

# Breve report sui risultati ottenuti tesi

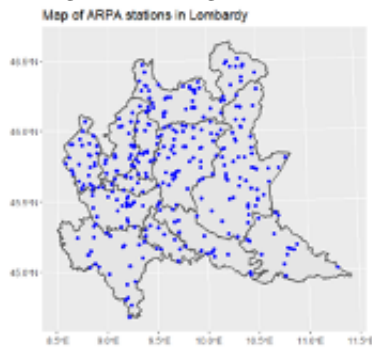
David Guzman Piedrahita and Marco Vinciguerra

17 febbraio 2022

## 1 Introduzione e reti neurali utilizzate

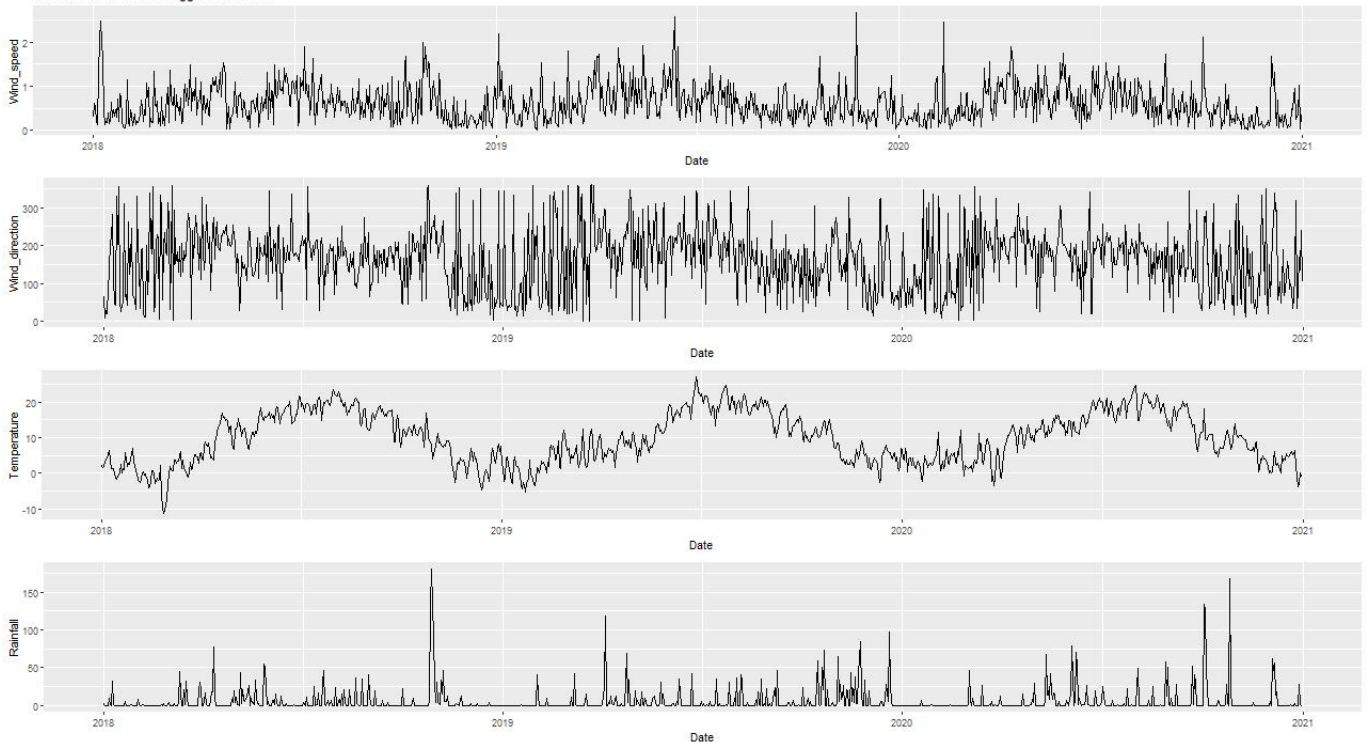
Per studiare la potenziale relazione tra ammoniaca e particolato in Lombardia, sono state utilizzate diverse strategie di machine learning al fine di trovare un modello in grado di descrivere nel modo più accurato possibile, il comportamento e le interazioni tra le due sostanze in questione.

In particolare, sono state utilizzate diverse tipologie di reti neurali LSTM (long-short term memory), strumenti adatti a modellizzare serie storiche i cui valori dipendono dal passato utilizzando le informazioni degli istanti temporali precedenti. I diversi modelli implementati permettono di gestire dati mancanti nel dataset, e sono in grado di migliorare il rendimento dei modelli, stabilizzandolo e rimuovendo le informazioni superflue.



Il problema dei dati mancanti in questo tipo di analisi riveste un ruolo importante in quanto i sensori di rilevamento degli inquinanti vengono spesso disattivati causando la perdita dei dati. In questo caso, come si può vedere dalla figura seguente, i dati mancanti sono assenti

Cassina Valsassina Moggio 2018-2020



## 2 Caso studio e dati

Le 3 centraline prese in considerazione sono state Moggio (ID=681, rurale), Cremona via Fatebenefratelli (ID=627, urbana) e Schivenoglia (ID=703, rurale). Per il caso studio sono stati utilizzati dati provenienti dalle 3 centraline di qualità dell'aria e da centraline meteorologiche dislocate in tutta la Lombardia.

Usando, per esempio, i dati della centralina di Moggio dal 2014 al 2020, possiamo ottenere risultati che sottolineano una forte relazione tra il particolato e l'ammoniac, tenendo conto anche di altre variabili atmosferiche, quali: la velocità e la direzione del vento, l'angolazione (tramite il quadrante), la temperatura e le precipitazioni.

## 3 Risultati delle reti neurali

Il dataset a disposizione è stato diviso in training e validazione (scrivere la percentuale). Infatti, usando per la creazione del modello dati dal 2014 al 2019 e per la validazione tutto il 2020. Otteniamo un notevole inseguimento della previsione rispetto ai dati reali, usando dati di 5 giorni successivi per determinare l'andamento del sesto giorno.

## 4 Regressione tramite reti neurali

La figura qui sotto rappresenta in arancione il modello da noi ottenuto, mentre quello in blu il modello reale.

