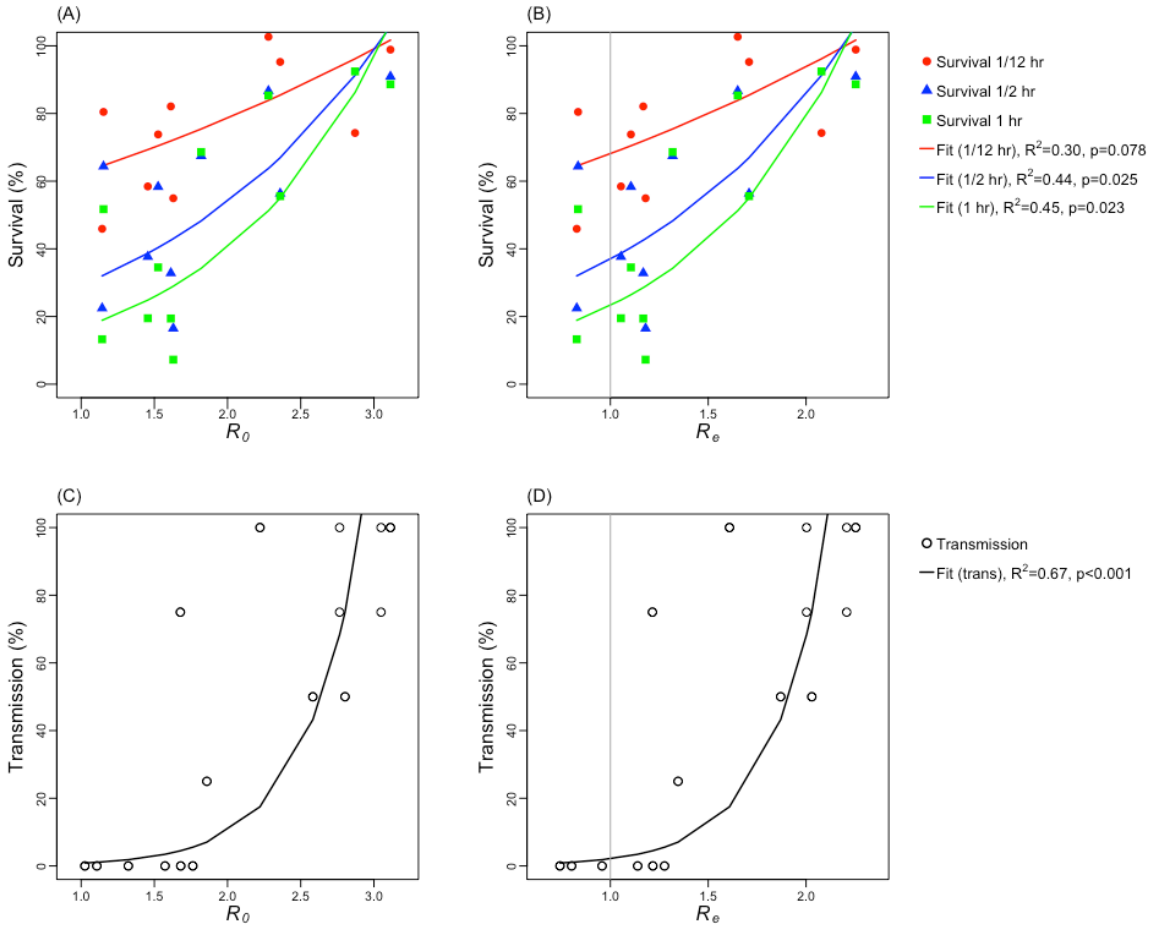


Fig 5 Comparison of the reproductive numbers estimated by the AH/T model with laboratory observed virus survival rate and transmission rate in guinea pigs. Left panel plots the viral survival rate (A) and transmission rate (C) against R_0 calculated using Eqn 3 and best-fit parameters for the AH/T model. Right panel plots the viral survival rate (B) and transmission rate (D) against R_e, where R_e is calculated as R_0 multiplied by the estimated mean population susceptibility during the study period. The grey vertical line indicates where R_e =1. The viral survival data came from Harper 1961 [35] and transmission rate data came from Lowen et al. 2007 and 2008 [39, 40].



Supporting Information

Modeling influenza seasonality in tropics and subtropics

Haokun Yuan, Sarah Kramer, Eric H. Y. Lau, Benjamin J. Cowling, Wan Yang

This supplemental document includes 1) Preliminary data processing; and 2) Simulation methods and modeling details.

