Forecasting methods

We applied the forecasting method presented in our retrospective forecasting studies [*1*, *2*], which combines an $SEIR$ compartment model of infection with influenza notification counts through the use of a bootstrap particle filter.

In brief, we randomly selected parameter values for the $SEIR$ model (Eqs 1–6, parameters in Table S1) to generate a suite of candidate epidemics, and defined a relationship between disease incidence in the $SEIR$ model and the influenza notification counts (an "observation model"). This allowed us to calculate the likelihood of "observing" the notification counts from each of the model simulations. The particle filter (Eqs 7–11) was used to update these likelihoods ("weights" w_i) as data ($\mathbf{y_t}$) were reported, and to discard extremely unlikely simulations in favour of more likely ones ("resampling").

As in our previous studies, we assumed the relationship between disease incidence in the $SEIR$ model and the weekly notification counts was defined by a negative binomial distribution with dispersion parameter k (Eqs 12–15), since the data are non-negative integer counts and are over-dispersed when compared to a Poisson distribution. The probability of being *observed* (i.e., of being reported as a notifiable case) was the product of two probabilities: that of becoming infectious (p_{inf}), and that of being identified (p_{id}, the likelihood of being symptomatic, presenting to a doctor, and having a specimen collected). The probability of becoming infectious was defined as the fraction of the *model* population that became infectious (i.e., transitioned from E to I), and subsumed symptomatic and asymptomatic infections.

Values for p_{id} and k were informed by retrospective forecasts using notifications data from previous seasons [*2*], while the background notification rate p_{bg} was estimated from out-of-season notification levels (March to May) for 2016 and for 2017. The climatic modulation signals $F(t)$ were characterised by smoothed absolute humidity data for each city in previous years, as previously described [*3*].

Forecasts were generated using the pypfilt [1] and epifx [2] packages for Python, which were developed as part of this project and are both available under permissive free software licenses.

[1] http://pypfilt.readthedocs.io/en/latest/
[2] http://epifx.readthedocs.io/en/latest/

Equations

$$\frac{dS}{dt} = -\beta SI - \theta_{seed} \tag{1}$$

$$\frac{dE}{dt} = \beta SI + \theta_{seed} - \sigma E \tag{2}$$

$$\frac{dI}{dt} = \sigma E - \gamma I \tag{3}$$

$$\frac{dR}{dt} = \gamma I \tag{4}$$

$$\beta = R_0 \cdot \gamma \cdot [1 + \alpha \cdot F(t)]) \tag{5}$$

$$\theta_{\text{seed}} = \begin{cases} \frac{1}{N} & \text{if } S(t) = 1 \text{ and } \theta(t) < p_{\text{seed}} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$$\mathbf{x_t} = [S(t), E(t), I(t), R(t), R_0, \alpha, \sigma, \gamma]^T \tag{7}$$

$$w_i(0) = (N_{\text{px}})^{-1} \tag{8}$$

$$w_i'(t \mid \mathbf{y_t}) = w_i(t-1) \cdot P(\mathbf{y_t} \mid \mathbf{x_t^i}; k) \tag{9}$$

$$w_i(t \mid \mathbf{y_t}) = w_i'(t) \cdot \left(\sum_{j=1}^{N_{\text{px}}} w_j'(t) \right)^{-1} \tag{10}$$

$$N_{\text{eff}}(t) = \left(\sum_{j=1}^{N_{\text{px}}} \left[w_j'(t) \right]^2 \right)^{-1} \tag{11}$$

$$p_{\text{inf}}(t, \Delta) = S(t-\Delta) + E(t-\Delta) - S(t) - E(t) \tag{12}$$

$$p_{\text{ili}}(t, \Delta) = p_{\text{inf}}(t, \Delta) \cdot p_{\text{id}} + [1 - p_{\text{inf}}(t, \Delta)] \cdot \Delta \cdot p_{\text{bg}} \tag{13}$$

$$P(\mathbf{y_t} \mid \mathbf{x_t}; k) = \frac{\Gamma(\mathbf{y_t} + k)}{\Gamma(k) \cdot \mathbf{y_t}!} \cdot (p_{\text{k}})^k \cdot (1 - p_{\text{k}})^{\mathbf{y_t}} \tag{14}$$

$$p_{\text{k}} = \frac{k}{k + N \cdot p_{\text{ili}}} \tag{15}$$

Tables

	Meaning	Value
β	Force of infection	Eq 5
R_0	Basic reproduction number	$\sim \mathcal{U}(1,2)$
σ	Incubation period (days $^{-1}$)	$\sim [\mathcal{U}(0.5,3)]^{-1}$
γ	Infectious period (days $^{-1}$)	$\sim [\mathcal{U}(0.5,3)]^{-1}$
F(t)	Climatic modulation signal	—
α	Scale of climatic modulation	$\sim \mathcal{U}(-0.2,0.2)$
p_{seed}	Daily probability of initial exposure	$\frac{1}{36}$
$\theta(t)$	Stochastic variable for seeding an initial exposure	$\sim \mathcal{U}(0,1)$
N_{px}	Number of particles (simulations)	$15,000$
N_{min}	Minimum number of effective particles	$0.25 \cdot N_{\text{px}}$
Δ	Observation period (days)	7
k	Dispersion parameter	100
p_{bg}	Background observation rate	see Table S2
p_{id}	Observation probability	see Table S2
N	Population size	see Table S2

Table S1: Parameter values for (i) the transmission model; (ii) the bootstrap particle filter; and (iii) the observation model. Here, $\mathcal{U}(x,y)$ denotes the continuous uniform distribution where x and y are the minimum and maximum values, respectively.

Location	Year	p_{bg}	p_{id}	p^*_{id}	N
NSW: Sydney	2016	200	0.0065	—	4,921,000
NSW: Sydney	2017	125	0.0076	0.0228	4,921,000
Qld: Brisbane	2016	62	0.004	—	2,308,700
Qld: Brisbane	2017	46	0.003	0.012	2,308,700
Qld: Gold Coast	2016	23	0.0045	—	555,608
Qld: Gold Coast	2017	20	0.0045	0.018	555,608
Qld: Toowoomba	2016	5	0.00625	—	163,232
Qld: Toowoomba	2017	5	0.00625	0.025	163,232
Vic: Melbourne	2016	80	0.00275	—	4,108,541
Vic: Melbourne	2017	70	0.00275	0.011	4,108,541

Table S2: Location-specific forecast parameter values. The p^*_{id} column lists the recalibrated observation probabilities for 2017.

References

1. Moss R, Zarebski A, Dawson P, McCaw JM. Forecasting influenza outbreak dynamics in Melbourne from Internet search query surveillance data. Influenza and Other Respiratory Viruses. 2016 Jul;10(4):314–323.

2. Moss R, Zarebski A, Dawson P, McCaw JM. Retrospective forecasting of the 2010–14 Melbourne influenza seasons using multiple surveillance systems. Epidemiology and Infection. 2017 Jan;145(1):156–169. Available from: http://dx.doi.org/10.1017/S0950268816002053.

3. Zarebski AE, Dawson P, McCaw JM, Moss R. Model selection for seasonal influenza forecasting. Infectious Disease Modelling. 2017 Apr;2(1):56–70.