Characterization and forecast of global influenza (sub)type dynamics

Francesco Bonacina[1,2], Pierre-Yves Boëlle[1], Vittoria Colizza[1,3], Olivier Lopez[2,4], Maud Thomas[2], Chiara Poletto[5*]

[1] INSERM, Sorbonne Université, Pierre Louis Institute of Epidemiology and Public Health, Paris, France

[2] Sorbonne Université, CNRS, Laboratoire de Probabilités, Statistique et Modélisation, LPSM, Paris, France

[3] Department of Biology, Georgetown University, Washington, DC, USA

[4] CREST Laboratory, CNRS, Groupe des Écoles Nationales d'Économie et Statistique, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France

[5] Department of Molecular Medicine, University of Padova, Padova, Italy

*Corresponding author. Email : chiara.poletto@unipd.it

Abstract

The (sub)type composition of seasonal influenza waves varies in space and time. (Sub)types tend to have different impacts on population groups, therefore understanding the drivers of their co-circulation and anticipating their composition is important for epidemic preparedness and response. FluNet provides data on influenza specimens by (sub)type for more than one hundred fifty countries. However, due to surveillance variations across countries, global analyses usually focus on (sub)type compositions, a kind of data which is difficult to treat with advanced statistical methods. We used Compositional Data Analysis to circumvent the problem and study trajectories of annual (sub)type compositions of countries. First, we examined global trends from 2000 to 2022. We identified a few seasons which stood out for the strong within-country (sub)type dominance due to either a new virus/clade taking over (2003/2004 season, A/H1N1pdm pandemic) or (sub)types' spatial segregation (COVID-19 pandemic). Second, we showed that the composition trajectories of countries between 2010 and 2019 clustered in two macroregions characterized by (sub)type alternation vs. persistent mixing. Finally, we defined five algorithms for forecasting the next-year composition and we found that taking into account the global history of (sub)type composition in a Bayesian Hierarchical Vector AutoRegressive model improved predictions compared with naive methods. The joint analysis of spatiotemporal dynamics of influenza (sub)types worldwide revealed a hidden structure in (sub)type circulation that can be used to improve predictions of the (sub)type composition of next year's epidemic according to place.

Introduction

Since 2009, the H1N1pdm and H3N2 subtypes of influenza type A and influenza type B have co-circulated in the human population, with highly variable occurrences in space and time (1, 2). The viral diversity profoundly impacts the epidemiological characteristics of epidemics: epidemics dominated by A/H3N2 strains are generally more severe as this strain is more transmissible (3) and causes increased morbidity in the elderly due to the immune the imprinting mechanism (4–6); while epidemics due to the A/H1N1 and B strains tend to cause more infections in the young (6–8). The diversity of strains also manifests with extended epidemic periods, as, for example, the peak of A infections and of B infections are typically separated by a few weeks in the northern temperate regions (9, 10). Anticipating the (sub)type composition of the coming season is therefore key to improving preparedness, e.g. planning hospital capacities, awareness campaign and vaccine distribution (10, 11).

Yet anticipating the (sub)type co-circulation is complicated because influenza viruses interact with one another. Evidence of viral interference has been found experimentally (3, 12), from mathematical models fitted to country-level incidence data (13, 14), and from country (10, 15) and multi-country (9, 16, 17) statistical analyses. In particular, past studies have shown that cross-immunity is an important ingredient of models aiming at reproducing plausible influenza dynamics (13, 14). In other words, Influenza (sub)types form a coupled ecological system that needs to be studied as a whole. A second source of complication is represented by the fact that influenza rapidly spreads globally (18, 19) and viral compositions in different countries are interdependent. This makes the study of worldwide influenza circulation critical for interpreting the viral patterns observed within a focal country.

In response to these needs, the Global Influenza Surveillance and Response System (GISRS) (1) gathers and makes available through the FluNet portal the weekly number of samples by (sub)types and country. The quality and quantity of the data is constantly increasing, but (sub)typing effort is not standardized across countries yet. To allow multi-countries comparison, it is therefore preferable to focus on rescaled quantities, such as the percentages of infections by each (sub)type. These data allowed to show strong patterns of alternation between A/H1N1pdm and A/H3N2 in Europe (20, 21), the domination of A/H3N2 among all (sub)types and inter-hemispheric synchrony in its circulation (9). It was also used to characterize the altered (sub)type circulation following the emergence of the A/H1N1pdm (sub)type (16) and the COVID-19 pandemic (22). However, these studies did not provide anticipation. Research work attempting to forecast the next-year (sub)type composition in a country is so far scarce (23).

Here, we undertook a systematic analysis of the coupled dynamics of (sub)types by analyzing their relative abundances across countries and years through the Compositional Data Analysis (CoDA) framework (24, 25). This approach has been used in ecology and geology to analyze percentage data, taking into account the sum-to-one constraint. Here, we defined for each country a trajectory in the (sub)type composition space. We quantified how these trajectories evolved in time and space. We then proposed an approach able to leverage this structure to forecast (sub)type relative abundance in each country one year ahead with improved accuracy compared with naive estimators.

Results

Studying the relative abundances of influenza (sub)types with Compositional Data Analysis (CoDA)

We studied the relative abundance of influenza (sub)types for different countries-years, defined by the percentages of the form (B%, H1%, H3%). For brevity, we will use in figures and equations H1 for the A/H1N1 strains - historical A/H1N1 before 2009, and the A/H1N1pdm after -, and H3 for A/H3N2. We consider weekly surveillance data reported in FluNet (1, 26) from 2000 to 2023, for up to 151 countries. We aggregated data annually - from May to April - and defined “compositions” (24), i.e. the percentages of each of the 3 (sub)types for each year and country (Fig. 1A) (for details see the Material and Methods). These compositions are “sum to one” data that can be shown in a ternary plot (Fig. 1B). The study of this type of data is complicated since the components are inherently non-independent (24, 27), making standard statistical tools inadequate (25). Modern compositional analysis solves the issue by transforming the original data to log-ratios (28). Here, we chose to use the isometric log-ratio transformation (*ilr*), i.e. the log-ratio of B to A and of A/H1N1 subtypes to A/H3N2, as it provided simple epidemiological interpretation - an alternative transformation, the additive log-ratio transformation was tested in the sensitivity analysis, see the SI Appendix. The transformation is as follows :

$$\begin{cases} u &= \sqrt{\frac{2}{3}} \ln \frac{B\%}{\sqrt{H1\% * H3\%}} \\ v &= \sqrt{\frac{1}{2}} \ln \frac{H1\%}{H3\%} \end{cases}$$

The ternary graph can then be remapped to these new coordinates (Fig. 1C), that makes it easy to represent trajectories of (sub)type abundance over time for a given country (Fig. 1D). In these graphs, we also highlighted “dominance” boundaries corresponding to compositions

with one (sub)type amounting to more than 50% of the samples. The central region, or "co-dominance" region, corresponds to neither (sub)type being dominant.