1. Preliminary data processing

1.1 Calculation of absolute humidity [1]

First, we calculated saturation vapor pressure as:

$$e_s(T) = e_s(T_0) \times e^{\frac{L}{R_v}\left(\frac{1}{T_0}-\frac{1}{T}\right)} \quad [S4]$$

where $e_s(T)$ is the saturation vapor pressure at temperature T (in K), $e_s(T_0)$ is the saturation vapor pressure at 273.15 K, L is the latent heat of evaporation for water, R_v is the gas constant for water vapor. We then computed vapor pressure at each time point as:

$$e = e_s(T)\left(\frac{RH}{100}\right) \quad [S5]$$

where RH represents relative humidity. We calculated the mixing ratio, mr, as:

$$mr = \frac{R_d}{R_v}\left(\frac{e}{p_0 - e}\right) \quad [S6]$$

where R_d is the gas constant for dry air and p_0 is the atmospheric pressure at sea level. Finally, absolute humidity was calculated as:

$$AH = \frac{mr}{1 + mr} \quad [S7]$$

2. Simulation methods and modeling details

2.1 Stochastic Model Runs

Similar to the climate forcing model used to model influenza in temperate region by Shaman et al. [2], we constructed a stochastic Markov chain, where stochastic is introduced by the transition between states. The number of individuals moving from one state to another (i.e. susceptible to infected, infected to recovered) is random draw from a Poisson distribution with a rate determined by Eqn 1.

2.2 Climate forcing model for temperate regions (Null2)

The climate forcing model for temperate regions, first introduced by Shaman et al [2], assumes influenza transmission decreases monotonically with increasing absolute humidity. This relationship is modeled as:

$$R_0(t) = exp(a \times q(t) + b) + R_{0min} \quad [S8]$$

Where $a = -180$, $b = \log(R_{0max} - R_{0min})$, R_{0max} is the maximum daily basic reproductive number, while R_{0min} is the minimum daily basic reproductive number.

2.3 Climate forcing model for (sub)tropical regions (AH/T, AH/T/Short, AH/T/Vary, AH/T/Strain)

Absolute humidity alone was proposed to have a bimodal effect on influenza transmission. And the formula can be written as follow:

$$R_0(t) = aq^2(t) + bq(t) + c \quad [S6]$$

Values of a, b, and c are defined as:

$$\begin{cases} a = \dfrac{-b}{q_{max} + q_{min}} \\ b = \dfrac{\left(R_{0_{max}} - \left(R_{0_{max}} - R_{0_{diff}}\right)\right)(q_{max} + q_{min})}{(q_{max} - q_{mid})(q_{min} - q_{mid})} \\ c = \left(R_{0_{max}} - R_{0_{diff}}\right) - aq_{mid}^2 - bq_{mid} \end{cases} \quad [S7]$$

Given equation S6, the derivation of equation S7 is shown as below:

$$R_{0_{max}} = aq_{min}^2 + bq_{min} + c \quad (1)$$
$$R_{0_{max}} = aq_{max}^2 + bq_{max} + c \quad (2)$$
$$R_{0_{min}} = aq_{mid}^2 + bq_{mid} + c \quad (3)$$

Subtracting (1) from (2), we can get:

$$a(q_{max}^2 - q_{min}^2) + b(q_{\max} - q_{\min}) = 0$$

Rearrange the equation:

$$a = -\frac{b}{q_{max} + q_{min}}$$

Moreover, by subtracting (3) from (2), we can get:

$$\mathrm{R}_{0_{\max}} - \mathrm{R}_{0_{\min}} = a(q_{max}^2 - q_{mid}^2) + b(q_{max} - q_{mid})$$

Plug in a:

$$\mathrm{R}_{0_{\max}} - \mathrm{R}_{0_{\min}} = b\left[(q_{max} - q_{mid}) - \frac{q_{max}^2 - q_{mid}^2}{q_{max} + q_{min}}\right]$$

Move b to the left-hand side and rearrange:

$$b = \frac{\left(\mathrm{R}_{0_{\max}} - \left(\mathrm{R}_{0_{\max}} - \mathrm{R}_{0_{\mathrm{diff}}}\right)\right)(q_{max} + q_{min})}{(q_{max} - q_{mid})(q_{min} - q_{mid})}$$

Plug a and b into Eqn. S7:

$$c = \left(R_{0_{max}} - R_{0_{diff}}\right) - aq_{mid}^2 - bq_{mid}$$

Reference:

1. Wallace J, Hobbs P. Atmospheric Science: An Introductory Survey. New York: Academic Press; 2006.
2. Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M. Absolute Humidity and the Seasonal Onset of Influenza in the Continental United States. Ferguson NM, editor. PLoS Biol. 2010 Feb 23;8(2):e1000316.

