Pacific Northwest
NATIONAL LABORATORY

PNNL-30778

EBSD seed LDRD Project: Does Corona Virus – 2019 (COVID-19) and Seasonal Flu have similar meteorology and air quality controls driving their spread?

September 2020

Aishwarya Raman
Susannah M Burrows
Ethan King
Lisa M Bramer

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical
Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

EBSD seed LDRD Project: Does Corona Virus – 2019 (COVID-19) and Seasonal Flu have similar meteorology and air quality controls driving their spread?

September 2020

Aishwarya Raman
Susannah M Burrows
Ethan King
Lisa M Bramer

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Report

Project summary: Seasonal influenza and Influenza like Illnesses (ILI) pose a serious public health risk and in turn affect the economy. Various factors affect ILI cases and mortality, including the pathogen and its interaction with the host, as well as environmental and socioeconomic factors such as meteorology, household structure, air pollution, urbanization, and population. Despite the growing number of studies on influenza and ILI, challenges remain in forecasting the timing of seasonal onset, outbreak patterns, and key factors affecting transmission. In particular, the impacts of meteorology and air quality on ILI have been challenging to understand, with linear regression studies focused on different geographic regions producing contradictory results. For example, influenza seasonality has been associated with cold-dry conditions in temperate mid-latitudes, but with humid-rainy conditions in tropical climates (Tamerius et al., 2013). These apparently contradictory results imply that the relationships between influenza cases and atmospheric variables may be too complex to be captured by linear regression models. In this seed project, we analyze meteorology and air quality variables from numerical models to determine which atmospheric variables are most helpful in predicting weekly changes in recorded flu cases. In contrast to most previous studies that relied on linear regression analysis to predict the timing of the flu onset or peak, we employ a robust machine learning algorithm to evaluate the contribution of atmospheric variables to weekly changes in recorded ILI cases. These results may also be relevant to the spread of other respiratory illnesses such as Corona Virus Disease – 2019 (COVID-19).

Introduction and Project background: Previous studies have determined that the meteorological variables associated with the timing of ILI cases differ between geographical regions. Most of these studies have identified correlations of influenza patterns with key climate variables like temperature, humidity, and precipitation (Liu et al., 2019). Additionally, previous exposure to air pollutants, particularly fine particulate matter ($PM_{2.5}$) and ozone, has been shown to be associated with greater susceptibility and poorer outcomes for respiratory illness (Chen et al., 2017). However, statistical correlations between air pollutants and influenza mortality are not well-understood. This project aimed to improve that understanding by developing a machine learning approach that employs Random Forest regression models to identify meteorological and air quality features that demonstrate significant impact on seasonal influenza outbreaks in the contiguous United States (CONUS). We also analyze the key meteorological and air quality variables that drive ILI percentage counts in separate National Climate Assessment (NCA) regions. In recent decades, machine learning approaches have been increasingly used in the study of infectious diseases to discern the key factors driving their patterns. For example, a Random Forest (RF) regression model was found to perform better than conventional moving average methods in identifying factors associated with H1B1pdm influenza infections (Petukhova et al., 2018).

Data and Approach: We compare the performance of three models for one-week-ahead prediction of ILI data to asses the impacts of meteorological features and air quality variables on influenza infection. The models contained information about the infection history and either no meteorological data (RFr), one week of meteorological data (RFM), or meteorological data including lags up to four weeks (RFMd). Influenza case numbers were obtained from the CDC; meteorological and air quality data were obtained from the NASA MERRA-2 reanalysis product.