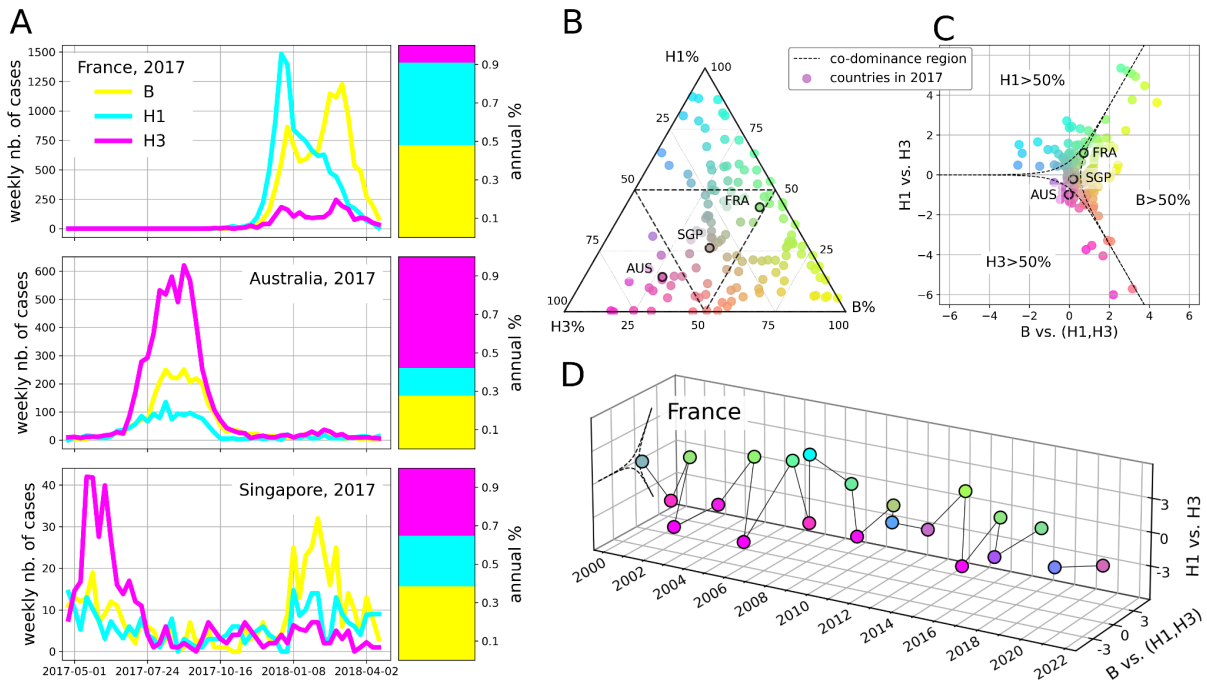

Fig. 1 - Use of CoDA to define (sub)type abundance trajectories. A) FluNet samples by (sub)types for France, Australia, and Singapore. We consider the year 2017 as an example, defined as the period going from Apr 2017 to Apr 2018, and compute the (sub)type abundances over the year. B) Proportions of A/H1N1 strains (A/H1N1 before 2009, and A/H1N1pdm after 2009), A/H3N2, and B represented in a ternary graph. Colored points represent the observations for each of the 132 countries considered in 2017. The dotted line separates the three dominance and the (central) co-dominance regions. Points are color-coded according to the composition in B (yellow), A/H1N1 (cyan) and A/H3N2 (magenta) - gray indicates perfect balance of circulating strains. C) Relative abundances of (sub)types of the same 132 points plotted in panel B) after isometric log-ratio transformation. D) Trajectory of relative abundances log-ratios in France, from 2000 to 2022. The point for 2020 is missing because less than 50 cases were uploaded on FluNet between Apr 2020 and Apr 2021. Points are colored with the same triangular code as in Fig. B) and C). In the figure, H3 is used to indicate A/H3N2, and H1 is used to indicate A/H1N1 before 2009, and A/H1N1pdm after 2009.

Global multiannual trends in (sub)type mixing

The composition trajectories for 151 countries are shown in Fig. 2A. The number of countries contributing to FluNet was initially limited - in 2000, only 30 countries contributed more than 50 samples, our threshold for inclusion -, but has substantially increased since then, e.g. 125 countries contributed in 2022.

We first analyzed global statistics on the ensemble of trajectories from one year to another. We introduced the *mixing score* to quantify the (sub)type mixing in a country/year. This is defined as the distance in the log-ratio plane between the country/year composition and the boundary of the co-dominance region. It ranges from positive values when the composition

is inside the co-dominance region to negative ones when it is outside, i.e. one (sub)type is dominant. The measure has an upper bound at 0.56 corresponding to equipartition of cases of a country/year among the three (sub)types, while it does not have lower bound. As an example, a strong dominance, where e.g. one (sub)type has frequency >75%, corresponds to mixing score values between -0.9 and -0.8. The distribution of countries' mixing scores in a given year provides a global overview of country-level (sub)type mixing. The multiannual comparison of Fig. 2B highlights anomalous events, when the mixing score was extraordinarily low, overall. A/H3N2 was strongly dominant (i.e. >75%) in 35 out of 45 countries in 2003, and A/H1N1pdm was strongly dominant in 80 out of 96 countries in 2009 (Fig. 2C). One (sub)type among A/H1N1pdm, A/H3N2, and B was strongly dominant in 13 out of 26 countries in 2020, and in 76 out of 100 countries in 2021.

Atypical values correspond to events occurring at the global scale in those years. The strong dominance of A/H3N2 in 2003 was caused by an emerging clade for which the vaccine had limited effect (29, 30). Similarly, the strong dominance of A/H1N1pdm in 2009 was due to the zoonotic emergence of the A/H1N1pdm subtype and the consequent pandemic (16, 31). In both examples, a punctuated change in the virus caused a strong level of dominance, which was associated with a high level of subtype synchrony across countries, and more intense viral circulation globally. During 2020 and 2021, on the other hand, different (sub)types strongly dominated in different regions - e.g. in 2020 A/H3N2 in Southeastern Asia, A/H1N1pdm in five African countries, and B in the Americas, Western Asia, and other countries (SI Appendix, Fig. S1). This was due to the unprecedented contact and mobility restrictions during COVID-19 (32–34), which caused a worldwide drop in influenza cases (only 26 countries reported at least 50 samples in 2020).

In the SI Appendix we analysed the mixing score under alternative log-ratio coordinates. We also show that the Shannon entropy provides a result similar to the mixing score. The advantage of the mixing score is that it provides an easy-to-read visualization of how far the composition is from co-dominance.

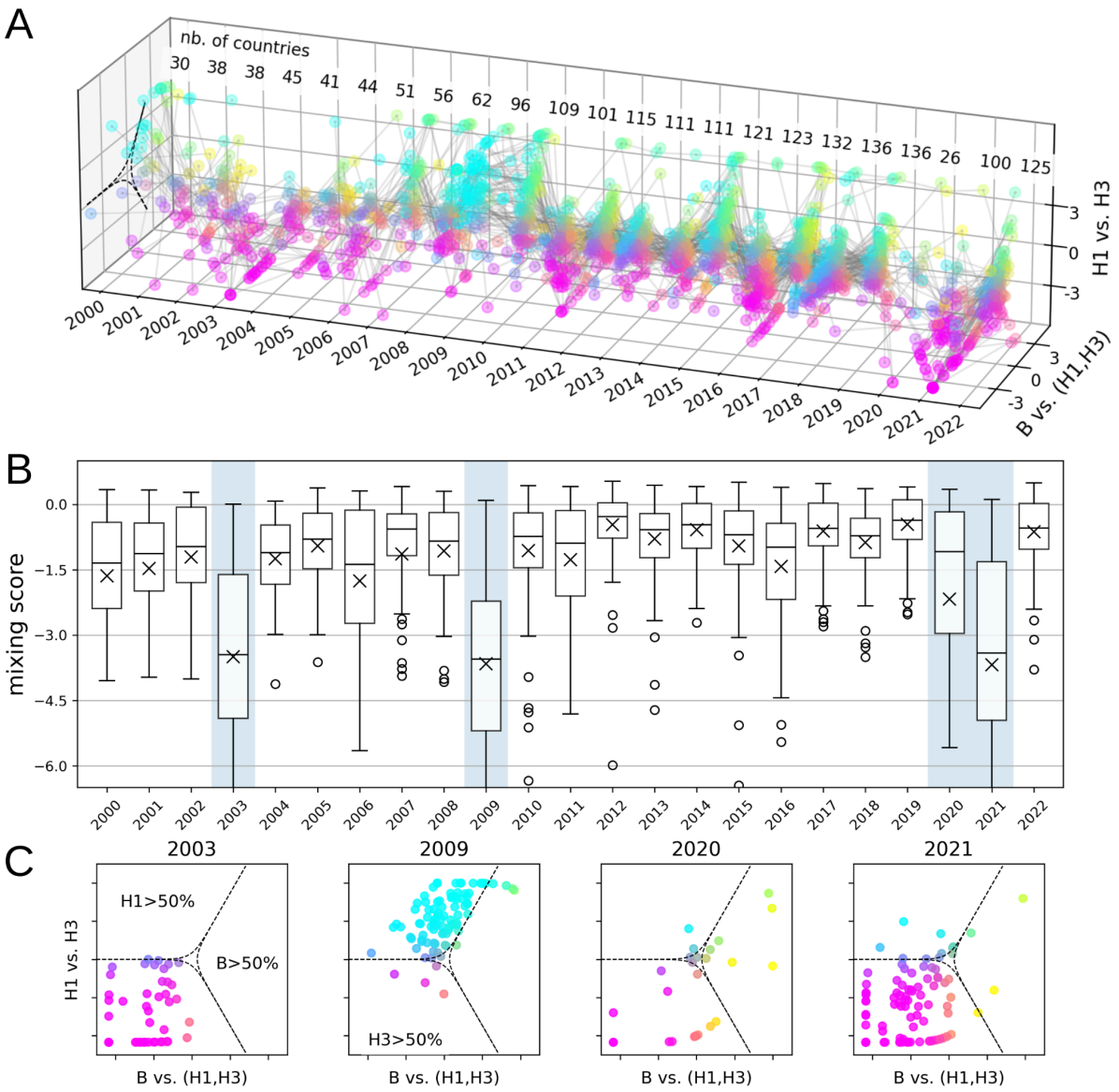

Fig. 2 - Trajectories of relative abundances of influenza (sub)types and global (sub)type mixing. A) Trajectories of relative abundances of influenza (sub)types. We present trajectories for 151 countries, including years from 2000 to 2022. Points in the plot represent countries/years for which the number of FluNet samples is above 50. The number of countries included by year is shown at the top. B) Degree of mixing of influenza (sub)types over time. For the years 2000 to 2022, the mixing score of influenza (sub)types was computed for each country, and their distributions are depicted through the boxplots. Positive scores represent countries where each (sub)type is responsible for <50% of cases, negative scores denote the dominance of one (sub)type. C) Influenza (sub)type abundances for atypical years. In 2003, 2009, 2020, and 2021, (sub)type mixing was unusually low. In 2003 and 2009, A/H3N2 and A/H1N1pdm, respectively, were dominant in almost all countries. Spatial segregation of influenza (sub)types occurred in 2020 and 2021, with different (sub)types dominant in different regions. In the figure, H3 is used to indicate A/H3N2, and H1 is used to indicate A/H1N1 before 2009, and A/H1N1pdm after 2009. Points are colored with the same triangular code as in Fig. 1.

