Light scattering enhancement factor, discrepancies between model and observations

January 9, 2020

Light scattering enhancement factor: Discrepancies between model and observations
Dominic Heslin-Rees, NeGI course 2019, Abisko.

1 Introduction

Aerosol particles have demonstrated a clear ability to affect the Earth's climate. The latest report by the Inter-governmental Panel on Climate Change (IPCC) stated that clouds and aerosols continue to contribute the largest uncertainty to our understanding of changes to the Earth's energy budget (IPCC, 2013). Aerosols can influence the Earth's climate directly through what is known as the radiative effects of aerosol-radiation interactions. The direct effects are produced as a result of a combination of the scattering (σ_{sp}) and absorbing (σ_{ap}) properties of aerosol particles. Furthermore, aerosol particles can also perturb the Earth's energy budget via an indirect process described as aerosol-cloud interaction, where the particles serve as cloud condensation nuclei (CCN).

In terms of the direct effects, the ability of particles to scatter light is dependent on both the size of the particles and also the chemical composition of the aerosol particles. Atmospheric conditions play an important role in determining the two aforementioned factors. The ambient relative humidity (RH) can influence a particle's ability to uptake water, which in turn alters the size and chemical composition of the particle. The light scattering of particles is therefore strongly dependent on RH.

The process by which aerosol particles grow as a result of their uptake of water is known as hygroscopicity. For climate models, the σ_{sp} under ambient conditions is required to calculate the aerosol radiative forcing. The study of hygroscopicity is therefore vital to understanding the climatic influence of aerosol particles. Given that ground based in situ instrumentation typically measure σ_{sp} at dry conditions (RH < 40%), then the influence from hygroscopocity is not accounted for.

Hygroscopicity is explored through a property known as the scattering enhancement factor $f(RH, \lambda)$, which is the ratio of scattering induced under ambient conditions and scattering at dry conditions. It is important to note that there are differing definitions of what is considered dry, and as such it is important in the study of hygroscopicity to note the exact RH at which the σ_{sp} is measured. The scattering enhancement factor is described in the following equation:

$$f(RH, \lambda) = \frac{\sigma_{sp}(RH, \lambda)}{\sigma_{sp}(RH_{dry}, \lambda)}$$

The focus of the work presented here is to investigate the direct effects from aerosol particles namely to explore $f(RH, \lambda)$. Model outputs will be used in conjunction with in situ measurements.

1.0.1 Previous studies

Despite the multitude of factors that go into determining the $f(RH, \lambda)$ values for aerosol particles, it has been suggested that there is a pattern among regions, whereby certain areas of the globe experience larger $f(RH, \lambda)$ values. There is a tendency for $f(RH, \lambda)$ values to produce the following inequality: Arctic > Marine > Rural > Desert (Zieger et al., 2010). The Arctic region tends to produce the largest $f(RH, \lambda)$ values, which are up to 3.41 (Zieger et al., 2010).

Burgos et al. (2019) produced a unique dataset of multi-wavelength particle light scattering and hemispheric backscattering coefficents for a range of relative humidities. As noted by Zieger at al. (2010) the largest $f(RH, \lambda)$ values are observed in the Arctic, including the Zeppelin Mountain observatory and the Barrow, Alaska, site. The dataset produced by Burgos et al. (2019) provides a vital resource for climate and atmospheric model-measurments intercomparsions. Furthermore, Burgoes et al. (2018) showed models such as CAM5.3-Oslo overestimates $f(RH = 85$ for Arctic sites.

1.0.2 Motivation

The following work wil explore how well climate models represent aerosol hygroscopic growth and includes these central aims:

- Compare model outputs with in situ observations for f(RH) at Barrow, Alaska
- Assess the influence of the different definitions of dry conditions (i.e. RH=0,40 %)

1.0.3 Import packages

In order to proceed, various packages were imported into the `.JupyterLab` notebook. The packages are listed below:

[20]:
```
import xarray as xr
import numpy as np
import matplotlib.pyplot as plt
import cartopy as cy
import glob
import pandas as pd
from pydap.client import open_dods, open_url
from matplotlib import pyplot as mp
import seaborn as sns
from scipy import stats
from datetime import datetime, timedelta
from matplotlib.lines import Line2D
import cartopy.crs as ccrs
import cartopy
from dask.distributed import Client
from datetime import datetime
import cartopy.crs as ccrs
import cartopy as cy
```

```
import cmocean
import seaborn as sns
import skill_metrics as sm
from matplotlib import rcParams
sns.set_style('white')
sns.set_style('ticks')
import functions as fu
%load_ext autoreload
%autoreload 2
```

```
The autoreload extension is already loaded. To reload it, use:
  %reload_ext autoreload
```

[21]:
```
client = Client(n_workers=2, threads_per_worker=2, memory_limit='8GB')
client
```

```
/opt/conda/lib/python3.7/site-packages/distributed/dashboard/core.py:72:
UserWarning:
Port 8787 is already in use.
Perhaps you already have a cluster running?
Hosting the diagnostics dashboard on a random port instead.
  warnings.warn("\n" + msg)
```

[21]:
```
<Client: 'tcp://127.0.0.1:38850' processes=2 threads=4, memory=16.00 GB>
```

2 Method

In this study, observational data is taken from the unique dataset provided by Burgos et al. (2019) and compared to data from model outputs. The models and the sites are described below in the subsequent subsections.