Results and Accomplishments: Inclusion of meterological and air quality data in the RF model improved the one-week-ahead prediction of ILI. The four unique meterological features that primarily contributed to improvement in predictability of ILI cases were identified for each of the models. For the RFM, we find that total column Ozone (O_3), Surface Pressure (PS), Surface sulfate mass concentrations (SO_4), and total precipitable water vapor (TQV) are the key variables driving percentage of ILI counts across CONUS (hereafter, RFM with SM). For the RFMd, in addition to O_3, TQV, and SO_4, an unexpected result was the finding that the quasi-geostrophic vertical velocity at the 500mb pressure level (OMEGA_500) played a key role in prediction of ILI. Box plots in Figure 1 shows the improvement in the RF model on adding meteorological variables compared to the control experiment. This is suggestive of a possible mediating role of synoptic-scale air stagnation events in contributing to air pollution effects on respiratory illness outcomes, although further work would be required to validate this hypothesis. Since aILI is tied to severity of symptoms it is difficult to discern whether the observed effects are the result of atmospheric impacts on airborne transmission (e.g., due to longer residence time of viral particulates in dry air), impacts on host susceptibility to infection, or impacts on the severity of symptoms among individuals previously exposed to pollutants.

Impacts/Benefits: Our results suggest that population-scale respiratory illness transmission and/or outcomes are tied to atmospheric variables, and therefore may be sensitive to changes in future climate. In particular, air stagnation events are projected to occur more fequently in many regions under future climate conditions, leading to more frequent exceedences of clean air thresholds for ozone and PM2.5. **Our results suggest that projected future increases in air stagnation and poor air quality have the potential to cause future flu seasons to be more severe than otherwise expected.** These findings may also prove relevant to the transmission of other infectious diseases, including COVID-19 and future emerging respiratory illnesses. **A manuscript describing these results is in preparation for submission to the journal *Geohealth*.**

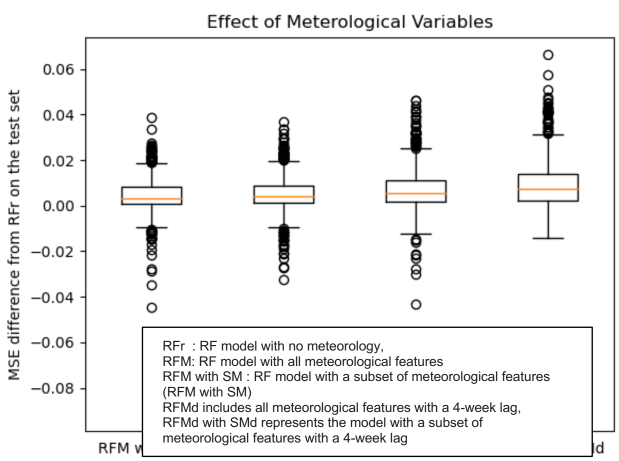

Figure 1. One-week-ahead flu forecast is improved by inclusion of meteorological and air quality variables in the random forest (RF) model. Box plots show the improvement in mean squared error (MSE) differences between the control RF model and experiment.

References

- Tamerius JD, Shaman J, Alonso WJ, Bloom-Feshbach K, Uejio CK, Comrie A, et al. (2013) Environmental Predictors of Seasonal Influenza Epidemics across Temperate and Tropical Climates. PLoS Pathog 9(3): e1003194.
- Liu, Z., Zhang, J., Zhang, Y., Lao, J., Liu, Y., Wang, H., & Jiang, B. (2019). Effects and interaction of meteorological factors on influenza: Based on the surveillance data in Shaoyang, China. *Environmental research*, *172*, 326-332.
- Chen, G., Zhang, W., Li, S., Zhang, Y., Williams, G., Huxley, R., ... & Guo, Y. (2017). The impact of ambient fine particles on influenza transmission and the modification effects of temperature in China: A multi-city study. *Environment international*, *98*, 82-88.
- Petukhova, T., Ojkic, D., McEwen, B., Deardon, R., & Poljak, Z. (2018). Assessment of autoregressive integrated moving average (ARIMA), generalized linear autoregressive moving average (GLARMA), and random forest (RF) time series regression models for predicting influenza A virus frequency in swine in Ontario, Canada. *PloS one*, *13*(6), e0198313.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov