



Prediction Model

>>>>>

2 3 r d N o v e m b e r



FluNet



- FluNet is a global web-based tool for influenza virological surveillance.
- The data at country level are publically available and updated weekly.
- NICs collect virus specimens in their country and perform preliminary analysis.





Collect Influenza Surveillance Data

FluNet functions

Influenza laboratory surveillance data

[Download data from the latest week](#) >

[Download data from any week](#) >

Select by: [Country, area or territory](#) ▾

- Country, area or territory
- WHO region
- Influenza transmission zone



Collect Influenza Surveillance Data

Select by: Filter by:

- North America
- Central America and Caribbean
- Temperate South America
- Tropical South America
- Northern Europe
- South West Europe

Year from	<input type="button" value="2021"/>	Week from	<input type="button" value="1"/>	Year to	<input type="button" value="2021"/>	Week to	<input type="button" value="1"/>	<input type="button" value="Display report"/>
	<input type="button" value="2020"/>		<input type="button" value="2"/>		<input type="button" value="2020"/>		<input type="button" value="2"/>	
	<input type="button" value="2019"/>		<input type="button" value="3"/>		<input type="button" value="2019"/>		<input type="button" value="3"/>	
	<input type="button" value="2018"/>		<input type="button" value="4"/>		<input type="button" value="2018"/>		<input type="button" value="4"/>	
	<input type="button" value="2017"/>		<input type="button" value="5"/>		<input type="button" value="2017"/>		<input type="button" value="5"/>	
	<input type="button" value="2016"/>		<input type="button" value="6"/>		<input type="button" value="2016"/>		<input type="button" value="6"/>	



Collect Influenza Surveillance Data

Surveillance site type : All sites Calendar week : ISO 8601

1 of 1 Find | Next

Influenza Laboratory Surveillance System

Data source: FluNet (www.who.int/flunet), Global Influenza Program

Global Influenza Surveillance and Response System (GISRS)

Country, area or territory	WHO region	Influenza transmission zone	Year	Week	Specimen date	Report date	Number of specimens		Number of influenza A viruses detected by subtype					
							Received/ Selected	Processed	A (H1)	A (H1N1)pdm09	A (H3)	A (H5)	A (not subtyped)	A (Total)
Bermuda	Region of the Americas of WHO	North America	2019	1	2018-12-31	2019-01-06								
Bermuda	Region of the Americas of WHO	North America	2019	2	2019-01-07	2019-01-13			1		1			1
Bermuda	Region of the Americas of WHO	North America	2019	3	2019-01-14	2019-01-20			2		2			2
Bermuda	Region of the Americas of WHO	North America	2019	4	2019-01-21	2019-01-27			1		1			1
Bermuda	Region of the Americas of WHO	North America	2019	5	2019-01-28	2019-02-03								
Bermuda	Region of the Americas of WHO	North America	2019	6	2019-02-04	2019-02-10								

Word
Excel
PowerPoint
PDF
TIFF file
MHTML (web archive)
CSV (comma delimited)
XML file with report data
Data Feed



Collect Influenza Surveillance Data

Country	WHOREGION	FLUREGION	Year	Week	SDATE	EDATE
Canada	Region of the Americas of WHO	North America	2019	1	2018.12.31	2019.1.6
Canada	Region of the Americas of WHO	North America	2019	2	2019.1.7	2019.1.13
Canada	Region of the Americas of WHO	North America	2019	3	2019.1.14	2019.1.20
Canada	Region of the Americas of WHO	North America	2019	4	2019.1.21	2019.1.27
Canada	Region of the Americas of WHO	North America	2019	5	2019.1.28	2019.2.3
Canada	Region of the Americas of WHO	North America	2019	6	2019.2.4	2019.2.10
Canada	Region of the Americas of WHO	North America	2019	7	2019.2.11	2019.2.17
Canada	Region of the Americas of WHO	North America	2019	8	2019.2.18	2019.2.24

INF_A	BYAMAGATA	BVICTORIA	BNOTDETERM	INF_B	ALL_INF
3338			31	31	3369
3139			43	43	3182
2299			40	40	2339
2108			48	48	2156
1994			43	43	2037
1667			42	42	1709
1488			33	33	1521
1518			40	40	1558
1670			64	64	1734
1688			71	71	1759



Example code: make Database





1. Create Table in MySQL

- now = datetime.datetime.now()
- nowDate = now.strftime('%Y%m%d')
- **tableName** = nowDate + "_fluNet"
- **connection** = pymysql.connect(host='localhost', user='_mysql(userID', password='_user_password', db='_your_db_name', charset='utf8', autocommit=True, cursorclass = pymysql.cursors.DictCursor)
- **cursor** = connection.cursor()
-
- **sql** = 'CREATE TABLE ' + **tableName** + '(id int(10) NOT NULL AUTO_INCREMENT PRIMARY KEY, year int, week int, s_month int, s_day int, e_month int, e_day int, flu_region text, who_region text, country text, h1_inf int, h3_inf int, all_inf int, influ_a int, influ_b int, b_victoria int, b_yamagata int);'
- **cursor**.execute(sql)
- **connection**.commit()
- **connection**.close()



