

# A machine learning approach to predict air passenger traffic affected by COVID-19



Raffaele Vergnani  
Ironhack Data Analytics  
final project

Friday 12th March 2021





**2020 has seen the  
sharpest traffic  
decline in aviation  
history.**





Passenger traffic was  
-66% at global level,  
compared to 2019.  
(source: IATA)



# How it started

Luckily, my flight  
to Berlin on  
11/03/2020  
wasn't cancelled



Temperature  
checks and  
longer  
procedures at  
borders controls

PARTENZE / DEPARTURES 12:02							
Airline	Flight	Destination	Time	Exp.	Check-in	Gate	Note
AZ	7168	MOSCOW SVO	11:45	C 54 - 57	24		Decollato
AE	683	ATHENS	12:15	B 25 - 25			Cancelled
AA	8550	MADRID	12:25	B 33 - 33			Cancelled
RY	6423	VALENCIA	12:35	A 01 - 08			Cancelled
AA	6607	LONDON LHR	12:40	B 30 - 31			Cancelled
RY	3407	LUTON	12:40	A 01 - 08	19		
RY	4895	LANZAROTE	12:40	A 01 - 08			Cancelled
RY	1085	CATANIA	13:00	A 01 - 08	6		
RY	4303	PARIS BVA	13:00	A 01 - 08	12		
EW	8241	MUNICH	13:10	B 38 - 39	1		
DL	9419	AMSTERDAM	13:10	C 43 - 48	3		
RY	137	BERLIN SXF	13:40	A 01 - 08	14		
easyJet	EZY990	LONDON LGW	13:50	B 25 - 25			Cancelled
QF	8094	DUBAI	14:30	A 13 - 21	17-18		
AC	9115	FRANKFURT	14:30	B			Cancelled
W6	3164	BUCHAREST OTP	14:30	C			Cancelled
RY	9367	BARCELONA	14:35	A			Cancelled
RY	5717	WARSAW WMI	14:40	A			
RY	7794	MALTA	15:30	A			Cancelled
RY	5810	PORTO	16:00	A			Cancelled
AZ	7770	LISBON	16:15	A			Cancelled
RY	4301	TRAPANI	16:25	A			
RY	4317	WROCLAW	16:25	A			
W6	3790	CHISINAU	16:50	C			Cancelled
Q8	3647	BARCELONA	16:55	C			Cancelled
RY	5375	MADRID	17:00	A			Cancelled
SN	3126	BRUSSELS	17:00	C			
W6	3560	TIMISOARA	17:15	C			Cancelled
AZ	7144	CASABLANCA	17:20	B			Cancelled
EW	8243	MUNICH	17:30	B			
AF	1773	LYON	17:50	C			
MK	9471	PARIS CDG	18:15	C			
BA	545	LONDON LHR	18:35	A			Cancelled