Supplementary figures

Fig S1. The change of mean ILI+ of each influenza (sub)type in circulation within a year. The mean ILI+ took average of the ILI+ observed in Hong Kong over 21 years (1998-2018).

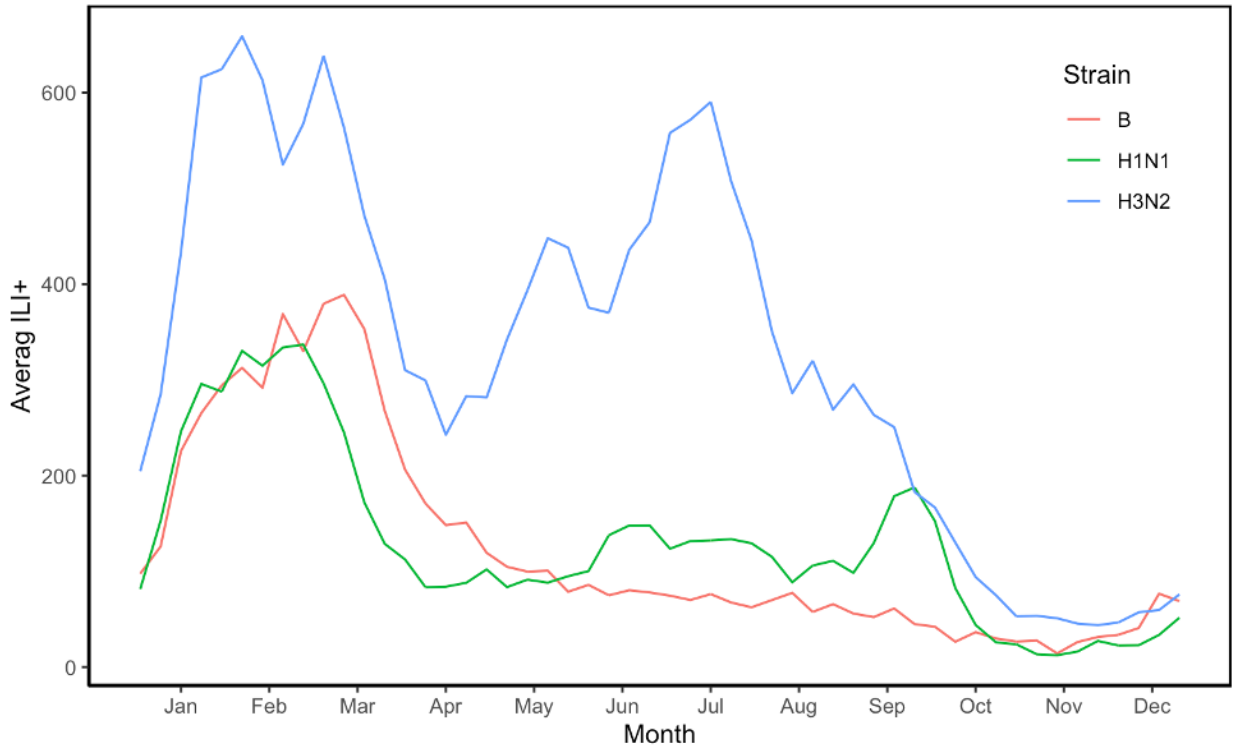


Fig S2. Top10 model fits for Null1 (A), Null2 (B), and AH (C) model. Black crosses show observed ILI+; the colored lines run through the crosses show the top 10 model estimates. The vertical dash line indicates the onset of the 2009 pandemic. The shaded regions indicate testing years (2013-2018); and the rest are the training years.