Geographical patterns in (sub)type cocirculation

We then looked for geographical patterns in countries' trajectories from 2010 to 2019. During this period, a large number of countries contributed to FluNet - 81 countries' satisfied our inclusion criteria. At the same time, it was the longest time window of regular influenza circulation, in between the periods of 2009 A/H1N1pdm pandemic and COVID-19 pandemic. We clustered countries showing similar trajectories in the log-ratio plane. There were two main groups of countries (Fig. 3B), each one overlaping different WHO Influenza Transmission Zones (ITZs) (35). Group I included 39 countries belonging to the ITZs of Europe, Northern Africa, and Western Asia, with the addition of Iran, belonging to the Southern Asian ITZ, and the Republic of Korea from the Eastern Asian ITZ. These countries were characterized by trajectories showing strong and synchronous alternation between (sub)types (Fig. 3A). The 42 countries of Group II belonged to the other ITZs, except for Oman and Qatar which are included in the Western Asian ITZ (SI Appendix, Fig. S2). These countries displayed flatter trajectories overall (Fig. 3A). Looking more closely, Group II includes countries in the tropics, which normally see the co-circulation of all (sub)types throughout the annual period. It also includes countries located in temperate regions across both hemispheres which are characterized by a marked (sub)type alternation, yet not synchronous with Group I countries. This is the case of six Central and South American countries (Mexico, Nicaragua, El Salvador, Costa Rica, Argentina, Chile) and five Northern Hemisphere countries (United States, Canada, Japan, Mongolia, and Kazakhstan). In the SI Appendix, Fig. S2, we show that these two groups cluster separately after successive algorithm iterations. Examples of trajectories of the subgroups of Group II are reported in the SI Appendix, Fig. S3.

Groups were robust when testing for other clustering techniques and data inclusion criteria, together with alternative log-ratio coordinates for representing trajectories (SI Appendix).

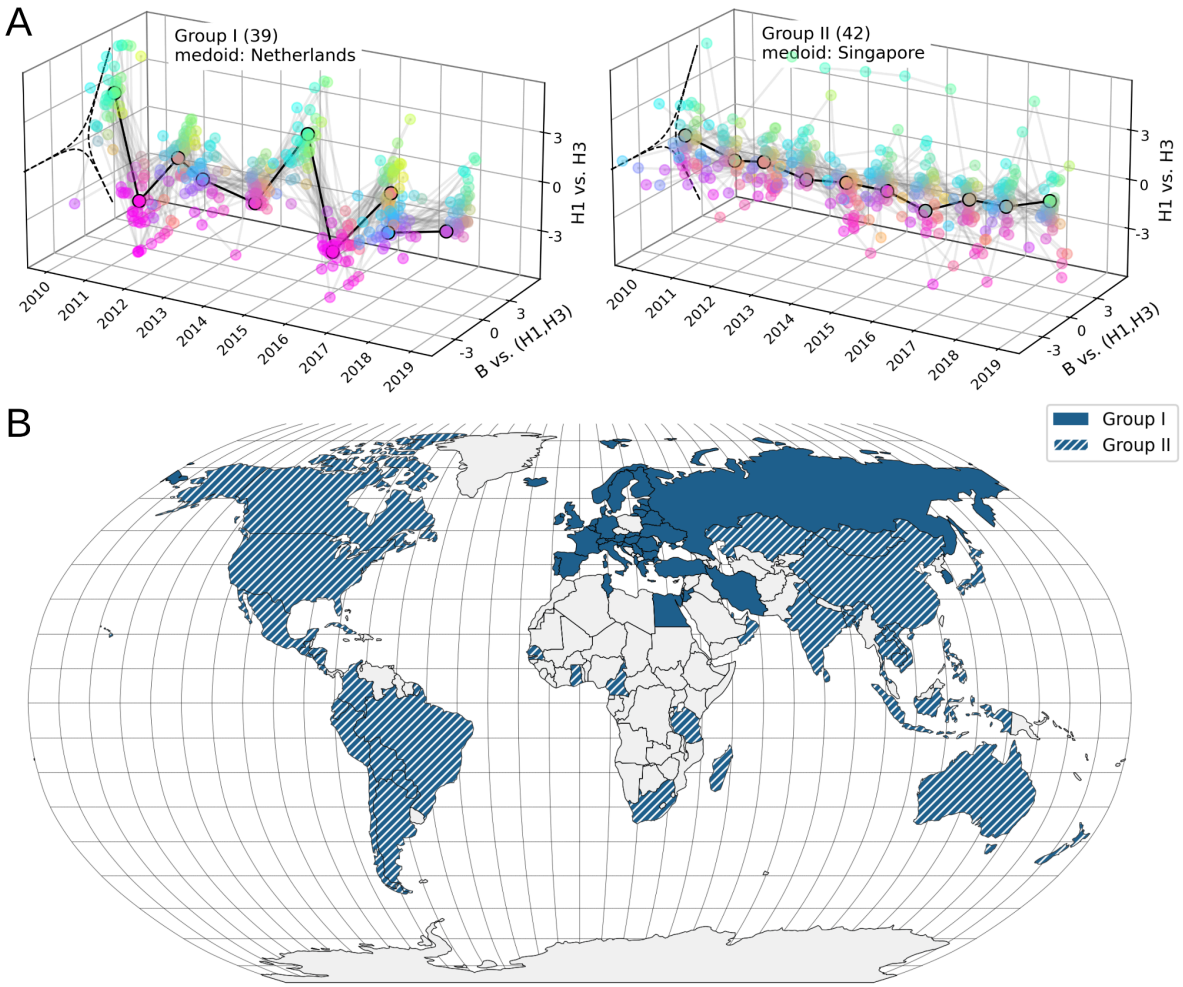

Fig. 3 - Geographical pattern of (sub)type trajectories. A) Hierarchical clustering of trajectories of A/H1N1, A/H3N2, and B relative abundances. Trajectories for 81 countries from 2010 to 2019 are clustered in two groups - groups I and II - with 39 and 42 countries, respectively. The medoid trajectories - i.e. the most central trajectories - of the two groups are highlighted in black and are reported in the legends. B) Geographic positioning of Group I and Group II countries. Source of shape files for the map: Natural Earth (36). In the figure, H3 is used to indicate A/H3N2, and H1 is used to indicate A/H1N1 before 2009, and A/H1N1pdm after 2009. Points are colored with the same triangular code as in Fig. 1. The co-dominance regions are shown in Panel A with dashed lines as a reference.

Forecasting of next-year influenza (sub)type composition

By examining (sub)type compositions as a trajectories over time, we can try to predict next-year's influenza (sub)type composition or dominance pattern. As there is currently no accepted approach to carry out this task, we consider several possibilities based on epidemiological considerations or using state of the art statistical models building on the CoDA framework. We focus on predicting compositions for 2017, 2018, and 2019 using the history of compositions observed worldwide since 2010. The models were as follows (see Material and Methods for details):

- *M1 past frequencies*: prediction for the coming year corresponds to the dominance/codominance status - between dominance of A/H1N1pdm, A/H3N2, B, or co-dominance - that occurred most often in previous years. Note that it does not predict a composition but only the dominance/co-dominance status.
- *M2 H1-H3 alternation:* we use the composition of the past year in the same country, exchanging the percentages of A/H1N1pdm and A/H3N2. This is based on the empirical evidence that A/H1N1pdm and A/H3N2 viruses tend to alternate, while influenza B often co-dominates (9, 16, 21).
- *M3 average composition:* we compute the average of previous years' compositions in the same country, after the log-ratio transformation.
- *M4 VAR*: we fit a Vector AutoRegressive (VAR) model using the data of all countries and obtain individual predictions for each country. In VAR models, the future composition is based on a linear combination of past ones, accounting for auto-correlation in time and for correlation between individual components (37). We used a VAR model with lag=1.
- *M5 HVAR*: we fit a Bayesian Hierarchical Vector AutoRegressive (HVAR) model to provide predictions for each individual country from past compositions of countries with similar trajectories (38). The approach extends the *M4 VAR* model and compensates for limited data over time by using information on (sub)type compositions over space. In practice, the trajectories of the training set were clustered as described in the previous section. The HVAR model was then fitted separately to each cluster, with the assumption that each trajectory follows a VAR process, but that the VAR processes of trajectories within the cluster cannot be too different. We tested both lag=1 and lag=2. Fitted VAR coefficients are analysed in the SI Appendix.

In models *M3*, *M4*, and *M5* we also computed prediction regions based on 95% probability intervals. Models *M1* and *M2* only allowed punctual predictions.