N: FLIGHT DEPARTURE ANNOUNCEMENT  
Aeroporto di Bologna





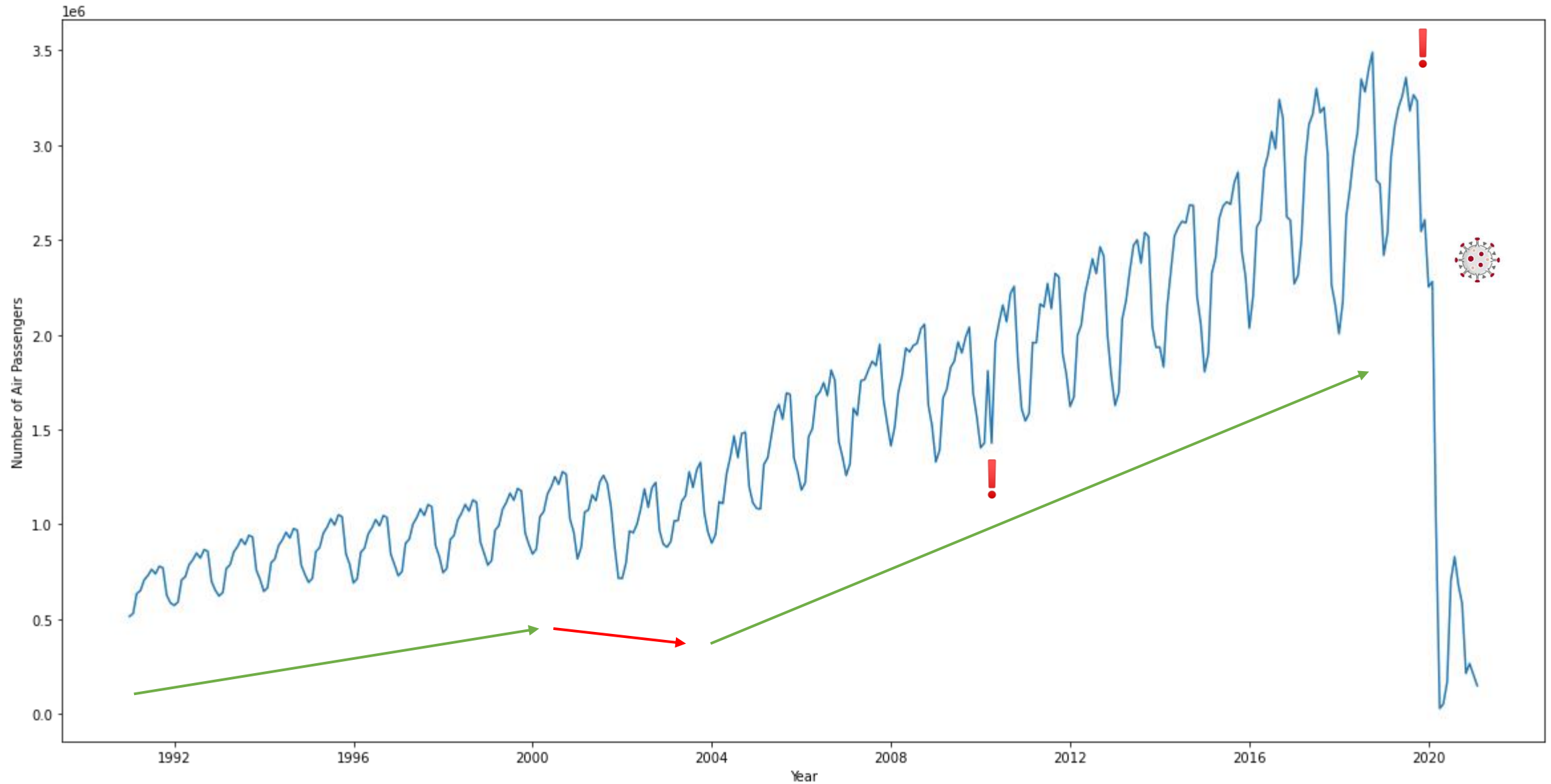
# How it is going on

Few flights operating in autumn and winter, due to COVID-19  
2nd wave and cross-border travel restrictions

Vaccinations are spreading, spring and summer are around the corner. As a result, higher passenger volumes should be (somehow) back.

MILAN BERGAMO AIRPORT				PARTENZE	
COMPAGNIA VOLO	PER	SCHED	STIM	BANCO	NOTE
RYANAIR FR 4733	BERLIN	16:10		SELF	✓ MALTA AIR
RYANAIR FR 7748	CAGLIARI	16:20		SELF	✓ MALTA AIR
RYANAIR FR 6451	PALERMO	17:15		SELF	✓ MALTA AIR
RYANAIR FR 6651	CATANIA	18:00		SELF	✓ MALTA AIR
RYANAIR FR 5984	MADRID	18:00		SELF	
Alitalia 30 458	CASABLANCA	19:15		31-34	
RYANAIR FR 8095	BRINDISI	19:55		SELF	✓ MALTA AIR
RYANAIR FR 4525	ARLEROI	21:10		SELF	
RYANAIR FR 4118	NAPOLI	21:15		SELF	✓ MALTA AIR
Wizz Air W6 3136	OPENI	21:50		13-14	
RYANAIR FR 2992	COPENHAGEN	06:25		SELF	
RYANAIR FR 8846	MARRAKECH	06:25		SELF	
RYANAIR FR 4111	NAPOLI	07:00		SELF	
RYANAIR FR 6354	PRAT	08:35		SELF	
SpiceJet SG 9724		K09:45			
RYANAIR FR 8838	TENERIFE	11:50		SELF	