2. Extract data from downloaded raw data

- df_idx = pd.DataFrame()
- df_idx = pd.read_csv(fileName, sep=',', skiprows=3)
- df_idx = df_idx.replace(np.nan,0)
-
- s_tmpDate = str(df_idx.SDATE[aa])
- s_parsDate = s_tmpDate.split('-')
- s_month = s_parsDate[1].strip()
- s_day = s_parsDate[2].strip()
-
- country = str(df_idx.Country[aa])
- country = country.replace('₩','₩^')
- country = country.strip()
-
- all_inf = str(df_idx.ALL_INF[aa])
- influ_a = str(df_idx.INF_A[aa])
- influ_b = str(df_idx.INF_B[aa])
-

textbox22		Textbox9					
generated on 20/11/2021 08:44:10 UTC		Data from: All sites					
Country	WHOREGION	FLUREGION	Year	Week	SDATE	EDATE	
Canada	Region of the Americas of WHO	North America	2019	1	2018-12-31	2019-01-06	
Canada	Region of the Americas of WHO	North America	2019	2	2019-01-07	2019-01-13	
Canada	Region of the Americas of WHO	North America	2019	3	2019-01-14	2019-01-20	
Canada	Region of the Americas of WHO	North America	2019	4	2019-01-21	2019-01-27	
Canada	Region of the Americas of WHO	North America	2019	5	2019-01-28	2019-02-03	
Canada	Region of the Americas of WHO	North America	2019	6	2019-02-04	2019-02-10	

INF_A	BYAMAGATA	BVICTORIA	BNOTDETERMINED	INF_B	ALL_INF
3338			31	31	3369
3139			43	43	3182
2299			40	40	2339
2108			48	48	2156



3. Send query to MySQL

- `connection2 = pymysql.connect(host='localhost', user='isahn', password='*E4r5i8u7', db='isahn', charset='utf8', autocommit=True, cursorclass = pymysql.cursors.DictCursor)`
- `cursor2 = connection2.cursor()`
- `sql = "INSERT INTO " + tableName + '(year, week, s_month, s_day, e_month, e_day, flu_region, who_region, country, h1_inf, h3_inf, all_inf, influ_a, influ_b, b_victoria, b_yamagata)' + ' VALUES (' + year + ", " + week + ', ' + s_month + ', ' + s_day + ', ' + e_month + ', ' + e_day + ', ' + ' + region + ' + who_region + ' + country + ' + h1_inf + ', ' + h3_inf + ', ' + all_inf + ', ' + influ_a + ', ' + influ_b + ', ' + b_victoria + ', ' + b_yamagata + ");"`
- `cursor2.execute(sql)`
- `connection2.commit()`
- `connection2.close()`

Basic Analysis





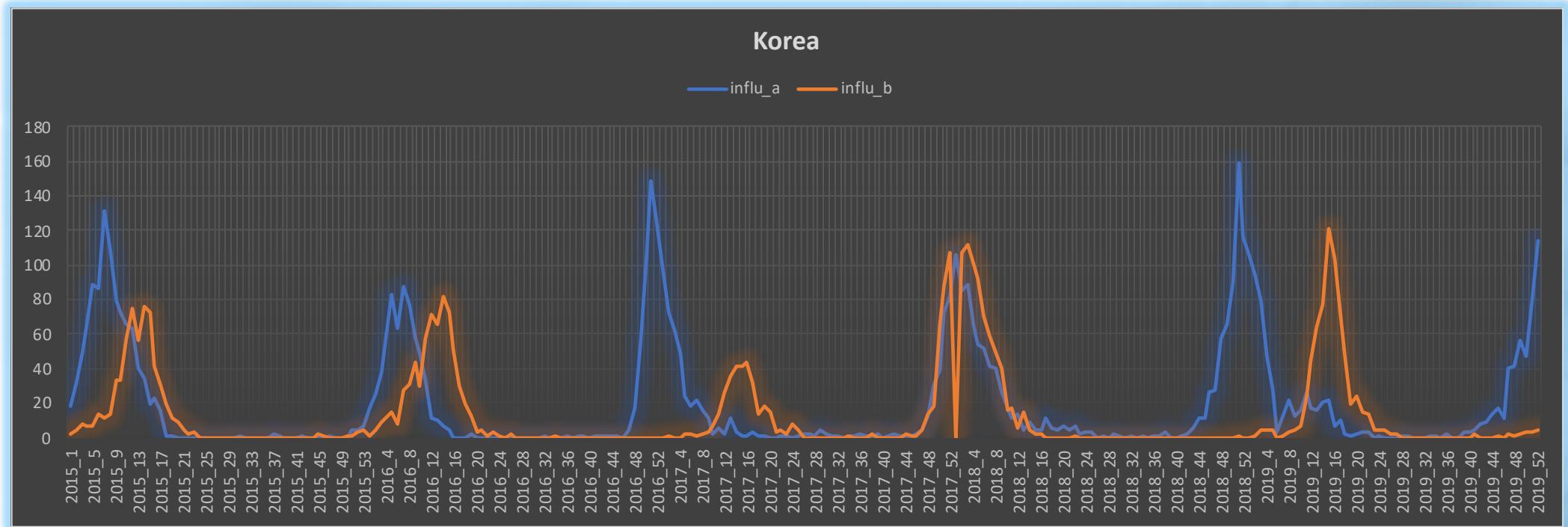
Pre-processed data file

fluNet_Afghanistan_weekly.csv
fluNet_Albania_weekly.csv
fluNet_Anguilla_weekly.csv
fluNet_AntiguaAndBarbuda_weekly.csv
fluNet_Argentina_weekly.csv
fluNet_Armenia_weekly.csv
fluNet_Aruba_weekly.csv
fluNet_Australia_weekly.csv
fluNet_Austria_weekly.csv
fluNet_Azerbaijan_weekly.csv
fluNet_Bahamas_weekly.csv