Table 1, top panel, shows the forecast performances of the models by looking at the percentages of correctly predicted dominance states - the *Dominance State Accuracy*. Results are presented globally and for Group I and Group II countries. Model *M1* had 28% accuracy overall, performing similarly to random guessing the dominance status out of the 4 possibilities, i.e, dominance of A/H1N1, A/H3N2, B, or co-dominance. Accuracy was lower for Group I (20%) than for Group II (34%), consistent with the difficulty in predicting the

marked subtype alternation of Group I countries. Other models (*M2*, *M3* and *M4*) had performances close to model *M1*. Estimates improved with *M5 HVAR*, with a global accuracy of 34%, improving in both Group I (27%) and Group II (41%).

Model	***M1 past frequencies***	***M2 H1-H3 alternation***	***M3 average composition***	***M4 VAR***	***M5 HVAR***
	Predicted observable: dominance state Score: Dominance State Accuracy				
Group I	0.200 ± 0.003	0.200 ± 0.003	0.200 ± 0.002	0.217 ± 0.003	0.270 ± 0.003
Group II	0.344 ± 0.003	0.328 ± 0.003	0.359 ± 0.003	0.289 ± 0.003	0.406 ± 0.003
all countries	0.276 ± 0.002	0.267 ± 0.002	0.284 ± 0.002	0.255 ± 0.002	0.342 ± 0.002
	Predicted observable: composition Score: Energy Score				
Group I	/	3.90 ± 0.01	1.85 ± 0.01	2.23 ± 0.01	1.49 ± 0.01
Group II	/	2.02 ± 0.01	1.33 ± 0.01	1.61 ± 0.01	1.17 ± 0.01
all countries	/	2.91 ± 0.01	1.58 ± 0.01	1.90 ± 0.01	1.32 ± 0.00

Table 1**. Evaluating influenza (sub)type forecasting.** Scores considered to compare the five predictive models are the *Dominance State Accuracy* for dominance state predictions and the *Energy Score* for predictions on compositions. Scores are averaged over predicted years (2017, 2018, 2019) and countries within the same group - i.e. Group I, Group II, and all the countries reported in different rows. Scores are reported with uncertainties expressed as Standard Errors of the Mean (SEM) calculated over 200 bootstrap samples. For the *M5 HVAR* model, we report the result with the lag that performs the best for each group, i.e. lag=2 for Group I and lag=1 for Group II. Methods that perform best by country groupings are highlighted for both the *Dominance State Accuracy* and the *Energy Score*. *Dominance State Accuracy* is positively oriented (i.e. it increases as a model performance increases), whereas the *Energy Score* is negatively oriented (i.e. it decreases as a model performance increases).

For models *M2, M3, M4,* and *M5*, which predict continuous (sub)type compositions, we can also use evaluation metrics for probabilistic forecasting commonly adopted in epidemiology (39), where both calibration (closeness to the target) and sharpness (uncertainty in the prediction) are taken into account. In Table 1, bottom panel, we show the values of the *Energy Score* (40) where smaller values correspond to better characteristics. We found again that the *M5 HVAR* model performed better than other models. Predictions for France (Group I) and Australia (Group II) using this model are illustrated in Fig. 4 as examples.

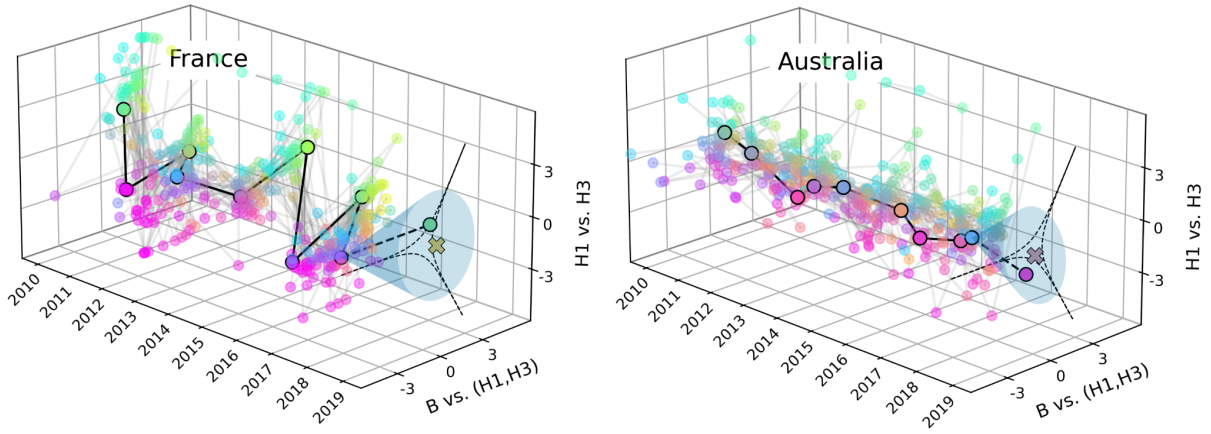

Fig. 4 - Prediction of relative abundances of influenza (sub)types for France and Australia in 2019. Predictions are computed using the *M5 HVAR* model of lag 2 and 1 (see Materials and Methods), respectively. For each country, the observed trajectory from 2010 to 2018 is represented with a black solid line, while the dashed segment links the points corresponding to 2019 - to be predicted - to the rest of the time series. The thinner gray lines correspond to the trajectories of the other countries within the respective group - Group I for France, and Group II for Australia - that were used to train the model. The crosses depict the predictions and the shadow areas the ellipses associated with the 95% confidence intervals. In the figure, H3 is used to indicate A/H3N2, and H1 is used to indicate A/H1N1pdm. Points are colored with the same triangular code as in Fig. 1. The co-dominance regions are shown with dashed lines as a reference. The *Energy Scores* for the 2019 predictions for France and Australia are 0.93 and 0.69, respectively.

Finally, we focused on simpler prediction tasks still relevant for public health. Precisely, we focused on each (sub)type in turn, and investigated whether it will be (i) dominant (>50% of cases) or not, and (ii) negligible (<10% of cases) or not. For model *M1 past frequencies,* the probability of negligibility of a (sub)type corresponded to the proportion of times the (sub)type was negligible in the previous years. To evaluate the performance of the five models in carrying out the two tasks we used the *Area Under the Receiver Operating Characteristic Curve (AUROC)*. Results are provided in Table 2.

The *M5 HVAR* model was the best performing model in predicting the negligibility of influenza type B and both dominance and negligibility of subtype A/H3N2. In particular, the *M5 HVAR* provided the greatest improvement when applied to the trajectories of Group I countries. For these countries, the *AUROC* score for the *M5 HVAR* model reaches 0.88 in predicting B negligibility, compared to scores of 0.37 to 0.60 obtained with other models. We found a similar improvement when predicting the dominance (negligibility) of A/H3N2, where the AUROC score for *M5 HVAR* was 0.82 (0.79) compared to 0.40-0.70 (0.36-0.69) for the other models. *M5 HVAR* was not always the best performing model, notably it was not always able to improve the predictions of the *M1 past frequencies* and *M3 average composition* models in predicting dominance of B and both dominance and negligibility of

A/H1N1. For the dominance of A/H1N1 for Group I countries no model performed significantly better than a random guess.

Predictions for all countries and years, obtained with the five methods are reported in the SI Dataset S1. In the sensitivity analysis we performed predictions with the alternative log-ratio transformation and using alternative metrics for the forecast evaluation. Results were overall consistent (see SI Appendix, Tables S1-S6).

Model	***M1 past frequencies***	***M2 H1-H3 alternation***	***M3 average composition***	***M4 VAR***	***M5 HVAR***
	Predicted observable: B will be dominant (>50%) in the next season: yes/no ? Score: AUROC				
Group I	0.452 ± 0.004	/	0.525 ± 0.004	0.435 ± 0.004	0.600 ± 0.004
Group II	0.606 ± 0.007	/	0.628 ± 0.006	0.457 ± 0.008	0.489 ± 0.013
all countries	0.541 ± 0.003	/	0.647 ± 0.003	0.510 ± 0.004	0.588 ± 0.004
	Predicted observable: H1 will be dominant (>50%) in the next season: yes/no ? Score: AUROC				
Group I	0.514 ± 0.004	/	0.423 ± 0.004	0.336 ± 0.004	0.450 ± 0.004
Group II	0.471 ± 0.004	/	0.521 ± 0.004	0.553 ± 0.004	0.612 ± 0.004
all countries	0.501 ± 0.003	/	0.491 ± 0.003	0.457 ± 0.003	0.529 ± 0.002
	Predicted observable: H3 will be dominant (>50%) in the next season: yes/no ? Score: AUROC				
Group I	0.401 ± 0.004	/	0.437 ± 0.005	0.703 ± 0.004	0.822 ± 0.003
Group II	0.566 ± 0.004	/	0.594 ± 0.004	0.540 ± 0.004	0.658 ± 0.004
all countries	0.497 ± 0.003	/	0.540 ± 0.003	0.614 ± 0.003	0.748 ± 0.003
	Predicted observable: B will be negligible (<10%) in the next season: yes/no ? Score: AUROC				
Group I	0.421 ± 0.004	/	0.366 ± 0.004	0.604 ± 0.004	0.876 ± 0.003
Group II	0.520 ± 0.006	/	0.529 ± 0.006	0.561 ± 0.005	0.649 ± 0.004
all countries	0.575 ± 0.003	/	0.523 ± 0.003	0.618 ± 0.003	0.804 ± 0.002
	Predicted observable: H1 will be negligible (<10%) in the next season: yes/no ? Score: AUROC				
Group I	0.620 ± 0.005	/	0.628 ± 0.007	0.257 ± 0.006	0.230 ± 0.005
Group II	0.639 ± 0.004	/	0.558 ± 0.005	0.594 ± 0.005	0.644 ± 0.005
all countries	0.598 ± 0.003	/	0.548 ± 0.004	0.442 ± 0.004	0.465 ± 0.005
	Predicted observable: H3 will be negligible (<10%) in the next season: yes/no ? Score: AUROC				
Group I	0.362 ± 0.004	/	0.425 ± 0.005	0.688 ± 0.004	0.788 ± 0.003
Group II	0.599 ± 0.004	/	0.503 ± 0.005	0.651 ± 0.006	0.641 ± 0.006
all countries	0.548 ± 0.003	/	0.556 ± 0.004	0.692 ± 0.003	0.743 ± 0.003