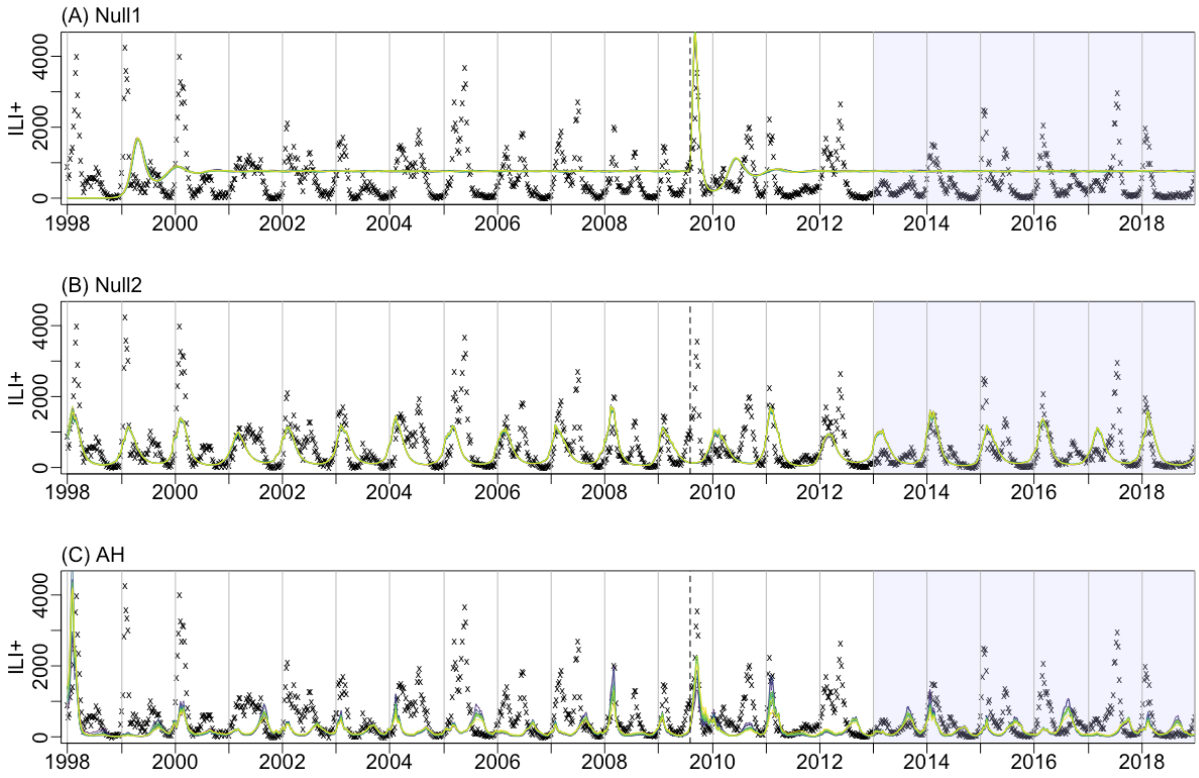
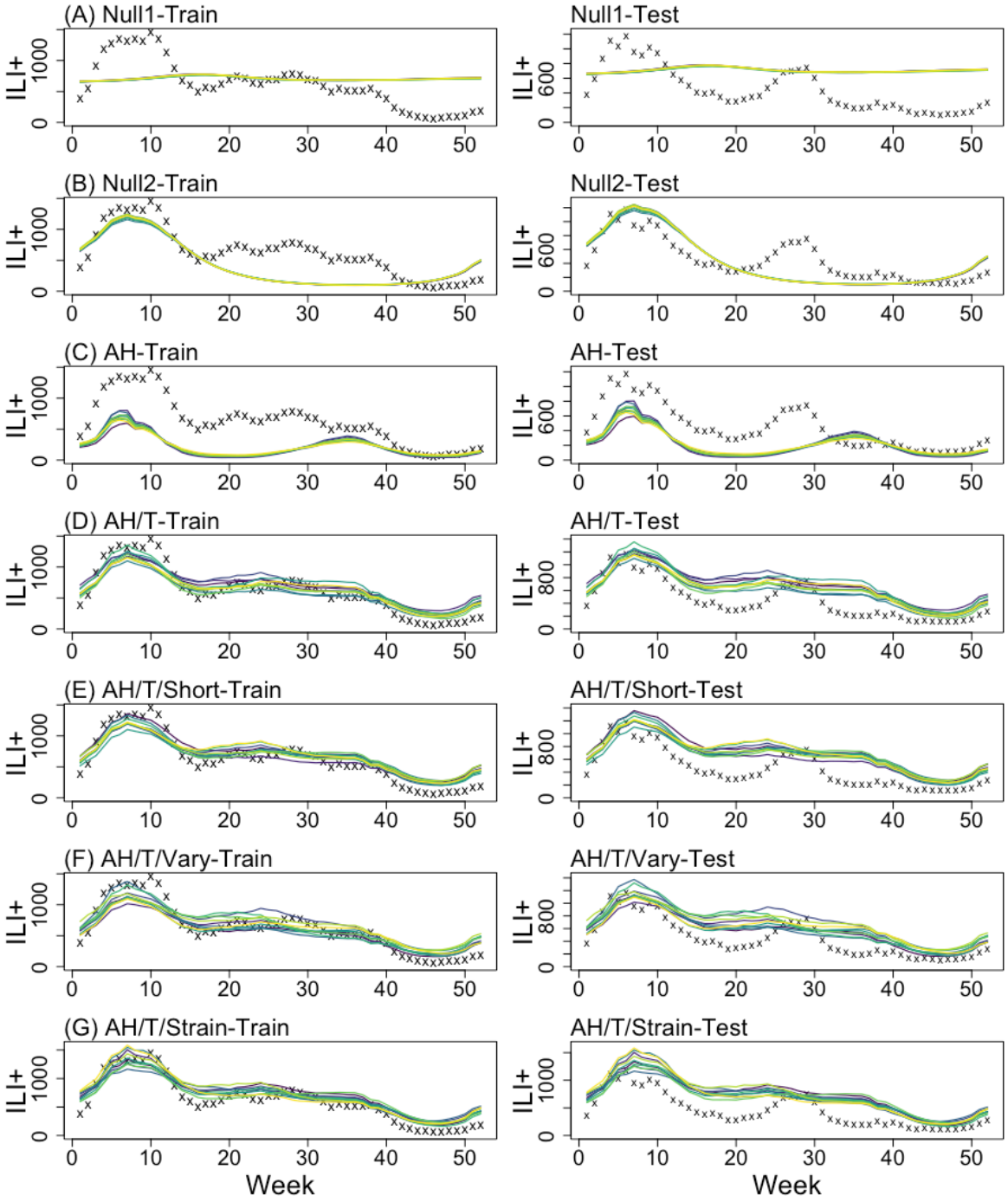