country	year	week	h1_inf	h3_inf	all_inf	influ_a	influ_b	b_victoria	b_yamagata
Republic of Korea	2015	1	0	15	20	18	2	0	0
Republic of Korea	2015	2	0	23	37	32	5	0	0
Republic of Korea	2015	3	0	43	57	49	8	0	0
Republic of Korea	2015	4	0	52	71	64	7	0	0
Republic of Korea	2015	5	0	82	96	89	7	0	0
Republic of Korea	2015	6	0	71	101	87	14	0	0
Republic of Korea	2015	7	0	116	144	132	12	0	0
Republic of Korea	2015	8	0	88	122	108	14	0	0
Republic of Korea	2015	9	0	68	113	80	33	0	0
Republic of Korea	2015	10	0	63	106	73	33	0	0
Republic of Korea	2015	11	0	59	124	66	58	0	0
Republic of Korea	2015	12	0	55	138	63	75	0	0
Republic of Korea	2015	13	0	29	97	40	57	0	0
Republic of Korea	2015	14	0	26	110	34	76	0	0
Republic of Korea	2015	15	0	13	93	20	73	0	0

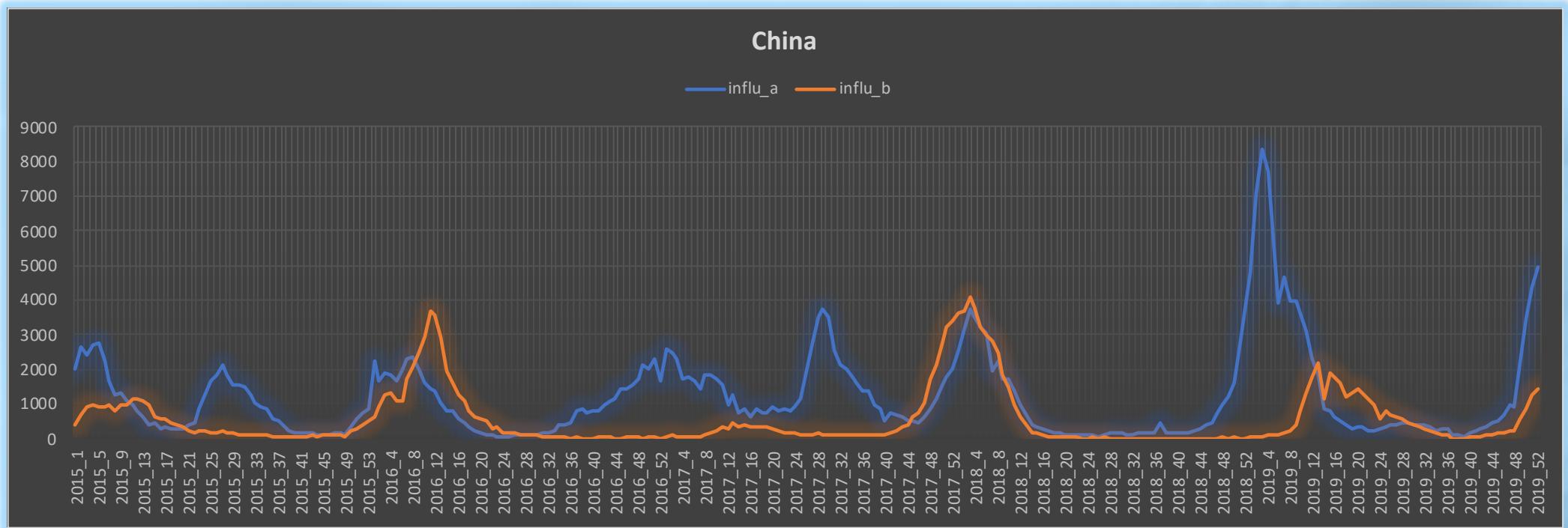


[Korea] Influenza A & B



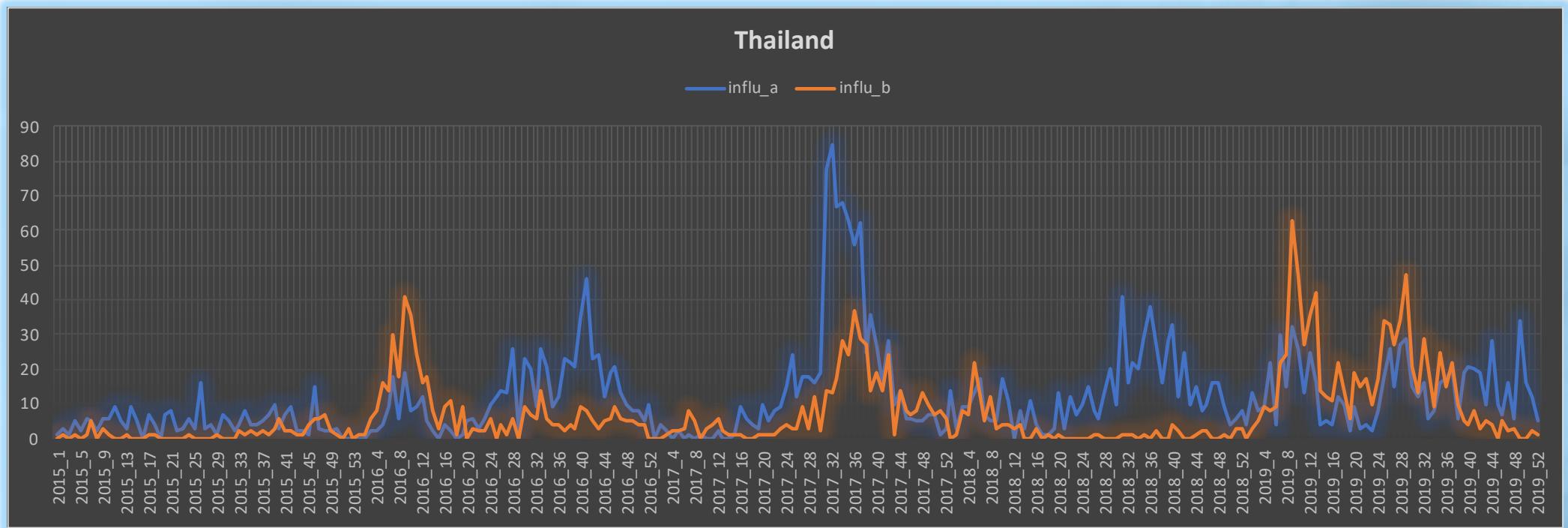


[China] Influenza A & B