[22]:
```
## The sites used in this study and their respective latitudes and longitudes
Barrow   =  {'lat' : 71.32, 'lon' : 156.61}
Zeppelin = {'lat' : 78.90, 'lon' : 11.88}
Coords = {'Barrow':Barrow,'Zeppelin':Zeppelin}

site = 'Barrow'
```

2.1 Model data

The model data used in this work includes GEOS-Chem, CAM5, OsloCTM3, GEOS5_GLOBASE and CAM5-ATRAS. Data from the year 2010 was used. The data was selected based on the latitude and longitude of the different sites, using the command `.sel(lat,lon,method='nearest')`. The files were gathered from the resources shared via the JupyterHub. The data was read using the xarray package as it is multi-dimensional and allows all the variables to be assembled together. Separate functions were created and saved in a .py file, so that the code could be removed from the final report, but still be accessible via to a supplementary material. For the purposes of creating a single dataset from multiple ones `xr.open_mfdataset` is used. The absorption and extinction variables were extracted from the files: `abs550dryaer`, `abs550rh40aer`, `abs550rh85aer`, `ec550dryaer`, `ec550rh40aer`, `ec550rh85aer`. The extinction coefficient (σ_{ep}) is equivalent to the sum of the scattering coefficient (σ_{sp}) and the absorption coefficient (σ_{ap}), as such the σ_{sp} at λ=550nm, for each RH value, was computed as follows: σ_{ep} - σ_{ap}. The scattering enhancement factor was subsquentlly calculated by dividing $\sigma_{sp}(RH = 85\%)$ by $\sigma_{sp}(RH = 40\%)$ and $\sigma_{sp}(RH = 0\%)$. The following equation was used to calculate the scattering enhancement factor (depending on the defintion of 'dry conditions'):

$$f(RH, \lambda) = \frac{\sigma_{sp}(RH, \lambda)}{\sigma_{sp}(RH_{dry}, \lambda)}$$

Attributes including the name of the variable and the units were added using `.attrs`.

The files from the models in question were gathered using the `.glob.glob` function. The files contained surface model outputs that were either of an hourly or daily time resolution.

[23]:
```
Chem_files = glob.glob("shared-ns1000k/inputs/hygro/Chem/surface/
↪aerocom3_GEOS-Chem-v10-01_AP3-CTRL2015_*_Surface_2010_hourly.nc")
CAM_files = glob.glob("shared-ns1000k/inputs/hygro/CAM/surface/
↪aerocom3_CAM5_INSITU_*_Surface_2010_hourly.nc")
Oslo_files = glob.glob("shared-ns1000k/inputs/hygro/OSLO-CTM3/
↪aerocom3_OsloCTM3v1.01-met2010_AP3-CTRL_*_Surface_2010_daily.nc")
Globase_files = glob.glob("shared-ns1000k/inputs/hygro/Globase/
↪aerocom3_GEOS5_GLOBASE_*_Surface_2010_daily.nc")
Atras_files = glob.glob("shared-ns1000k/inputs/hygro/Atras/
↪aerocom3_CAM5-ATRAS_INSITU_*_Surface_2010_hourly.nc")
```

Xarray datasets were constructed for each model as follows:

[24]:
```
Chem = fu.read_Chem_files(Chem_files)
CAM = fu.read_CAM_files(CAM_files)
Oslo = fu.read_Oslo_files(Oslo_files)
Globase = fu.read_Globase_files(Globase_files)
Atras = fu.read_Atras_files(Atras_files)
```

Using the functions in the supplementary `.py` file, the absorption and extinction coefficents were used to generate the enhancement factors using the following functions:

[25]:
```
Chem_frh = fu.generate_frh(Chem)
CAM_frh = fu.generate_frh(CAM)
Oslo_frh = fu.generate_frh(Oslo)
Globase_frh = fu.generate_frh(Globase)
Atras_frh = fu.generate_frh(Atras)
```

[26]:
```
variable = 'frh_85_0'

frh_85_0_df = []
df_Chem = fu.generate_site_df(Chem_frh, site, variable, 'Chem')
df_CAM = fu.generate_site_df(CAM_frh, site, variable, 'CAM')
df_Oslo = fu.generate_site_df(Oslo_frh, site, variable, 'Oslo')
df_Globase = fu.generate_site_df(Globase_frh, site, variable, 'Globase')
```

```
df_Atras = fu.generate_site_df(Atras_frh, site, variable, 'Atras')

frh_85_0_df = pd.concat([df_Chem,df_CAM,df_Oslo,df_Globase,df_Atras], axis=1,␣
↪join='inner')
```

```
/home/0e8bf28b-2d6f01-2d43fc-2d9d3f-2de1c44f80cee3/functions.py:77:
RuntimeWarning: Converting a CFTimeIndex with dates from a non-standard
calendar, 'julian', to a pandas.DatetimeIndex, which uses dates from the
standard calendar.  This may lead to subtle errors in operations that depend on
the length of time between dates.
  df.index = df.index.to_datetimeindex()
/home/0e8bf28b-2d6f01-2d43fc-2d9d3f-2de1c44f80cee3/functions.py:74:
RuntimeWarning: Converting a CFTimeIndex with dates from a non-standard
calendar, 'noleap', to a pandas.DatetimeIndex, which uses dates from the
standard calendar.  This may lead to subtle errors in operations that depend on
the length of time between dates.
  df.index = df.index.to_datetimeindex()
```

[27]:
```
variable = 'frh_85_40'

frh_85_40_df = []
df_Chem = fu.generate_site_df(Chem_frh, site, variable, 'Chem')
df_CAM = fu.generate_site_df(CAM_frh, site, variable, 'CAM')
df_Oslo = fu.generate_site_df(Oslo_frh, site, variable, 'Oslo')
df_Globase = fu.generate_site_df(Globase_frh, site, variable, 'Globase')
df_Atras = fu.generate_site_df(Atras_frh, site, variable, 'Atras')

frh_85_40_df = pd.concat([df_Chem,df_CAM,df_Oslo,df_Globase,df_Atras], axis=1,␣
↪join='inner')
```

```
/home/0e8bf28b-2d6f01-2d43fc-2d9d3f-2de1c44f80cee3/functions.py:77:
RuntimeWarning: Converting a CFTimeIndex with dates from a non-standard
calendar, 'julian', to a pandas.DatetimeIndex, which uses dates from the
standard calendar.  This may lead to subtle errors in operations that depend on
the length of time between dates.
  df.index = df.index.to_datetimeindex()
/home/0e8bf28b-2d6f01-2d43fc-2d9d3f-2de1c44f80cee3/functions.py:74:
RuntimeWarning: Converting a CFTimeIndex with dates from a non-standard
calendar, 'noleap', to a pandas.DatetimeIndex, which uses dates from the
standard calendar.  This may lead to subtle errors in operations that depend on
the length of time between dates.
  df.index = df.index.to_datetimeindex()
```

The model data was concatenated together. Monthly medians and the thresholds for the 25th and 75th percentiles were calculated as well in the following functions:

[28]:
```
frh_df = pd.concat([frh_85_0_df,frh_85_40_df], axis=1, join='inner')
frh_df['month'] = frh_df.index.month
```

```
frh_df_monthly = frh_df.groupby(['month']).median()
frh_df_25_monthly = frh_df.groupby(['month']).quantile(q=0.25)
frh_df_75_monthly = frh_df.groupby(['month']).quantile(q=0.75)
frh_df['year'] = frh_df.index.year
frh_df_std = frh_df.groupby(['year']).std()
```

2.2 Observational data

For the Arctic region, there was only enhancement factor data available for Barrow, Alaska and Zeppelin Observatroy, Svalbard. The measurements from the observatory in Barrow recorded enhancement factor values for ten out of the twelve months of the year, whilst the Zeppelin Observatory had a dataset consisiting of only 3 months.
The observational data was downloaded from the EBAS webpage (http://ebas.nilu.no/). The year 2010 was selected to collocate with the model data. The data was cleaned by removing the 'flagged' data. The flags are present in the dataset due to ivalid data. dditionally, for the monthly medians a minimum count of 10 was imposed on the data such that a month which contained less than 10 data points was excluded. The in situ $f(RH)$ values from Barrow were measured using a humidified nephelometer system.

[29]:
```
appended_data = []

files = glob.glob("Barrow/enhancement/*.nas")
for file in files:
    df = pd.read_csv(file,sep='\s+',parse_dates=True,skiprows=79)
    start_time = file[27:41]
    datetime_object = datetime.strptime(start_time, '%Y%m%d%H%M%S')
    df['endtime'] = df.endtime.astype(float)
    df['Datetime'] = datetime_object + pd.to_timedelta(df['endtime'], unit='d')
    df = df.set_index(pd.DatetimeIndex(df['Datetime']))
    df.index = df.index.round('H')
    appended_data.append(df)
appended_data = pd.concat(appended_data)

appended_data[['scat_enh.2','scat_enh.3']]

scat550_frh = appended_data[['scat_enh.2','scat_enh.3']]
scat550_frh['year'] = scat550_frh.index.year
scat550_frh = scat550_frh[scat550_frh.year==2010]
scat550_frh = scat550_frh.rename(columns={"scat_enh.2": "frh_85_40", "scat_enh.
 ↪3": "frh_85_0"})

scat550_frh.frh_85_40[scat550_frh.frh_85_40 >= 99] = np.nan
scat550_frh.frh_85_0[scat550_frh.frh_85_0 >= 99] = np.nan

scat550_frh_std = scat550_frh.groupby(['year']).std()

scat550_frh = scat550_frh.drop(['year'],axis=1)
```

```
scat550_frh['month'] = scat550_frh.index.month

scat550_frh_monthly = scat550_frh.groupby(['month']).median()
scat550_frh_monthly_count = scat550_frh[['frh_85_40','frh_85_0','month']].
 ↪groupby(['month']).count()
scat550_frh_monthly[scat550_frh_monthly_count[['frh_85_40','frh_85_0']] < 10] =␣
 ↪np.nan

scat550_frh_25 = scat550_frh.groupby(['month']).quantile(q=0.25)
scat550_frh_25[scat550_frh_monthly_count[['frh_85_40','frh_85_0']] < 10] = np.
 ↪nan

scat550_frh_75 = scat550_frh.groupby(['month']).quantile(q=0.75)
scat550_frh_75[scat550_frh_monthly_count[['frh_85_40','frh_85_0']] < 10] = np.
 ↪nan
```

```
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:18:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:22:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
/opt/conda/lib/python3.7/site-packages/ipykernel_launcher.py:23:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
```

2.2.1 Merge model and observational data

The model and observational data were merged in order to create a dataframe with collocated data points based on the `.datetime` index. Temporal collocation was done based on hourly values.

[30]:
```
model_obser = pd.concat([scat550_frh, frh_df], axis=1, join='inner')
model_obser = model_obser.drop(['month'],axis=1)

monthly_model_obser = pd.concat([scat550_frh_monthly, frh_df_monthly], axis=1,␣
 ↪join='inner')
monthly_model_obser_25 = pd.concat([scat550_frh_25, frh_df_25_monthly], axis=1,␣
 ↪join='inner')
```

```
monthly_model_obser_75 = pd.concat([scat550_frh_75, frh_df_75_monthly], axis=1,␣
↪join='inner')
```

3 Results and discussion

The following section explores the results generated from model and measurement comparsions. The results are presented within the notebook, however the majority of the code required to generate the plots has been saved in a complimentary file called `functions.py`.