Fig S3. Top10 model fits for the observed seasonality (averaged over training or testing years) for the seven models: Null1 (A), Null2 (B), AH (C), AH/T (D), AH/T/Short(E) AH/T/Vary (F) and AH/T/Strain (G). Black crosses show observed averaged ILI+ over training or testing years; the colored lines run through the crosses are the top10 model estimates. Left panels are the model fit towards training data, and the right panels plot the fit for testing.



Tables:

Table S1. The model performance ranks for each comparison metric using training and testing data (excluding the 2009 pandemic). The rankings are determined by the model's absolute mean rank differences with the best-ranked model. Different rankings between models indicate significantly different absolute mean ranks (i.e., p-value<0.007). The mean value of the corresponding metric is shown in the parenthesis.

	Models	Null1	Null2	AH	AH/T	AH/T/ Strain	AH/T/ Short	AH/T/ Vary
Train	full.RMSE	6 (781)	4 (644)	7 (821)	1 (606)	5 (701)	1 (602)	1 (612)
	avg.RMSE	6 (452)	5 (346)	7 (572)	1 (264)	4 (318)	1 (266)	1 (271)
	full.Correlation	7 (-0.08)	5 (0.51)	6 (0.27)	2 (0.55)	4 (0.52)	1 (0.56)	2 (0.55)
	avg.Correlation	7 (0.13)	5 (0.78)	6 (0.66)	1 (0.89)	1 (0.90)	1 (0.90)	1 (0.90)
	Average rank	6.5	4.75	6.5	1.25	3.5	1	1.25
Test	full.RMSE	6 (621)	2 (478)	5 (609)	2 (484)	7 (705)	2 (493)	1 (474)
	avg.RMSE	5 (434)	1 (238)	5 (421)	3 (256)	5 (499)	4 (274)	1 (235)
	full.Correlation	7 (-0.0007)	5 (0.49)	6 (0.11)	1 (0.54)	1 (0.54)	1 (0.54)	4 (0.53)
	avg.Correlation	7 (-0.0007)	5 (0.76)	6 (0.11)	1 (0.79)	1 (0.78)	1 (0.80)	4 (0.78)
	Average rank	6.25	3.25	5.5	1.75	3.5	2	2.5