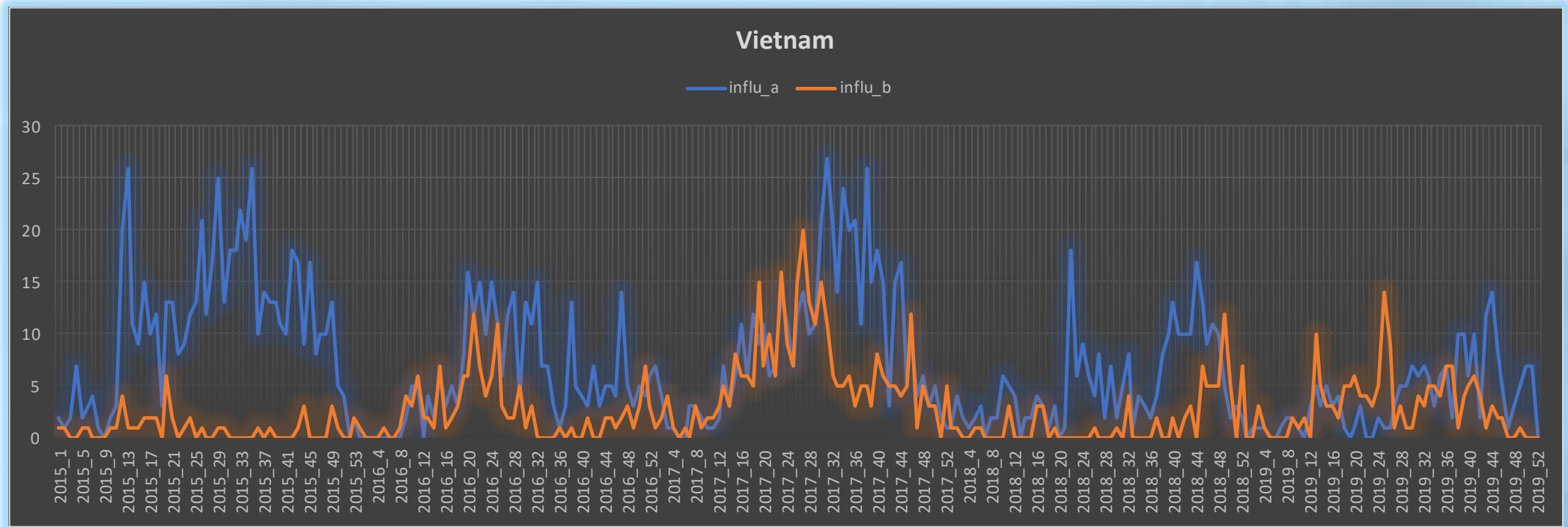


[Thailand] Influenza A & B



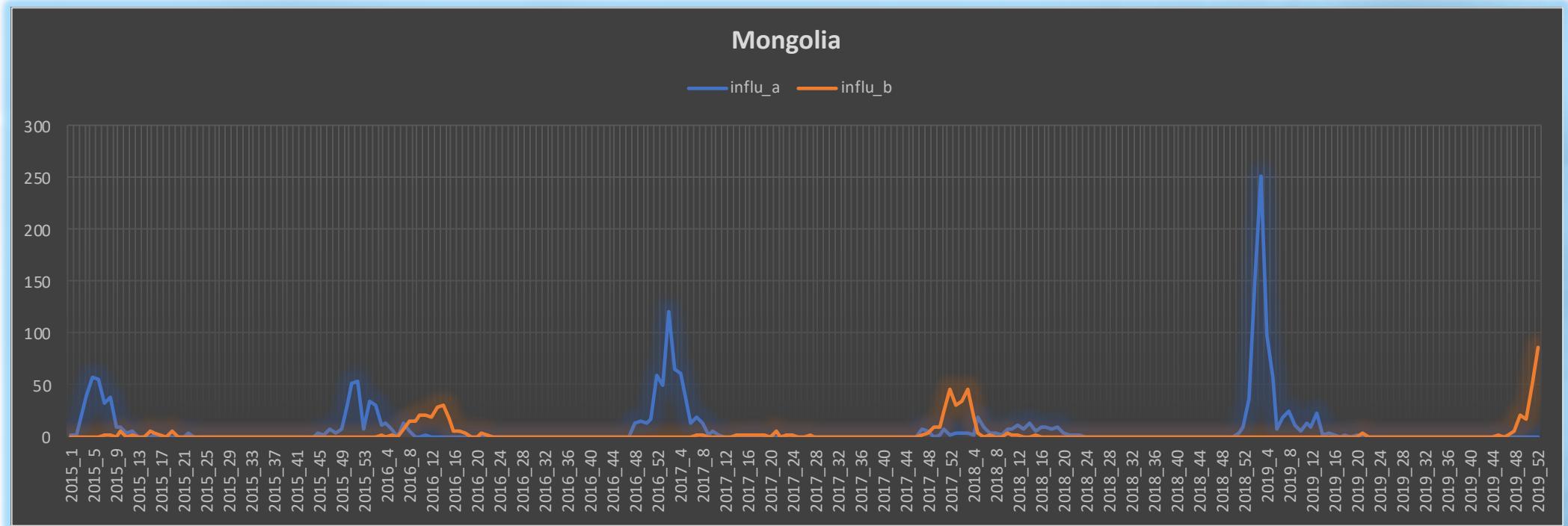


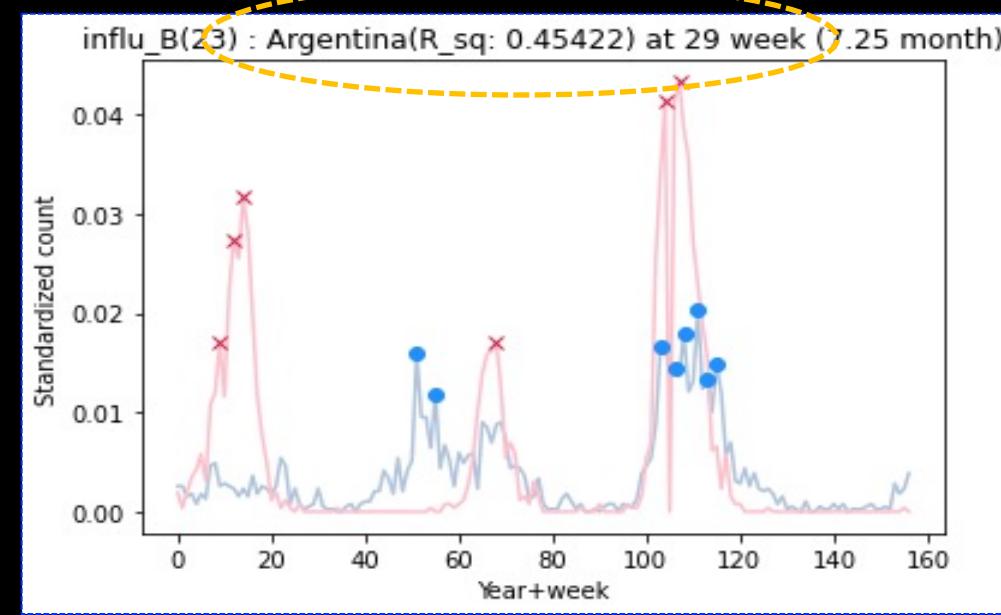
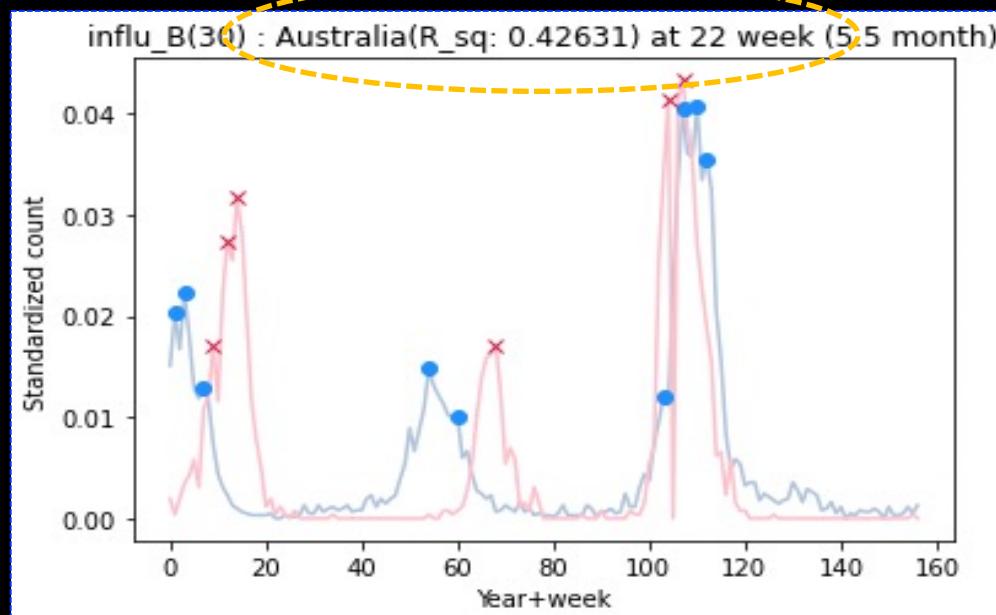
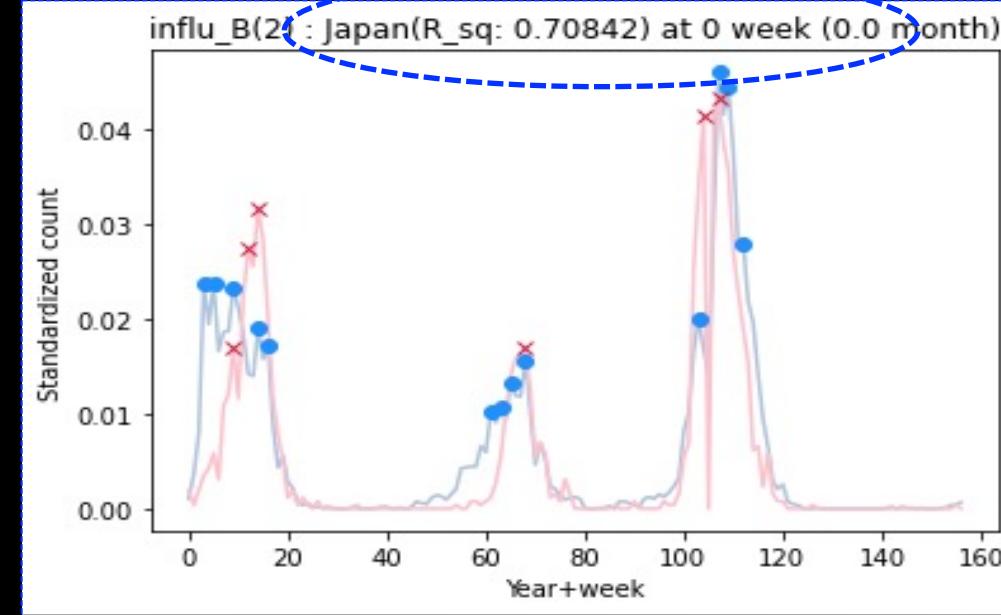
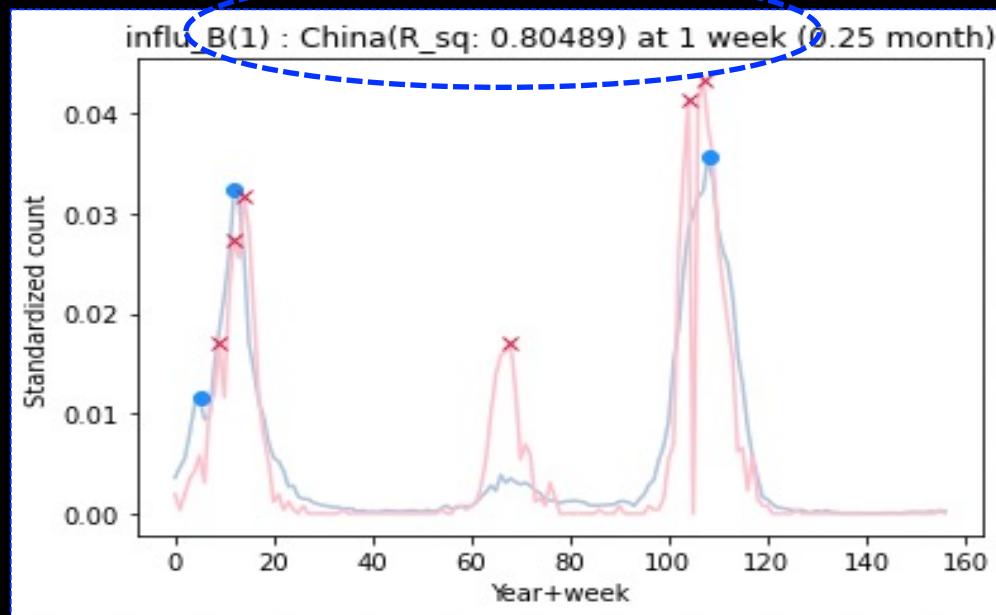
[Vietnam] Influenza A & B





[Mongolia] Influenza A & B







Advanced analysis

Long-term prediction of influenza virus



[1] Data preparation

- Dataset
 - ILI data were collected from the Korea Centers for Disease Control and Prevention (**KCDC**).
 - **FluNet's Surveillance data** from the 160 countries from the 40th week of 2010 until the 52nd week of 2019.
 - **Google Trends:** We included the search keywords “*A hyeong dokgam*” and “*B hyeong dokgam*”, which are Korean words for influenza A virus and influenza B virus, respectively, in South Korea from October 2010 to December 2019,