Table 2. **Forecast evaluation for predicting the dominance (yes/no) and negligibility (yes/no) of (sub)types.** Results are summarized in six panels: the three top ones focus on dominance and the three bottom ones on negligibility. As a score of prediction we used the *Area Under the Receiver Operating Characteristic*

Curve (AUROC). Scores were averaged over predicted years (2017, 2018, 2019) and countries within the same group - i.e. Group I, Group II, and all the countries reported in different rows. Scores are reported with uncertainties expressed as Standard Errors of the Mean (SEM) calculated over 200 bootstrap samples. The probability of dominance and negligibility of (sub)types could not be calculated with the *M2 H1-H3 alternation* model therefore the AUROC cannot be computed for this model. We highlight the method performing best by country grouping.

Discussion

Viral composition plays a key role in shaping influenza dynamics and the population burden of epidemics. A better anticipation of seasonal epidemics requires an improved ability to monitor, quantify, and predict (sub)type composition. This is increasingly within our reach thanks to the rapid increase of virological data on the global scale. Still, novel quantitative frameworks are needed to better visualize and analyze these data to extract meaningful information for response planning. Here we have shown that CoDA enables defining (sub)type composition trajectories thus opening the door to state-of-the-art statistical tools for analyzing their spatiotemporal properties and forecasting their future evolution.

We have analyzed (sub)type composition at the global level considering trajectories of annual compositions by country. This has revealed meaningful spatiotemporal patterns. First, we have characterized the evolution in time of the ensemble of trajectories by focusing on the extent of (sub)type mixing through the mixing score. We have highlighted years of strong dominance of a given (sub)type within each country – the strong dominance of A/H3H2 and A/H1N1pdm in almost all countries globally, respectively occurring in 2003 and 2009, and the spatial segregation of (sub)types occurring concomitantly to the COVID-19 pandemic. Past events of disruption of seasonal influenza activity were mainly analyzed by looking at changes in peak timing and severity of seasonal waves (16, 41). The mixing score introduced here provides a concise metric to quantify the impact of these events on (sub)type mixing for epidemiological interpretation. Strong (sub)type dominance can have marked effects on the age profile of cases. Mortality shifted toward young individuals during the 2009 A/H1N1pdm pandemic (42). This was linked to early-in-life infection with the A/H1N1 subtype circulating before 1950 that conferred protection against A/H1N1pdm to the older cohort (43). In addition, a sudden reduction of viral diversity can have long-term consequences on influenza ecology. The case of influenza during the COVID-19 pandemic provides a paradigmatic example as one clade of influenza B, B/Yamagata, has not been reported since the COVID-19 pandemic period (34, 44).

Second, considering a period free from disruption events, we investigated the geographical pattern of the (sub)type alternation. We used clustering to group countries with similar (sub)type trajectories. Grouping regions with similar characteristics is essential for surveillance. To this end, WHO defined the Influenza Transmission Zones as countries, areas or territories with similar influenza transmission patterns (35). A number of studies analyzed the similarity between countries' seasonal profiles to evaluate the validity of the partition (20, 45, 46). However, limited work addressed the problem by including (sub)type proportion as relevant information (20). Here we found that Group I - composed of countries from Europe, North Africa, and West Asia - was characterized by a synchronous alternation of (sub)types, clearly distinguishable from the rest of the world. The spatial synchrony in Europe following the 2009 influenza pandemic was noted before (20). It can be explained by the strong mobility coupling within the region and the consistent climatic conditions (20). Still, North America clustered in Group II despite being strongly connected with Europe and having a similar climate. The fact that the United States and Canada had a pattern of A/H1N1pdm and A/H3N2 alternation different from Europe the years that followed the A/H1N1pdm emergence was noted before (16, 21). It was hypothesized that this was due to the higher vaccination coverage in North America which mitigated the circulation of A/H3N2 preventing it from dominating over A/H1N1pdm during the 2011/2012 season (16). Interestingly, Australia and New Zealand had a different (sub)type alternation than Northern Hemisphere countries a few months later. They clustered in Group II, meaning that they were markedly different from Europe. They were also different from the US and Canada which were assigned to a different subcluster within Group II. Southern hemisphere countries are often regarded as sentinels of the approaching influenza season in northern countries (47–49). Still, the epidemiological evidence supporting this practice is contrasting (23, 50–52). For instance, the correlation between spatiotemporal indicators of Australia's wave and the waves in the UK, the US, and Canada was limited (51, 52). A similar clustering analysis can be repeated in the future to see whether the pattern remains robust following the disruption of influenza circulation caused by the COVID-19 pandemic, and therefore to identify robust priors predictive of sub-type circulation in the different areas.

Finally, the CoDa transformation made it possible to use state-of-the-art statistical forecasting algorithms such as VAR and hierarchical VAR to predict the (sub)type composition the following year. VAR is trained on past data, therefore we limited our analysis to the period 2010-2019 excluding the periods of disruption of circulation. The short length of the time series limited the power of VAR. Thus, the algorithm did not improve over naive estimates. Still, we could reach better performances through the hierarchical VAR model which leveraged the geographical structure identified by the clustering. Interestingly, despite

Group I countries being characterized by a marked A (sub)type alternation from one year to another the *M2 H1-H3 alternation* model did not perform well, likely due to its inability to capture the trend of B proportion. Besides forecasting (sub)type proportion and the dominant (sub)type we attempted to forecast binary outcome, e.g. whether one specific (sub)type would be dominant (yes/no) or negligible (yes/no). We found surprisingly accurate results when predicting the negligibility of type B and dominance and negligibility of subtype A/H3N2. For instance, for subtype A/H3N2 prediction of dominance, the AUROC score was 0.82 for Group I countries. Predicting the dominance A/H3N2 is important for preparedness since this is often associated with high epidemic burden.

Epidemic forecasting is a central problem in infectious disease epidemiology. Consortia of modelers are dedicating a great effort in projecting seasonal influenza epidemics in the short and long term. In the context of these initiatives, a wide range of mechanistic and statistical models are being combined in ensemble modelling to forecasting influenza incidence weeks ahead or to identify possible evolution scenarios months ahead (53–57). This provides critical information for response planning. Other studies focus on the evolution of influenza viruses, combining genetic and antigenic analyses to predict changes in the frequency of circulating clades for each (sub)type (58, 59). The identification of the clade more likely to be prevalent in the future informs the choice of vaccine composition which needs to be done six months in advance. Here we focused on (sub)type abundance and attempted to forecast this observable one year ahead. The (sub)type composition affects the expected epidemic size, the peak timing, and the age distribution of cases (6, 9, 13, 15). Therefore, its knowledge is precious for planning intervention measures. Predictions of next-year (sub)type composition, if reliable, would provide a long time horizon to plan the response, e.g. resource allocation, and targeted recommendation and vaccination. They could support scenario modeling and real-time incidence forecast. For instance, probabilistic projections of (sub)type composition could be used to weight possible scenarios. Knowing the dominant (sub)type at the start of the season would enable using statistical forecasting algorithms trained on the past seasons with the similar (sub)type frequency only, leading to higher prediction accuracy.

With this goal in mind, the accuracy of models tested is still limited and the analysis provided here should be regarded as a proof-of-concept. Still, the CoDA framework makes it possible to design more sophisticated models, incorporating demography, climate, air travel data, vaccine coverage, and other relevant covariates. In particular, vaccine coverage could provide critical information. Vaccination may impact (sub)type abundance because vaccine efficacy is different across (sub)types and age groups. Yet, no global database exists on vaccination coverage, which prevents the inclusion of this ingredient in global analyses. In

addition, new models could be designed to capture causal relationships between epidemics in different countries/years, e.g. through Bayesian networks (60). It is important to note that statistical forecasting models can be applied to periods of stable strain circulation, as was the case for 2010-2019. Following the COVID-19 pandemic, the global circulation of influenza viruses was highly perturbed, making any prediction more difficult.

CoDA is under-exploited in epidemiology. Analyses similar to the ones presented here could be applied to influenza (sub)type abundances at different spatiotemporal scales - weekly data to study the dynamics within a season (15), regional, or within-country spatial scale. More broadly, there is a variety of circulating viruses with distinct strains interfering with one another - SARS-CoV-2 variants, RSV types, HPV genotypes, and Dengue serotypes, to name a few. Percentage data are also used in monitoring the co-circulation of bacterial strains. This is the case of studying the dynamics of Staphylococcus aureus strains with different antibiotic resistance profiles (61), or vaccine-targeted vs. non-targeted strains of Streptococcus pneumonie (62).

Materials and Methods

FluNet data

FluNet publishes the weekly number of samples tested for influenza by country worldwide, classified as positive/negative and by influenza A subtype and B lineages (1, 26). For each country, we determine influenza B infections by summing B\Yamagata, B\Victoria, and unspecified B samples. We grouped the pre-pandemic A/H1N1 and A/H1N1pdm. Throughout the Material and Methods section we will use for brevity H1 for A/H1N1 and A/H1N1pdm and H3 for A/H3N2. The un-subtyped influenza A samples were distributed between H1 and H3 according to the respective proportions of these subtypes among the subtyped A samples that week. If no A samples were subtyped that week, we looked at the proportions of H1 and H3 in the five weeks centered around the week or in the year. We aggregated influenza B, H1, and H3 samples over one year and calculated the respective percentages to define the relative abundance of the three influenza strains. The time frame of a year considered in the analyses starts from the first Monday following April 23. By aggregating all samples worldwide and by averaging annual profiles between 1995 and 2019 we found that this week had the lowest proportion of positive influenza cases over the year. Thus setting the cut-off date this week minimized the risk of splitting the influenza epidemics in temperate countries into two consecutive years. In all analyses, we discarded

countries/years with fewer than 50 positive samples. We also tested an alternative threshold of 500 cases for robustness check.

Data pre-processing

Neither isometric log-ratio transformation defined in the section Results nor the *additive log-ratio transformation*, tested for robustness checks (see SI Appendix), are defined when any component equals zero. We thus first dealt with zeros assuming that they were due to insufficient sample sizes, rather than a real absence of the virus. We used a Geometric Bayesian-multiplicative treatment (63, 64). We first replaced the zero components with the Bayes estimator of a multinomial model, then rescaled all components to maintain their sum to one keeping at the same time their ratios. The multinomial model assumes that the number of infections per (sub)type follows a multinomial distribution, whose parameters are distributed with a Dirichlet prior. This choice leads to a posterior distribution which is still a Dirichlet distribution, from which the Bayes estimator (i.e. the expectation of the posterior distribution) is easily computed.

Definition of the *mixing score*

The *mixing score* is defined as the distance between the point in the log-ratio plane (u,v), representing the (sub)type composition, and the boundary of the co-dominance region, taken with a positive sign when the point is within that region and with a negative sign otherwise. The boundary of the co-dominance region in the simplex is identified by the points (B%, H1%, H3%) for which one component corresponds to exactly 50%. As an example, the points such that H1%=50%, are mapped by the *ilr* transformation into the coordinates (u, v(u)) such that

$$v(u) = \sqrt{2}[\ln(e^{u\sqrt{3/2}} + \sqrt{e^{u\sqrt{6}} + 4}) - \ln 2].$$

Clustering of trajectories

We defined the distances between any pair of country trajectories as the average Euclidean distance of the corresponding points in the (u,v) plane. We then applied the Ward's linkage hierarchical clustering (65). Hierarchical clustering provides a hierarchy of nested partitions. We used the Silhouette coefficients (66) to choose the optimal number of clusters, i.e. the level of hierarchy. The optimal partition grouped countries into two groups. Robustness checks and sensitivity analyses (including alternative clustering algorithms) are described in the SI Appendix.

Forecasting of trajectories

We tested five forecasting algorithms to predict next-year (sub)type composition in each country. We attempted to predict observables of different nature: (i) the multi-label categorical variable corresponding to the dominance state (dominance of B, H1, H3, or co-dominance); (ii) the (sub)type composition; (iii) the binary variables corresponsidng to the dominance (yes or no) and the negligibility (yes or no) of each (sub)type. Predictions of the three observables where associated to a qantification of the uncertainty in the prediction and a metric for prediction evaluation. However, not all observables, quantifications of uncertainty and prediction evaluations were computable with all prediction methods. The five forecasting algorithms and the evaluation metrics used in the main paper were introduced in the Results section. We provide hereafter additional details. A complete summary of what can or cannot be computed across predicted observables and prediction methods is reported in the SI Appendix, Table S1.

Algorithms for trajectory forecasting. We studied annual bivariate trajectories of compositions of the form (y_1, … , y_T), such that y_t=(u,v)'$_t$=*ilr*((B%, H1%, H3%)'$_t$), with t=1,...,T and the single quote mark indicating transposed. We forecasted y_{T+1}, i.e. the composition of 2017, 2018 or 2019, each time using the preceeding period as training set, i.e. (y_1, … , y_T) with y_1 the composition in 2010. Below we present the five forecasting methods.

M1 past frequencies: the dominance state was predicted to be the most frequently observed state during the preceeding years, if that state was unique. Otherwise, it was randomly chosen with uniform probability among the dominant states observed most often.

M2 H1-H3 alternation: the prediction for the y_{T+1} composition was $\hat{y}_{T+1}$=(u,-v)'$_T$, i.e., in percentage form, (B%$_{T+1}$=B%$_T$, H1%$_{T+1}$=H3%$_T$, H3%$_{T+1}$=H1%$_T$)'.

M3 average composition: the predicted composition was $\hat{y}_{T+1} = \frac{1}{T}\sum_{t=1}^{T} y_t$, and the prediction's confidence interval was given by the covariance matrix $\hat{\Sigma} = \frac{1}{T(T-1)}\sum_{t=1}^{T}(y_t - \hat{y}_{T+1})'(y - \hat{y}_{T+1})$.

M4 VAR: assuming that the trajectory has been generated by a VAR process of lag *p,* then composition *t* can be written as a linear function of the previous *p* compositions: $y_t = \nu + A^{(1)}y_{t-1} + \cdots + A^{(p)}y_{t-p} + \epsilon_t$, where ν is the vector intercept, $A^{(i)}$ are 2x2

coefficient matrices, and ϵ is the Gaussian noise. This can be rewritten in matrix form as $Y = BZ + U$ (37), where

$$
\begin{aligned}
Y &= (y_{p+1}, \ldots, y_T), \\
B &= (\nu, A^{(1)}, \ldots, A^{(p)}), \\
Z_t &= \begin{pmatrix} 1 \\ y_p \\ \vdots \\ y_1 \end{pmatrix}, \\
Z &= (Z_p, \ldots, Z_{T-1}), \\
U &= (\epsilon_{p+1}, \ldots, \epsilon_T).
\end{aligned}
\tag{2}
$$

The resulting least squares estimator is $\hat{B} = YZ'(ZZ')^{-1}$. It follows that the prediction of the composition T+1 is $\hat{y}_{T+1} = BZ_T$, and the prediction's confidence interval is estimated via the empirical covariance matrix corrected for short time-series:
$\hat{\Sigma} = \frac{T-p+1}{(T-p)(T-3p-1)}(YY' - YZ'(ZZ')ZY')$.
We only considered VAR processes with p=1 , since for the shortest trajectories (prediction of 2017, with training set between 2010 and 2016), when p>1, the estimator of the covariance matrix is not defined.

M5 HVAR: the trajectories of the training set were clustered as described above in the Method section. For all three forecasted years, this yealded two groups that were almost identical to Group I and Group II obtained for the whole period 2010-2019 and discussed in the Results. For each cluster, we assumed that each trajectory followed a VAR process, such that the process for country *c* is: $y_{t,c} = \nu_c + A_c^{(1)} y_{t-1,c} + \cdots + A_c^{(p)} y_{t-p,c} + \epsilon_{t,c}$. Similarly to the previous model, the coefficients can be encoded in the matrix $B_c = (\nu_c, A_c^{(1)}, \ldots, A_c^{(p)})$, and the Gaussian noise in the matrix U_c. Then, the hierarchical structure is imposed by assuming that the VAR processes for the trajectories in the group are similar. Specifically, we define $B_c = W + V_c$, with W being the matrix of coefficients encoding the average behavior of the group, that is the same for all the trajectories in the group, and V_c being the coefficient matrix for the single trajectory adjustment. Moreover, we assume that elements in W, V_c, U_c are independent random variables, sampled from distributions parametrized by some latent variables. All model equations are shown in the SI Appendix. The model coefficients were optimized by maximum likelihood estimation using Monte Carlo methods. In particular, since for each coefficient it was possible to write the mode'sl likelihood conditional on the other coefficients, efficient estimation was performed using a Gibb sampler. For the conditional

likelihood distributions and the code to implement the Gibb sampler, we followed (38), modifying their model to introduce the intercept terms. From the Monte Carlo chains, for each country, we sampled the distribution of predictions $\hat{y}_{T+1}$. Then, from those samples, we computed the mean prediction and the covariance matrix for the confidence intervals. We tested HVAR models with p=1 and p=2.