3.0.1 Annual plots

[31]:
```
variable = 'frh_85_40'
df = monthly_model_obser
df_25 = monthly_model_obser_25
df_75 = monthly_model_obser_75

fu.create_annual_plot(df, df_25, df_75, variable)
```

[31]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fd058f72278>
```

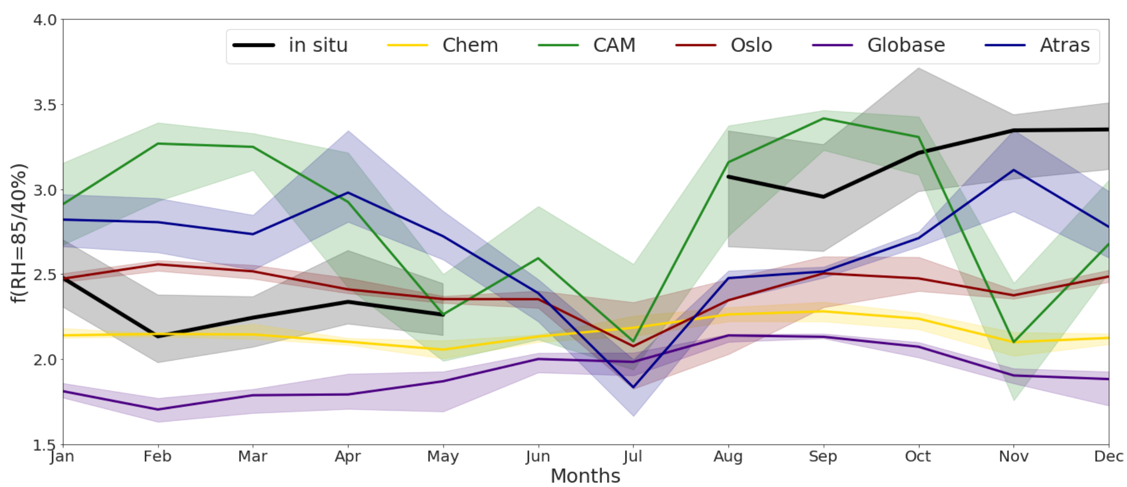


Fig. 1: shows the annual cycle for $f(RH,85/40\%)$. Each respective colour represents a model output and the black line represents the in situ measurements.

[32]:
```
\
variable = 'frh_85_0'
df = monthly_model_obser
df_25 = monthly_model_obser_25
df_75 = monthly_model_obser_75

fu.create_annual_plot(df, df_25, df_75, variable)
```

[32]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fcfef70e588>
```

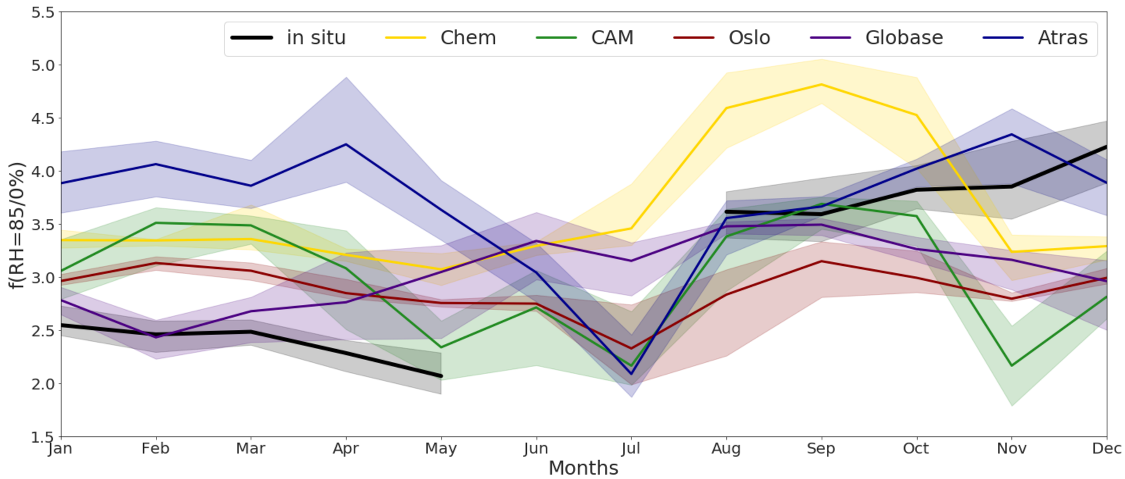


Fig. 2: shows the annual cycle for $f(RH, 85/0\%)$. Each respective colour represents a model output and the black line represents the in situ measurements.

Figures 1 and 2 display the annual variability among various model outputs. The annual cycle for the measurements taken at Barrow are displayed for comparsion. It is noticeable that the $f(RH)$ values are greater when the 0% definition for dry conditions is used, as opposed to 40%.

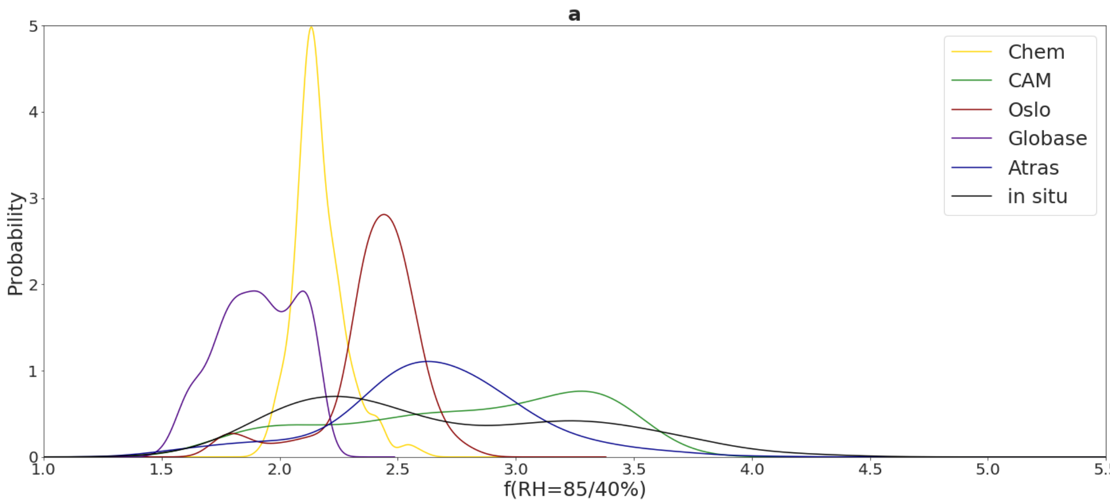
Overall, the variability of the models is small; specifically, the models Chem, Globase and Oslo show little annual variability. The median value changes little throughtout the year. Oslo, however, demonstrates an increased interquartile range during the summer. The variability is greater in figure 2 compared with figure 1.

