





**Figure S2.** Correlation matrix of some variables from CHIMERE simulation output for a typical day (15 October 2021). The meaning of the variable : lon:Longitude, usta:Friction velocity. Lat:Latitude, wsta:Convection velocity, relh:Relative humidity, soim:Soil moisture, temp:Temperature, sreht:Surface relative humidity, winz:Zonal wind, pres:Pressure, winm:Meridional wind, sshf:Surface sensible heat flux, winw:Vertical wind, slhf:Surface latent heat flux, sphu:Specific humidity, copc:Convective precipitation, pNH3:Ammonium particles, lspc:Large scale precipitation, pHNO3:Nitrates particles, topc:Precipitation, TOL:Toluene, aerr:Aerodynamic resistance, albedo:Surface albedo, atte:Radiation attenuation factor, hght:Boundary layer height, swrd:Shortwave radiation flux.

**Table S1.** The network architecture search space.

Hyperparameter	Minimum Value	Maximum Value	Increment
Number of layers	10	50	10
Number of perceptrons	10	50	10
Batch size	5000	20000	1000

**Table S2.** The best architecture for the NN regressor.

Layer Type	Output Shape	Number of Parameters
Batch normalization	96	384
Dense	30	2910
Dense	20	620
Dense	10	210
Dense	1	11