● A형 독감
검색어

+ 비교

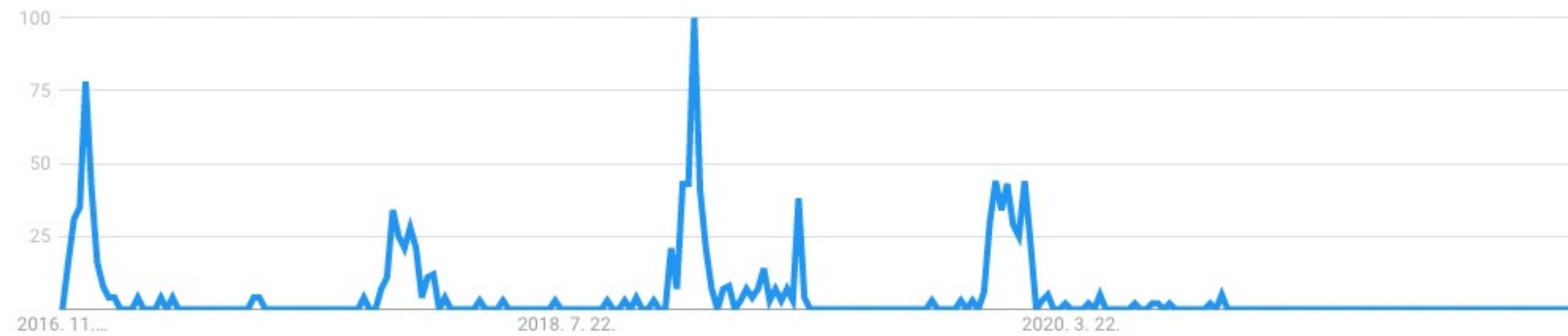
대한민국 ▾

지난 5년 ▾

모든 카테고리 ▾

웹 검색 ▾

시간 흐름에 따른 관심도 변화 ?



● B형 독감
검색어

+ 비교

대한민국 ▾

지난 5년 ▾

모든 카테고리 ▾

웹 검색 ▾

시간 흐름에 따른 관심도 변화 [?](#)

[⬇](#) [↔](#) [◀](#)





[1] Data preparation

- Dataset
 - **Weather data** in South Korea were obtained from the National Weather Data Release Portal.
 - Weekly temperature data and the average values for Seoul in South Korea during each week



[2] Statistical analysis

- **Cross-correlations** were analyzed using Pearson's correlation, with a time lag range of ± 30 weeks from the 40th week of 2010 until the 52nd week of 2019.
- **Linear regression analyses** were used to evaluate the relationship between the ILI in South Korea and ILI in selected variables by cross-correlation analysis with a time lag from the 40th week of 2010 to the 52nd week of 2019.