For the in situ measurements, the months corresponding to late winter and spring have relatively low magnitudes, whilst the months towards the end of year show increased medians. The fact that the instrumentation was down during parts of the summer and that when it began measuring again it show increased values may signify potential errors in the data. The increased values in the late summer and autumn may represent the influence of increased sea salt aerosol (SSA) particles. SSA particles generally exhibit higher enhancement factor values. However, there are numerous factors that contribute to the overall $f(RH)$ values including the size of the particles, so additional evidence is require to further support this hypothesis.

3.0.2 Frequency plots

[33]:
```
variable = 'frh_85_40'
ax = fu.generate_frequency_plot(variable, frh_df, scat550_frh)

variable = 'frh_85_0'
ax = fu.generate_frequency_plot(variable, frh_df, scat550_frh)
```



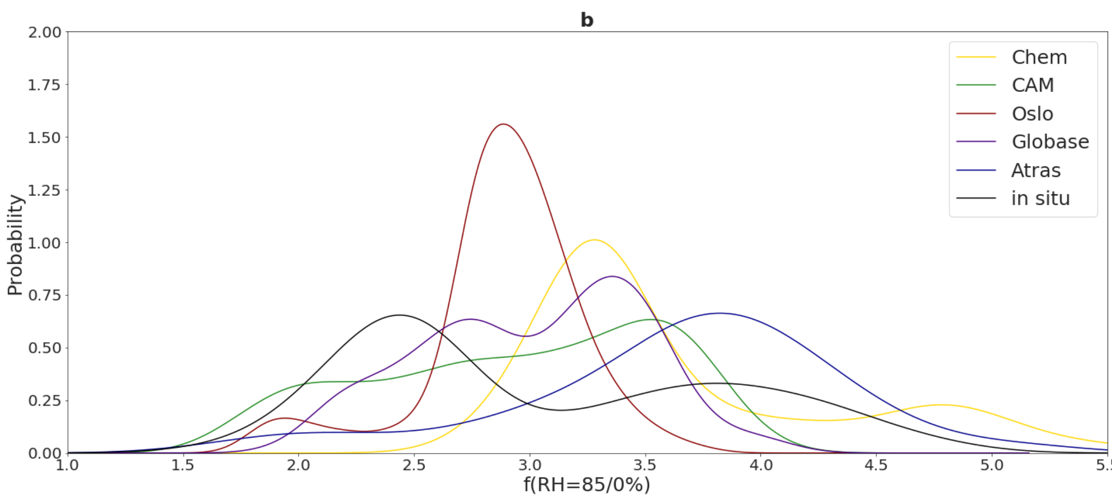


Fig. 3: a,b shows a normalised frequency plot, whereby the area under the probability density functions (PDFs) corresponds to 1. The colours yellow, green, dark red, indigo and dark blue correspond to the model outputs from Chem, CAM, Oslo, Globase and Atras respectively. a displays the variable $f(RH, 85/40\%)$, whilst b displays the PDF for $f(RH, 85/0\%)$.

The majority of the models produce a unimodal output (Chem, Atras and Oslo), however the in situ measurements display a bimodal output one. The lack of variability in the model outputs, compared with the observational data, is apparent in Figure 3.

3.1 Regression plots

Regression plots are used to compare the outputs from models to the in situ measurements, and explore which models and variables have the greatest discrepancies.

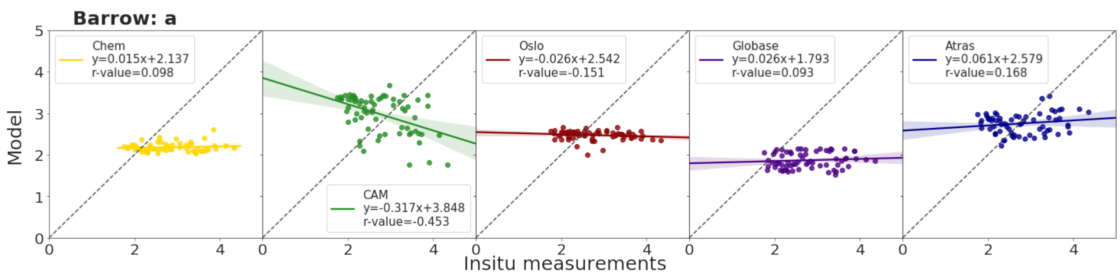
[34]:
```
variable = 'frh_85_40'
merged_df = model_obser
site = 'Barrow'

fu.generate_regression_plots(merged_df, variable, site)

variable = 'frh_85_0'
merged_df = model_obser
site = 'Barrow'

fu.generate_regression_plots(merged_df, variable, site)
```