Metrics for evaluating the goodness of predictions. To enable model comparison, the evaluation scores of different countries/years were averaged to obtain a single value for the model performance. Confidence interval was computed with bootstrap. Specifically, the values reported in Table 1 correspond to the averages of the *Dominance State Accuracy* and *Energy Score* values for different countries/years, while the values in Table 2 are averages of the country/year *Area Under the Receiver Operating Characteristic Curve (AUROC)* scores.

To compare dominance state predictions (Table 1) we defined the *Dominance State Accuracy* as the percentage of dominance states correctly predicted. To compare predictions of compositions (Table 1) we instead relied on probabilistic forecasting evaluation scores. Specifically, the *Energy Score* compares the composition $y \in \mathbb{R}^2$ corresponding to the observation with the forecast distribution *F* defined by N samples $X_1, \ldots, X_N$, $X_i \in \mathbb{R}^2$. The formula for the *Eenergy Score* is

$$ES(F, y) = \frac{1}{N} \sum_{i=1}^{N} ||X_i - y|| - \frac{1}{2N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} ||X_i - X_j||$$

It is worth noting that this is a multivariate generalization of the more common *Continuous Rank Probability Score*, often used in epidemiology (39). Moreover, in the case of point forecast, it coincides with the *Mean Absolute Error*. Thus it can be also used for evaluating the *M2 H1-H3 alternation* method for which we do not have a confidence interval. The score is negatively-oriented, so that it decreases when the forecast improves, and has a lower bound of zero for an ideal model. Furthermore, it is a *strictly proper* score, i.e. designed to be optimial only when the forecast distribution coincides with the true distribution of the observations (40). Binary predictions - i.e., dominance (yes/no) and negligibility (yes/no) of each (sub)type - were evaluated through the *Area Under the Receiver Operating Characteristic Curve (AUROC)* score (Table 2). It quantifies the overlap between the positive and negative classes and ranges from 0 to 1: 1 for an ideal model where the classes are perfectly separated, 0.5 for a random guess, and 0 for the worst possible model that misclassifies all observations.

For the *M5 HVAR* model, we tested p=1 and p=2. The results for the *M5 HVAR* model in all tables are obtained using the value of p with the lowest *Energy Score* for each group of countries, i.e. p=1 for Group II countries and p=2 for Group I countries.

Other evaluation metrics where tested as sensitivity analysis. They are described in the SI Appendix, Tables S1, while results of these metric are reported in the SI Appendix, Tables S2-S6.

Code and data availability

Analysis was implemented in R (version 4.3.2) and Python (version 3.8.5). Other than standard packages for data treatment, plots and calculations (mainly the python packages *pandas, matplotlib,* and *numpy*), we relied on the following packages for specific tasks:

- *zComposition 1.4.0-1* (R) for zero imputation in the pre-processing of compositions (64);
- *robCompositions 2.3.1* (R) for mapping compositions from the Simplex to the Euclidean space and back (67);
- *ternary* (python) for drawing ternary plots (68);
- *scipy 1.6.2* (python) for clustering analysis;
- *sklearn 1.3.2* (python) for clustering analysis and computation of some of the prediction evaluation scores (69);
- R code developed by Lu et al (38), on the basis of which we performed the VAR and HVAR predictions;
- *scoringRules 1.0.2* (R) for calculation of proper scores for probabilistic forecast evaluation (70).

Code and data for reproducible analyses are available at https://github.com/FrancescoBonacina/coupled-dynamics-flu-subtypes/ .

Conflict of interest

The authors declare no conflict of interest.

Acknowledgment

We thank Eugenio Valdano for the inputs on the analysis. This work was partially supported by: the EU H2020 grant MOOD (H2020-874850) to V.C.; the ANR project DATAREDUX(ANR-19-CE46-0008-03) to V.C.; the Municipality of Paris through the programme Emergence(s) to C.P. and F.B.; Cariparo Foundation through the program

Starting Package to C.P.; Department of Molecular Medicine through the program SID from BIRD funding to C.P.

Author contribution

F.B., P.Y.B, V.C., O.L., M.T, and C.P. designed research; F.B. performed research; F.B, P.Y.B., and C.P. analyzed data; F.B., P.Y.B, V.C., O.L., M.T, and C.P. wrote the paper.

References

1. GISRS, FluNet database - National Influenza Centres (NICs) of the Global Influenza Surveillance and Response System (GISRS) and World Health Organisation (WHO). (2022). Available at: https://www.who.int/tools/flunet [Accessed 8 April 2022].
2. P. Palese, T. T. Wang, Why Do Influenza Virus Subtypes Die Out? A Hypothesis. *mBio* **2**, e00150-11 (2011).
3. K. L. Laurie, *et al.*, Interval Between Infections and Viral Hierarchy Are Determinants of Viral Interference Following Influenza Virus Infection in a Ferret Model. *J. Infect. Dis.* **212**, 1701–1710 (2015).
4. S. Caini, *et al.*, Distribution of influenza virus types by age using case-based global surveillance data from twenty-nine countries, 1999-2014. *BMC Infect. Dis.* **18**, 269 (2018).
5. A. Gagnon, E. Acosta, M. S. Miller, Age-Specific Incidence of Influenza A Responds to Change in Virus Subtype Dominance. *Clin. Infect. Dis.* **71**, e195–e198 (2020).
6. M. C. Vieira, *et al.*, Lineage-specific protection and immune imprinting shape the age distributions of influenza B cases. *Nat. Commun.* **12**, 4313 (2021).
7. M. An der Heiden, U. Buchholz, Estimation of influenza-attributable medically attended acute respiratory illness by influenza type/subtype and age, Germany, 2001/02-2014/15. *Influenza Other Respir. Viruses* **11**, 110–121 (2017).
8. S. Puzelli, *et al.*, Co-circulation of the two influenza B lineages during 13 consecutive influenza surveillance seasons in Italy, 2004-2017. *BMC Infect. Dis.* **19**, 990 (2019).
9. B. S. Finkelman, *et al.*, Global Patterns in Seasonal Activity of Influenza A/H3N2, A/H1N1, and B from 1997 to 2005: Viral Coexistence and Latitudinal Gradients. *PLoS ONE* **2**, e1296 (2007).
10. A. Suzuki, K. Mizumoto, A. R. Akhmetzhanov, H. Nishiura, Interaction Among Influenza Viruses A/H1N1, A/H3N2, and B in Japan. *Int. J. Environ. Res. Public. Health* **16**, 4179 (2019).
11. E. Goldstein, S. Cobey, S. Takahashi, J. C. Miller, M. Lipsitch, Predicting the Epidemic Sizes of Influenza A/H1N1, A/H3N2, and B: A Statistical Method. *PLOS Med.* **8**, e1001051 (2011).
12. K. L. Laurie, *et al.*, Evidence for Viral Interference and Cross-reactive Protective Immunity Between Influenza B Virus Lineages. *J. Infect. Dis.* **217**, 548–559 (2018).
13. W. Yang, E. H. Y. Lau, B. J. Cowling, Dynamic interactions of influenza viruses in Hong Kong during 1998-2018. *PLOS Comput. Biol.* **16**, e1007989 (2020).
14. J. Truscott, *et al.*, Essential epidemiological mechanisms underpinning the transmission dynamics of seasonal influenza. *J. R. Soc. Interface* **9**, 304–312 (2011).
15. E. Goldstein, S. Cobey, S. Takahashi, J. C. Miller, M. Lipsitch, Predicting the Epidemic Sizes of Influenza A/H1N1, A/H3N2, and B: A Statistical Method. *PLOS Med.* **8**, e1001051 (2011).
16. D. He, *et al.*, Global Spatio-temporal Patterns of Influenza in the Post-pandemic Era. *Sci. Rep.* **5**, 11013 (2015).
17. L. Gatti, *et al.*, Cross-reactive immunity potentially drives global oscillation and opposed alternation patterns of seasonal influenza A viruses. *Sci. Rep.* **12**, 8883 (2022).

18. T. Bedford, *et al.*, Global circulation patterns of seasonal influenza viruses vary with antigenic drift. *Nature* **523**, 217–220 (2015).
19. F. Parino, *et al.*, Integrating dynamical modeling and phylogeographic inference to characterize global influenza circulation. [Preprint] (2024). Available at: https://www.medrxiv.org/content/10.1101/2024.03.14.24303719v1 [Accessed 18 April 2024].
20. S. Caini, W. J. Alonso, C. E.-G. Séblain, F. Schellevis, J. Paget, The spatiotemporal characteristics of influenza A and B in the WHO European Region: can one define influenza transmission zones in Europe? *Eurosurveillance* **22**, 30606 (2017).
21. P. Mook, *et al.*, Alternating patterns of seasonal influenza activity in the WHO European Region following the 2009 pandemic, 2010-2018. *Influenza Other Respir. Viruses* **14**, 150–161 (2020).
22. L. Zheng, *et al.*, Global variability of influenza activity and virus subtype circulation from 2011 to 2023. *BMJ Open Respir. Res.* **10**, e001638 (2023).
23. S. B. Choi, J. Kim, I. Ahn, Forecasting type-specific seasonal influenza after 26 weeks in the United States using influenza activities in other countries. *PLOS ONE* **14**, e0220423 (2019).
24. J. Aitchison, The Statistical Analysis of Compositional Data. *J. R. Stat. Soc. Ser. B Methodol.* **44**, 139–160 (1982).
25. P. Filzmoser, K. Hron, M. Templ, *Applied Compositional Data Analysis: With Worked Examples in R* (Springer International Publishing, 2018).
26. A. Flahault, *et al.*, FluNet as a Tool for Global Monitoring of Influenza on the Web. *JAMA* **280**, 1330–1332 (1998).
27. F. Chayes, On correlation between variables of constant sum. *J. Geophys. Res. 1896-1977* **65**, 4185–4193 (1960).
28. J. J. Egozcue, V. Pawlowsky-Glahn, G. Mateu-Figueras, C. Barceló-Vidal, Isometric Logratio Transformations for Compositional Data Analysis. *Math. Geol.* **35**, 279–300 (2003).
29. E. Ghedin, *et al.*, Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature* **437**, 1162–1166 (2005).
30. CDC, “Update: Influenza Activity --- United States and Worldwide, 2003--04 Season, and Composition of the 2004--05 Influenza Vaccine” (2004).
31. C. Fraser, *et al.*, Pandemic Potential of a Strain of Influenza A (H1N1): Early Findings. *Science* **324**, 1557–1561 (2009).
32. V. Dhanasekaran, *et al.*, “Human seasonal influenza under COVID-19 and the potential consequences of influenza lineage elimination” (In Review, 2021).
33. F. Bonacina, *et al.*, Global patterns and drivers of influenza decline during the COVID-19 pandemic. *Int. J. Infect. Dis.* **128**, 132–139 (2023).
34. Z. Chen, *et al.*, COVID-19 pandemic re-shaped the global dispersal of seasonal influenza viruses. [Preprint] (2023). Available at: https://www.medrxiv.org/content/10.1101/2023.12.20.23300299v1 [Accessed 2 April 2024].
35. Influenza Transmission Zones. Available at: https://www.who.int/publications/m/item/influenza_transmission_zones [Accessed 18 April 2024].
36. Natural Earth, Natural Earth - Free vector and raster map data. Available at: https://www.naturalearthdata.com/ [Accessed 8 February 2024].
37. H. Lütkepohl, *New introduction to multiple time series analysis: with ... 36 tables*, 1. ed., corr. 2. print (Springer, 2007).
38. F. Lu, Y. Zheng, H. Cleveland, C. Burton, D. Madigan, Bayesian hierarchical vector autoregressive models for patient-level predictive modeling. *PLOS ONE* **13**, e0208082 (2018).
39. J. Bracher, E. L. Ray, T. Gneiting, N. G. Reich, Evaluating epidemic forecasts in an interval format. *PLOS Comput. Biol.* **17**, e1008618 (2021).
40. T. Gneiting, A. E. Raftery, Strictly Proper Scoring Rules, Prediction, and Estimation. *J.*

Am. Stat. Assoc. **102**, 359–378 (2007).
41. M. A. Miller, C. Viboud, M. Balinska, L. Simonsen, The Signature Features of Influenza Pandemics — Implications for Policy. *N. Engl. J. Med.* **360**, 2595–2598 (2009).
42. M. Lemaitre, F. Carrat, Comparative age distribution of influenza morbidity and mortality during seasonal influenza epidemics and the 2009 H1N1 pandemic. *BMC Infect. Dis.* **10**, 162 (2010).
43. I. Skountzou, *et al.*, Immunity to pre-1950 H1N1 influenza viruses confers cross-protection against the pandemic swine-origin 2009 A (H1N1) influenza virus. *J. Immunol. Baltim. Md 1950* **185**, 1642–1649 (2010).
44. S. Caini, *et al.*, Probable extinction of influenza B/Yamagata and its public health implications: a systematic literature review and assessment of global surveillance databases. *Lancet Microbe* **0** (2024).
45. C. Chen, *et al.*, The global region-specific epidemiologic characteristics of influenza: World Health Organization FluNet data from 1996 to 2021. *Int. J. Infect. Dis.* **129**, 118–124 (2023).
46. W. J. Alonso, *et al.*, A global map of hemispheric influenza vaccine recommendations based on local patterns of viral circulation. *Sci. Rep.* **5**, 17214 (2015).
47. D. F. Hughes, E. O'Neill, Flu season started early in Australia – countries in the northern hemisphere took note. *The Conversation* (2023). Available at: http://theconversation.com/flu-season-started-early-in-australia-countries-in-the-northern-hemisphere-took-note-207967 [Accessed 17 April 2024].
48. M. G. Baker, H. Kelly, N. Wilson, Pandemic H1N1 influenza lessons from the southern hemisphere. *Eurosurveillance* **14**, 19370 (2009).
49. E. Harris, Southern Hemisphere Outcomes Support Flu Vaccine Effectiveness. *JAMA* **330**, 1318–1319 (2023).
50. Y. Zhang, L. Yakob, M. B. Bonsall, W. Hu, Predicting seasonal influenza epidemics using cross-hemisphere influenza surveillance data and local internet query data. *Sci. Rep.* **9**, 3262 (2019).
51. C. Viboud, *et al.*, Influenza Epidemics in the United States, France, and Australia, 1972–19971. *Emerg. Infect. Dis.* **10**, 32–39 (2004).
52. P. H. A. of Canada, Does the Australian influenza season predict the Canadian influenza season? CCDR 49(11/12). (2023). Available at: https://www.canada.ca/en/public-health/services/reports-publications/canada-communicable-disease-report-ccdr/monthly-issue/2023-49/issue-11-12-november-december-2023/does-australian-influenza-season-predict-canadian-influenza-season-2014-2020.html [Accessed 17 April 2024].
53. S. M. Mathis, *et al.*, Evaluation of FluSight influenza forecasting in the 2021–22 and 2022–23 seasons with a new target laboratory-confirmed influenza hospitalizations. *medRxiv* 2023.12.08.23299726 (2023). https://doi.org/10.1101/2023.12.08.23299726.
54. RespiCast | European Respiratory Diseases Forecasting Hub. Available at: https://respicast.ecdc.europa.eu/ [Accessed 18 April 2024].
55. CDC, About Flu Forecasting | CDC. (2023). Available at: https://www.cdc.gov/flu/weekly/flusight/how-flu-forecasting.htm [Accessed 6 January 2024].
56. Influcast. *Influcast*. Available at: https://influcast.org/ [Accessed 18 April 2024].
57. Home - Scenario model hub. Available at: https://scenariomodelinghub.org/ [Accessed 27 June 2024].
58. M. Łuksza, M. Lässig, A predictive fitness model for influenza. *Nature* **507**, 57–61 (2014).
59. R. A. Neher, T. Bedford, R. S. Daniels, C. A. Russell, B. I. Shraiman, Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E1701-1709 (2016).
60. D. Heckerman, Bayesian Networks for Data Mining. *Data Min. Knowl. Discov.* **1**, 79–119 (1997).
61. F. Pinotti, É. Fleury, D. Guillemot, P.-Y. Böelle, C. Poletto, Host contact dynamics

shapes richness and dominance of pathogen strains. *PLOS Comput. Biol.* **15**, e1006530 (2019).
62. F. Blanquart, S. Lehtinen, C. Fraser, An evolutionary model to predict the frequency of antibiotic resistance under seasonal antibiotic use, and an application to Streptococcus pneumoniae. *Proc. R. Soc. B Biol. Sci.* **284**, 20170679 (2017).
63. J.-A. Martín-Fernández, K. Hron, M. Templ, P. Filzmoser, J. Palarea-Albaladejo, Bayesian-multiplicative treatment of count zeros in compositional data sets. *Stat. Model.* **15**, 134–158 (2015).
64. J. Palarea-Albaladejo, J. A. Martín-Fernández, zCompositions — R package for multivariate imputation of left-censored data under a compositional approach. *Chemom. Intell. Lab. Syst.* **143**, 85–96 (2015).
65. J. H. Ward Jr., Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* **58**, 236–244 (1963).
66. P. J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).
67. M. Templ, K. Hron, P. Filzmoser, “robCompositions: An R-package for Robust Statistical Analysis of Compositional Data” in *Compositional Data Analysis*, (John Wiley & Sons, Ltd, 2011), pp. 341–355.
68. M. Harper, *et al.*, python-ternary: Ternary Plots in Python. (2015). https://doi.org/10.5281/zenodo.34938. Deposited 7 December 2015.
69. F. Pedregosa, *et al.*, Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
70. A. Jordan, F. Krüger, S. Lerch, Evaluating Probabilistic Forecasts with scoringRules. *J. Stat. Softw.* **90**, 1–37 (2019).