[3] Prediction model

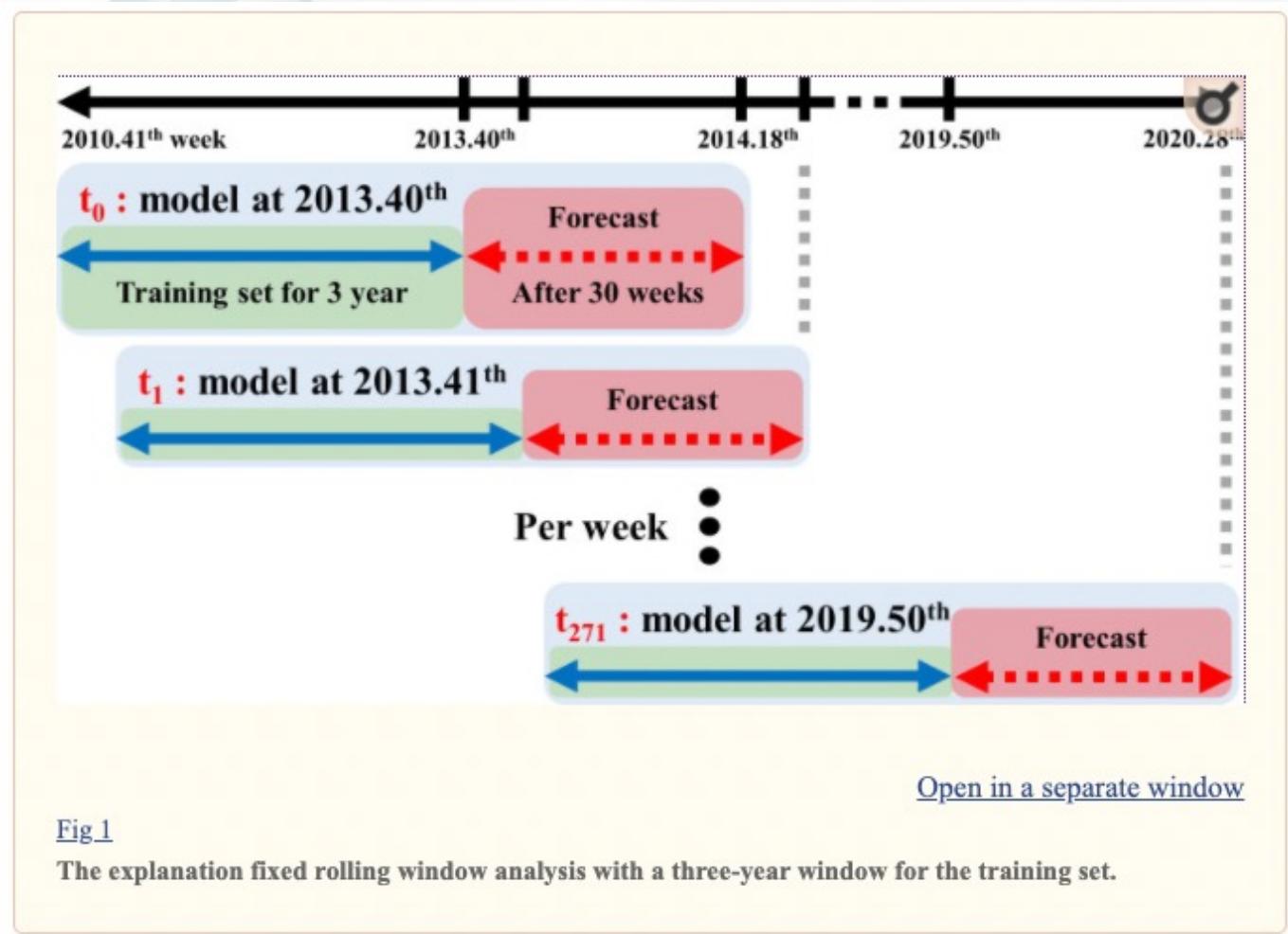


Fig 1

The explanation fixed rolling window analysis with a three-year window for the training set.

- The input variables of a prediction model for ILI after 30 weeks in South Korea were 1) previous seasonal ILI, 2) average vapor pressure, and 3) total INF in Argentina.



[4] Cross-correlation analysis

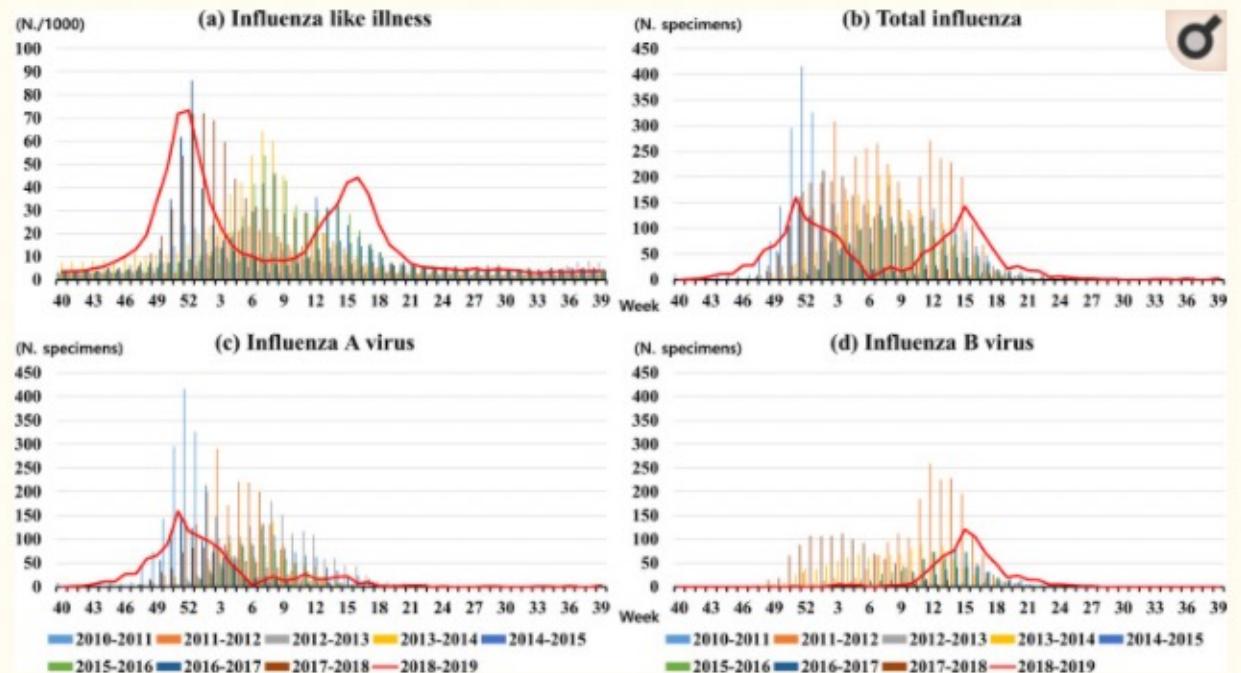


Fig 2

The nine seasonal patterns of ILI (a), total INF (b), INF A (c), and INF B (d) activity from 2010 to 2019 in South Korea. The red line denotes influenza cases during the 2018–2019 influenza season, and the bar graphs are for the rest of the seasons. INF, Influenza; ILI, Influenza-like illness.

- This figure shows the nine seasonal patterns of ILI, total INF, INF A, and INF B activity from 2010 to 2019 in South Korea.



Assignments





Assignments

1. **Collect** influenza outbreak data of countries of interest from **FluNet** and save it as a database.
2. **Compare** the differences between Influenza seasons.
3. **Search** the Time series trends in **Google Trends** using a word related in influenza.
4. **Elijah Hwejin Lee** from KIST and **Ahmad Masudi** from KIST
Prepare a 5-minute presentation of what these people have done.

THANK YOU