[34]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fcfd7416940>
```



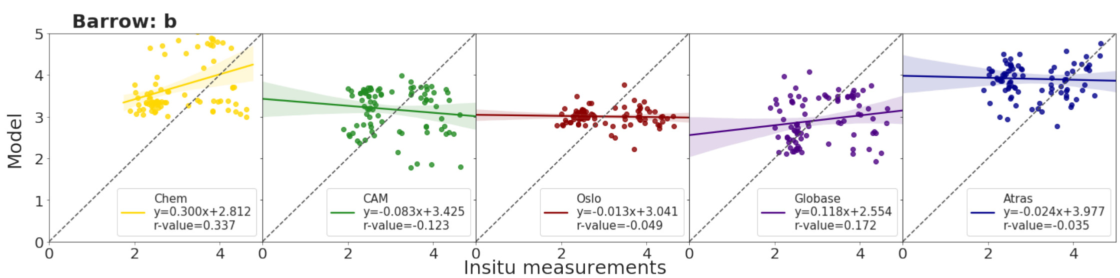


Fig. 4: a, b shows the regression plots comparing model output with in situ measurements for the Barrow. The colours yellow, green, dark red, indigo and dark blue correspond to the model outputs from Chem, CAM, Oslo, Globase and Atras respectively. a displays the variable $f(RH, 85/40\%)$, whilst b displays the regression plots for $f(RH, 85/0\%)$.

It is clear from the regression plots (see Fig. 4) that there is very little correlation between model outputs and the in situ measurements. The Chem model displays the greatest correlation, however the r-values are essentially insignificant. The CAM model displays the worst correlations between model and measurement. There is noticeably more variability in the model outputs for $f(RH, 85/0\%)$, as opposed to $f(RH, 85/40\%)$. For $f(RH, 85/40\%)$, all the models except CAM show little variability.

3.2 Taylor diagrams

Taylor diagrams are used to summarise the main statistics in model and observation comparison. The diagrams present the correlation coefficient and the standard deviation. Notice that the observations have a correlation coefficient of 1, as the data correlates perfectly with itself. The Python package `.SkillMetrics` was utilised to develope the Taylor diagrams.

[35]:
```
sm = fu.generate_taylor(merged_df,'frh_85_40', site, frh_df_std,␣
↪scat550_frh_std) #ccoef_frh_85_40
