**3. Results**

**3.1 Database of physicochemical and optical properties of black carbon fractal aggregates: dataset statistics**

- show the spread of the MSTM data with respect to the various features

* 1. **Machine learning techniques for black carbon optical properties at various stages of ageing**
* For random split method
* Wavelength fixed to 660 nm
* 3 plots for 3 optical properties (qabs,qsca,g)
* each plot has 9 subplots of particles
* each sub plots have type line (optical property vs size)
* each sub-plot : three lines for KRR, NN, and MSTM
  1. **Performance of the machine learning algorithms**
* For random split method
* Wavelength fixed to 660 nm
* 1 plot showing for 3 optical properties in three rows (qabs,qsca,g)
* 3 columns for 3 cases of particles (df=1.5, fcoating=0%; df=2.1,fcoating=50%; df=2.7,fcoating=90%)
* each sub plots have type point (KRR/NN vs MSTM)
* separate color points for KRR/NN
  1. **Technical things about the machine learning algorithms**
* If any technical plots – errors, time, etc
  1. **Generalization performance**

**–** Same as 3.3 for extrapolation

**–** Same as 3.3 for interpolation

* 1. **Error analysis of the machine learning algorithms**
* box plots done by Jaikrishna – need to be restructured
* three plots each for df/fcoating/wavelength
* each plot has 3 optical properties in three rows (qabs,qsca,g)
* each plot has 3 rows for random split, interpolation, extrapolation
* Each subplot boxes of error vs df/fcoating/wavelength
* X-axis df/fcoating maybe divided into 3-4 groups (for e.g. 1.5-1.8; 1.9-2.2; 2.2-2.5; 2.6-2.9)
* for each df/fcoating/wavelength point: two boxplots for KRR and NN (different color)
  1. **Guidelines to operate the machine learning algorithm**
* Text only