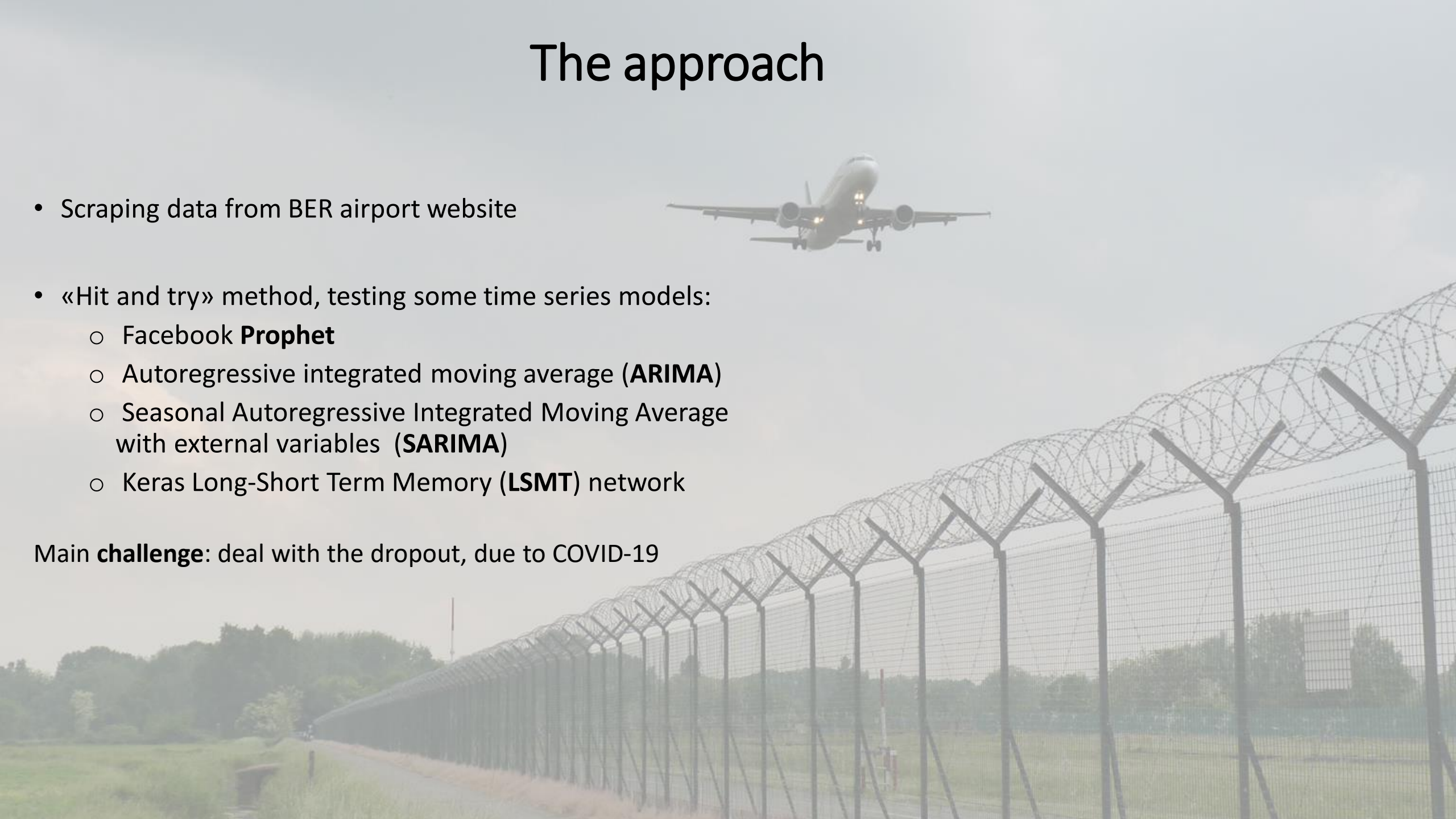
# Total number of passengers for Berlin's airports (Jan 1991- Feb 2021)

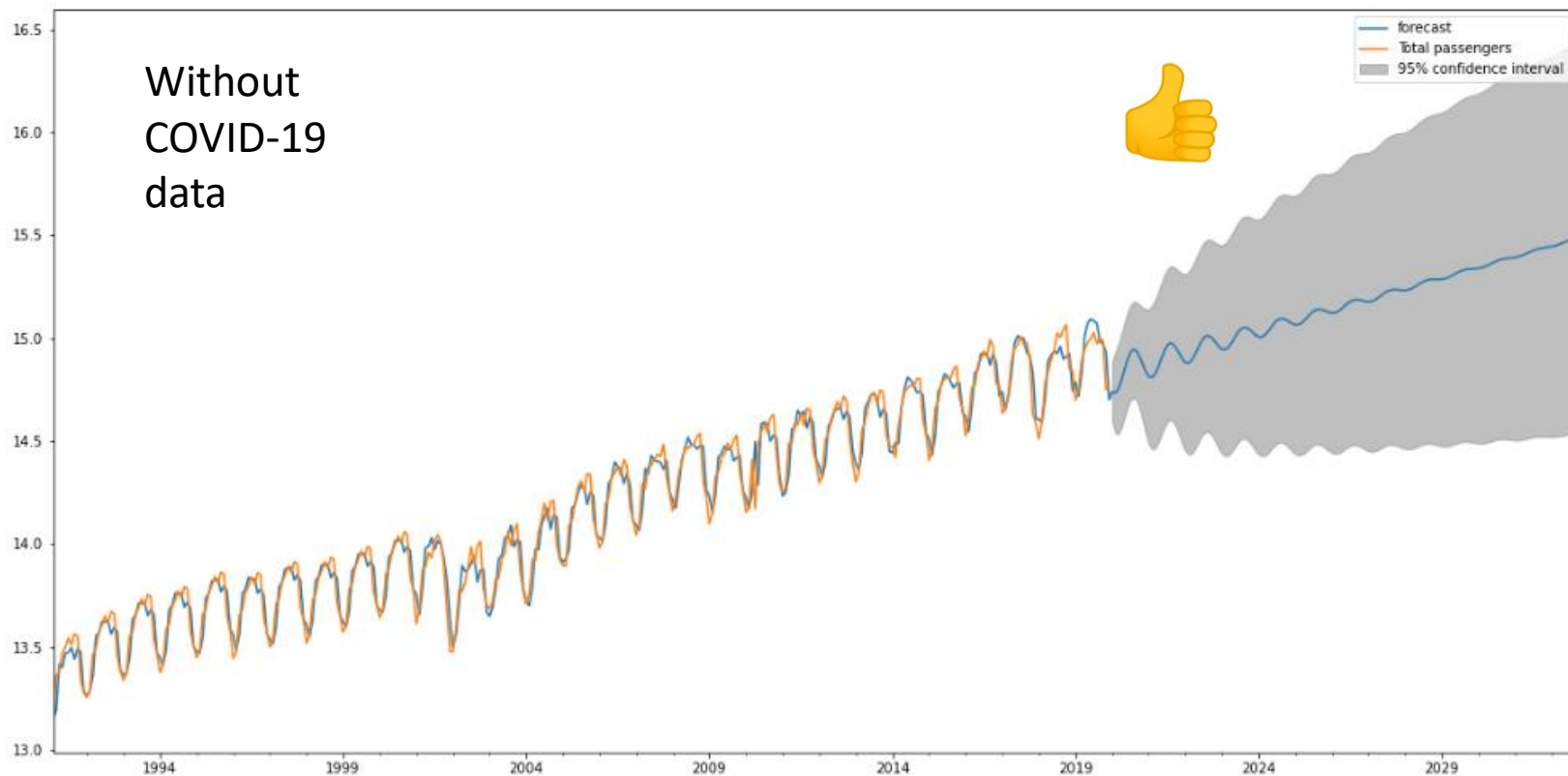


# The approach

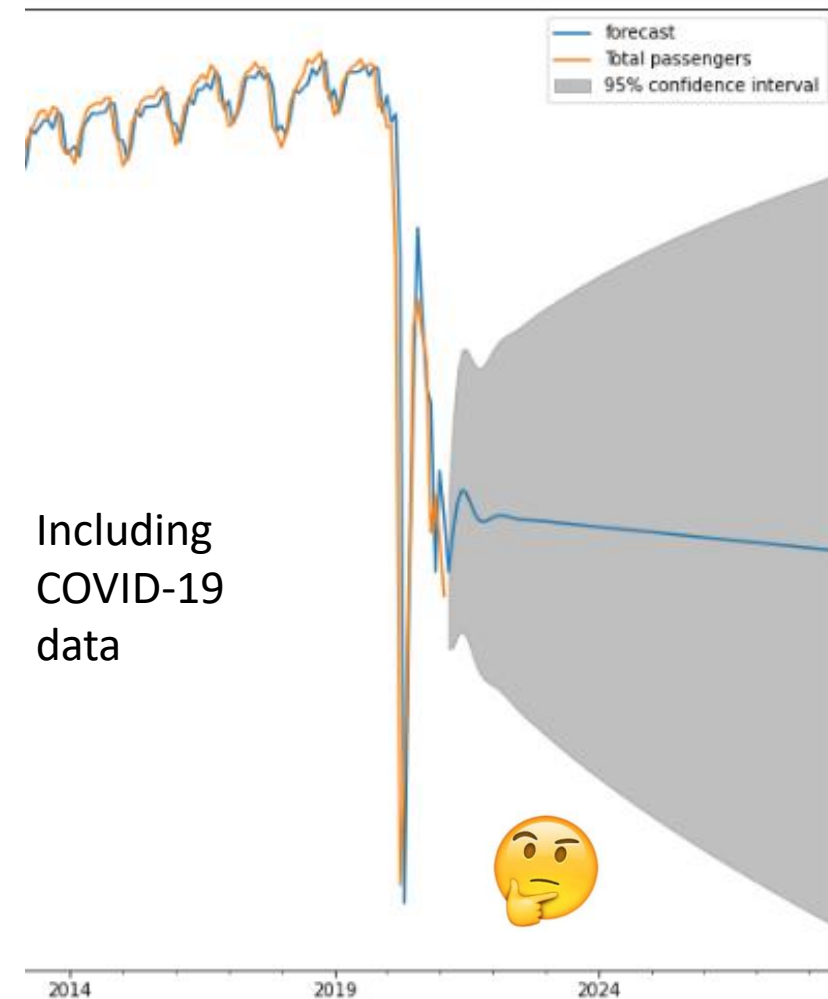
- Scraping data from BER airport website
- «Hit and try» method, testing some time series models:
  - Facebook **Prophet**
  - Autoregressive integrated moving average (**ARIMA**)
  - Seasonal Autoregressive Integrated Moving Average with external variables (**SARIMA**)
  - Keras Long-Short Term Memory (**LSMT**) network

Main **challenge**: deal with the dropout, due to COVID-19



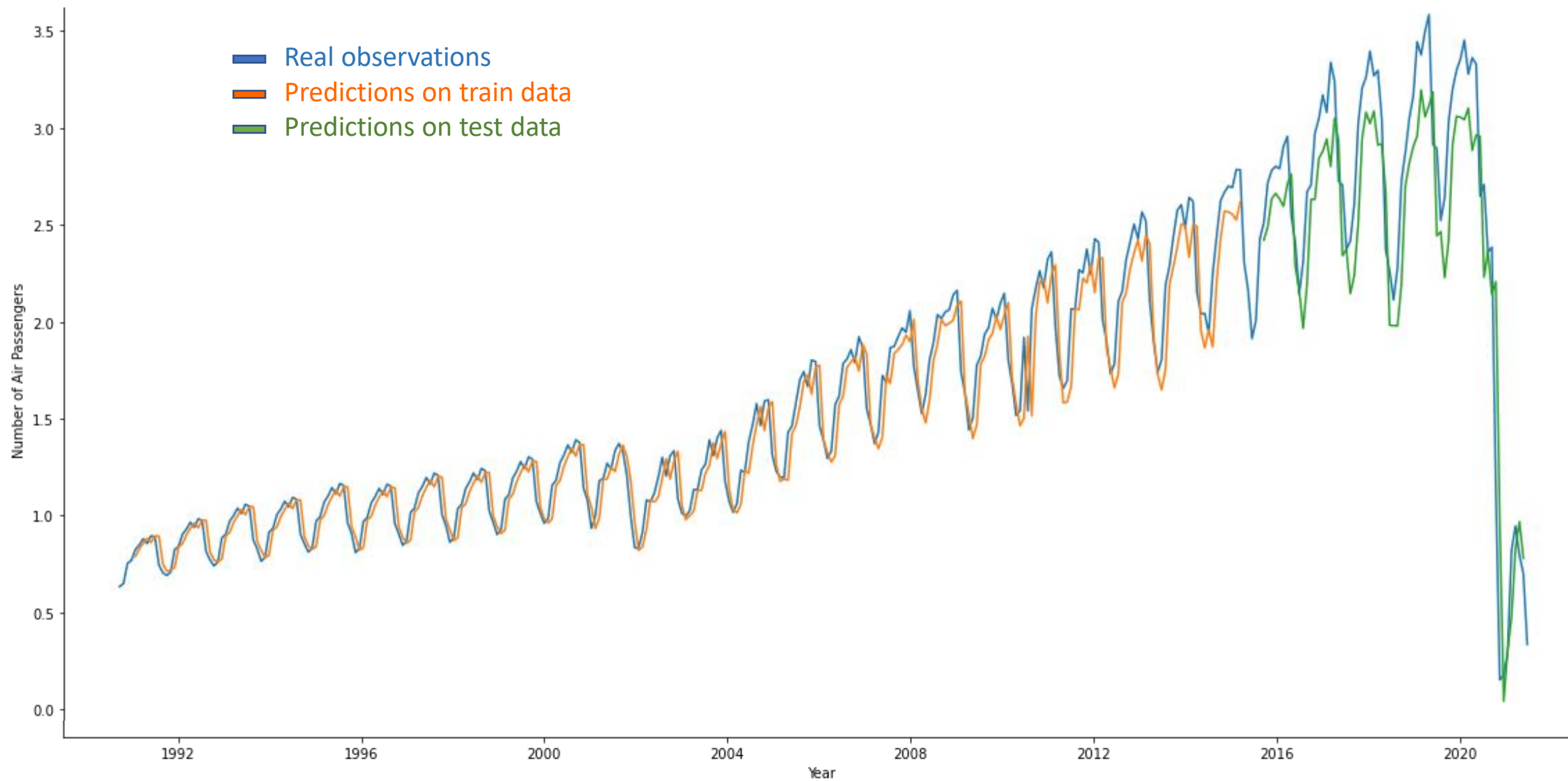


**Attempt with ARIMA model**





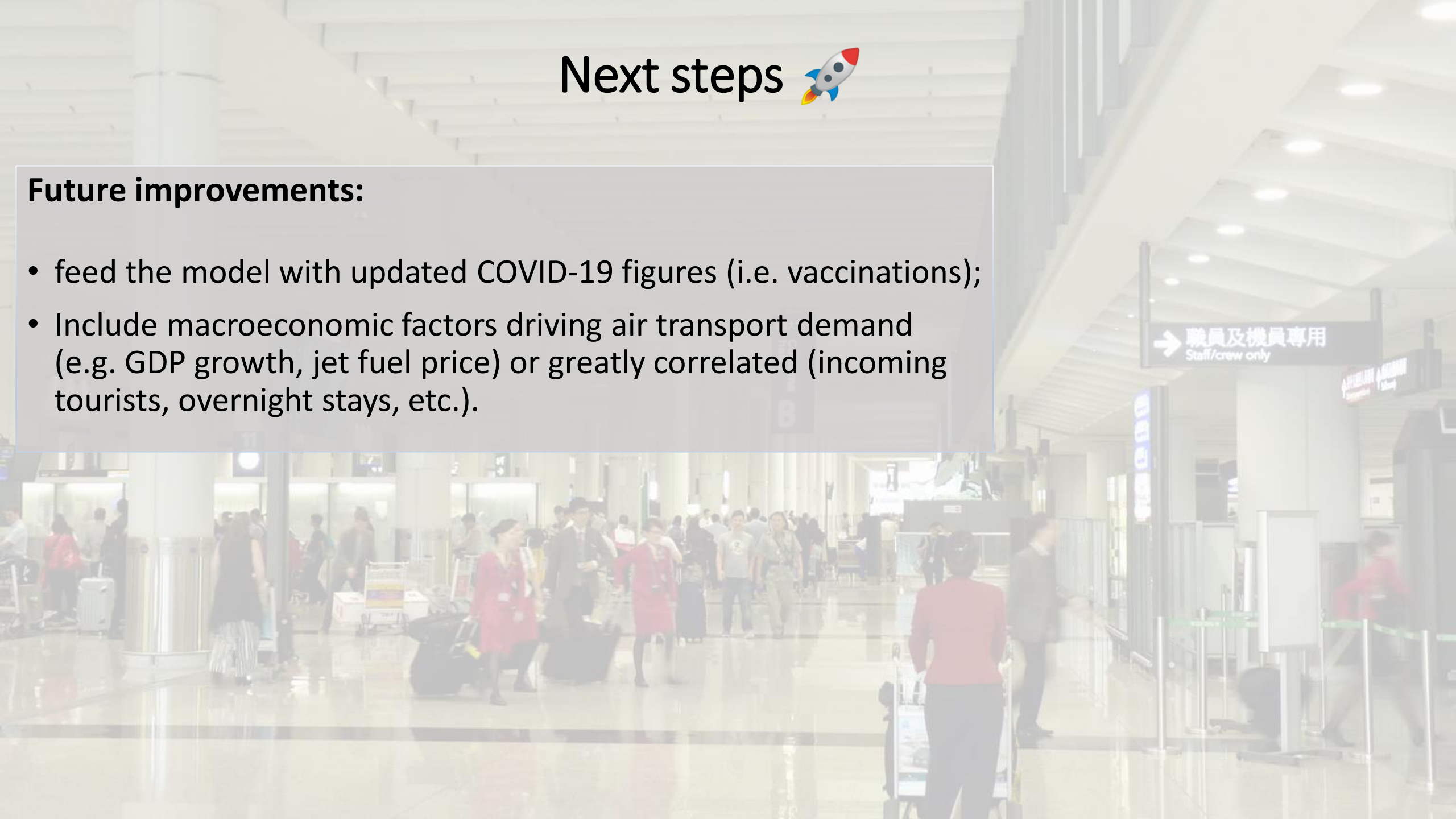
# Output Keras LSTM model



# Next steps

## Future improvements:

- feed the model with updated COVID-19 figures (i.e. vaccinations);
- Include macroeconomic factors driving air transport demand (e.g. GDP growth, jet fuel price) or greatly correlated (incoming tourists, overnight stays, etc.).





# Thank you for your attention!

**Contacts:**

vergnanir@gmail.com

**LinkedIn:**

Raffaele Vergnani



And hopefully  
flying again  
anytime soon!

Photos and plots: R. Vergnani