```

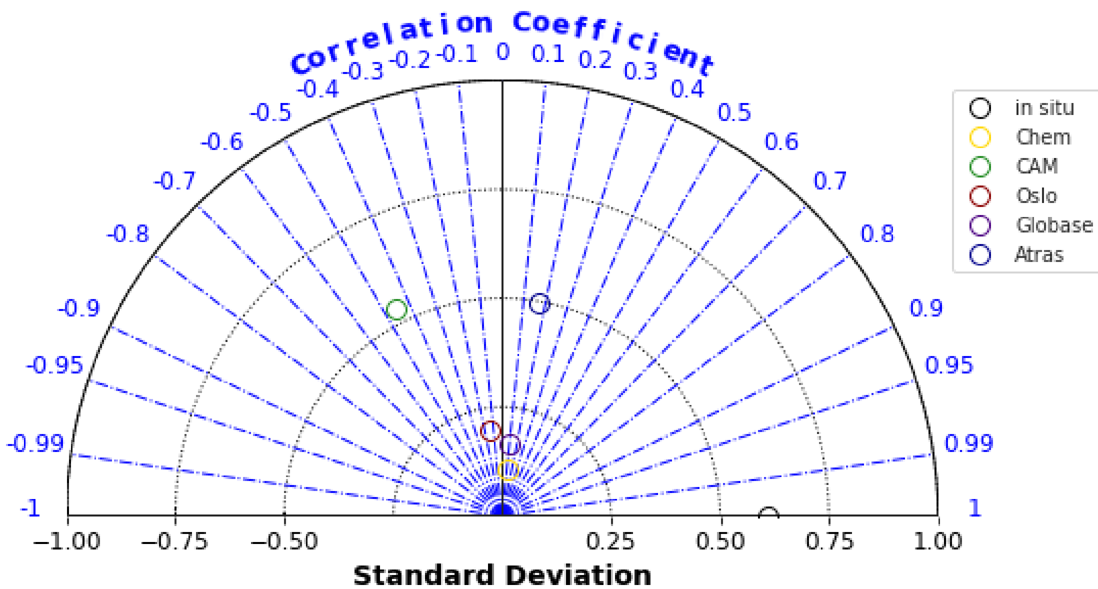


Fig. 5: shows the Taylor diagram for the variable $f(RH, 85/40\%)$, comparing model output with in situ measurements for the Barrow. The colours yellow, green, dark red, indigo and dark blue correspond to the model outputs from Chem, CAM, Oslo, Globase and Atras respectively.

[37]:
```
sm = fu.generate_taylor(merged_df,'frh_85_0', site, frh_df_std, scat550_frh_std)
```

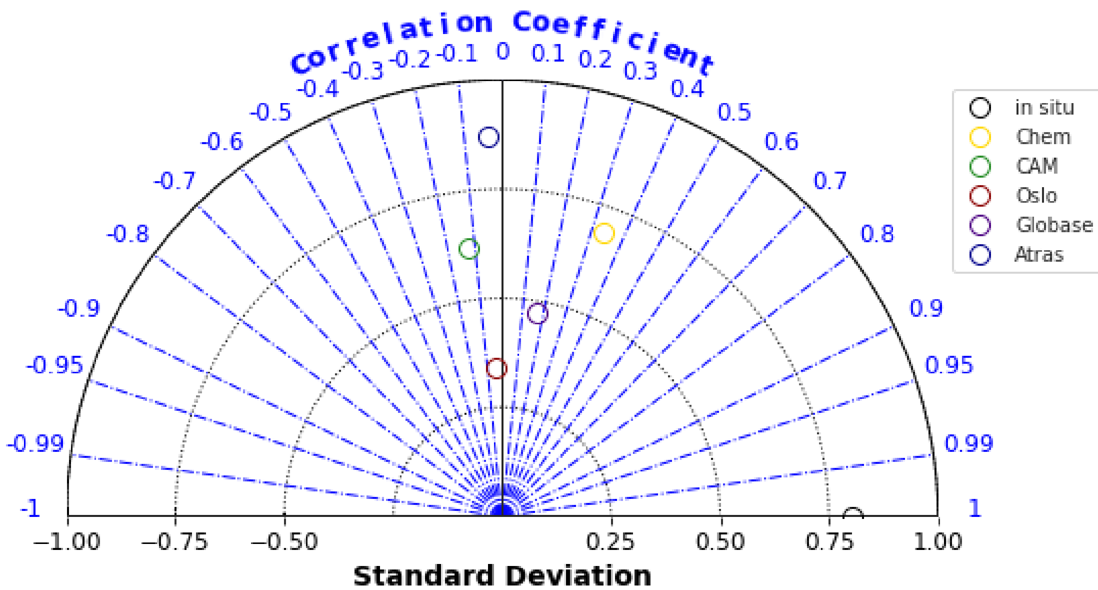


Fig. 6: shows the Taylor diagram for the variable $f(RH, 85/0\%)$, comparing model output with in situ measurements for the Barrow. The colours yellow, green, dark red, indigo and dark blue correspond to the model outputs from Chem, CAM, Oslo, Globase and Atras respectively.

From Figures 6 and 7, it is quite apparent that the correlation coefficients for the model are extremely poor and that in some cases negative. It is also noticeable that the standard deviation is greater for $f(RH, 85/0\%)$ as opposed to $f(RH, 85/40\%)$.

[38]:
```
month_to_season =  {1:'DJF',  2:'DJF',
                    3:'MAM',  4:'MAM',  5:'MAM',
                    6:'JJA',  7:'JJA',  8:'JJA',
                    9:'SON', 10:'SON', 11:'SON',
                   12:'DJF'}
merged_df['month'] = merged_df.index.month
merged_df['season_abb'] = merged_df.month.map(month_to_season).values
```

[39]:
```
df_median = merged_df.groupby(['season_abb']).median()
df_25 = merged_df.groupby(['season_abb']).quantile(0.25)
df_75 = merged_df.groupby(['season_abb']).quantile(0.75)
```

[40]:
```
site = 'Barrow'
variable = 'frh_85_40'
ax = fu.generate_seasonal_plot(variable, site, df_median, df_25, df_75)
```

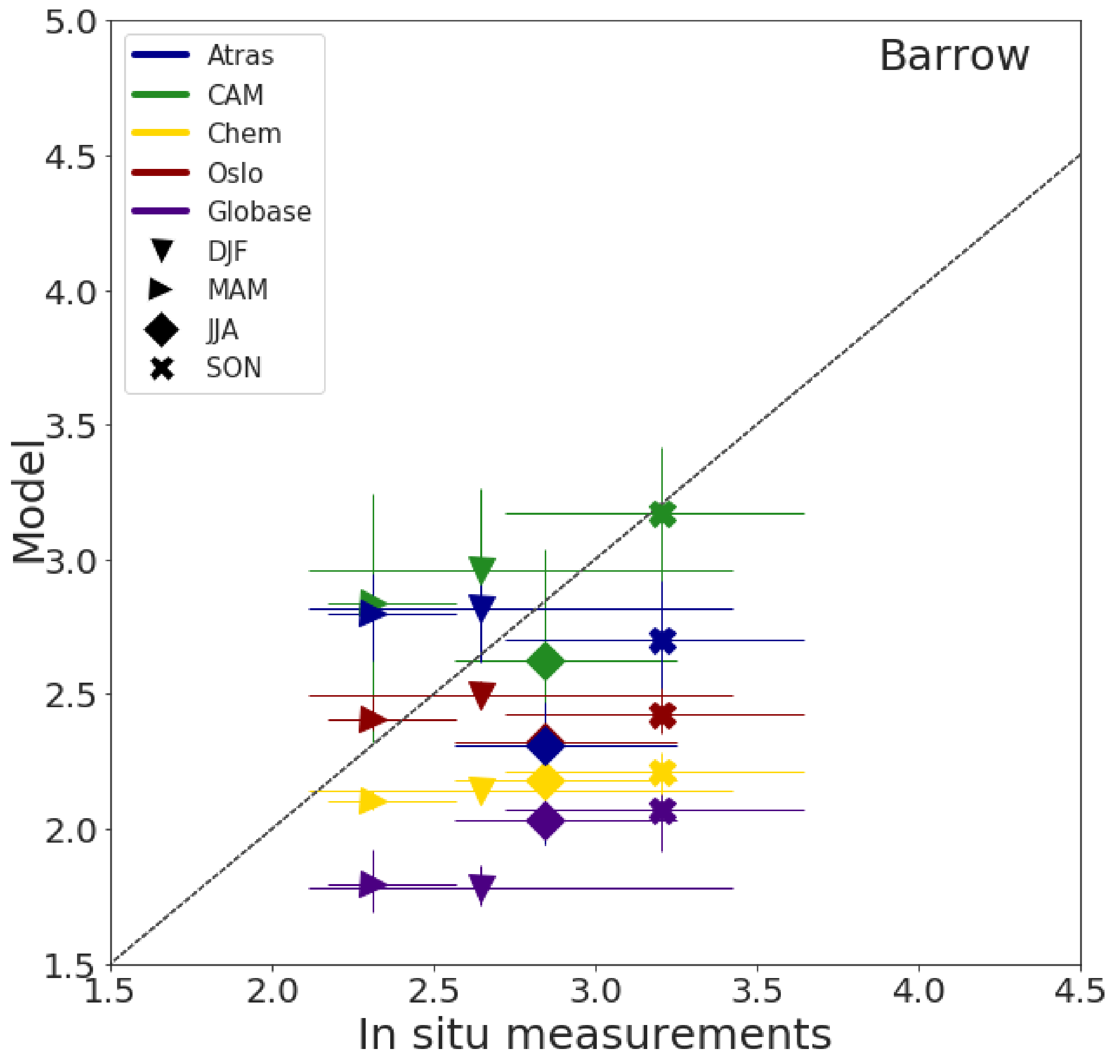


Fig. 7: shows the regression plots comparing model output with in situ measurements for $f(RH, 85/40\%)$ at Barrow. The colours yellow, green, dark red, indigo and dark blue correspond to the model outputs from Chem, CAM, Oslo, Globase and Atras respectively.

[42]:
```
site = 'Barrow'
variable = 'frh_85_0'
ax = fu.generate_seasonal_plot(variable, site, df_median, df_25, df_75)
```

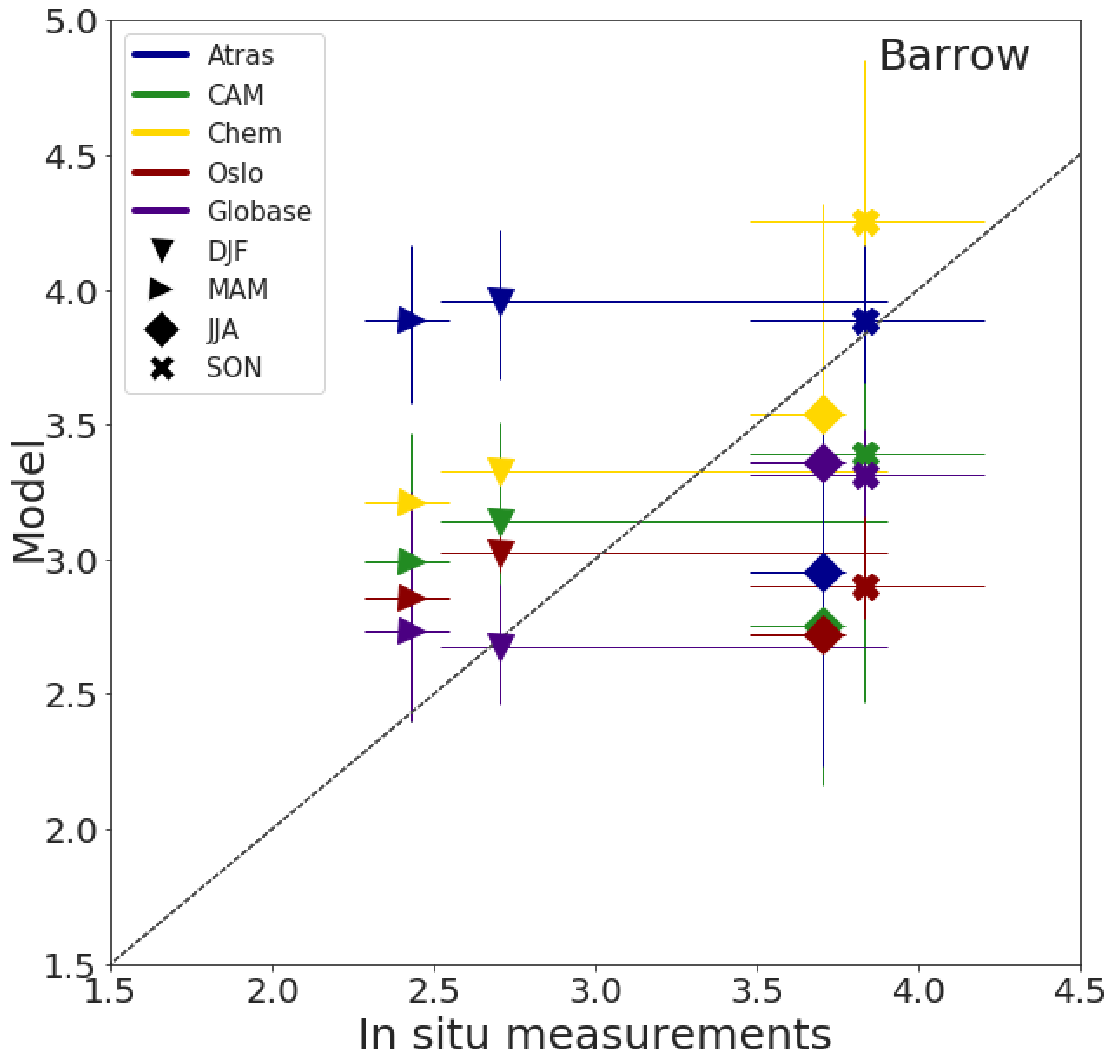


Fig. 8: shows the regression plots comparing model output with in situ measurements for $f(RH, 85/0\%)$ at Barrow. The colours yellow, green, dark red, indigo and dark blue correspond to the model outputs from Chem, CAM, Oslo, Globase and Atras respectively.

There is a clear seaonal dependency with model and in situ meaurement comparsions. Spring and winter tend to be overestimated by the model, whilst, summer and autumn are underestimated. The seaonal dependency is displayed in both figures 7. and 8, however it is clearer for $f(RH, 85/0$.

3.3 Mapped model outputs

[44]:
```
frh_85_40, frh_85_0 = Chem_frh
frh_85_40_season, frh_85_0_season = fu.generate_season(frh_85_40, frh_85_0)
p = fu.seasonal_plot(frh_85_0_season, 'Chem', 'frh_85_0')

p = fu.seasonal_plot(frh_85_40_season, 'Chem', 'frh_85_40')
```

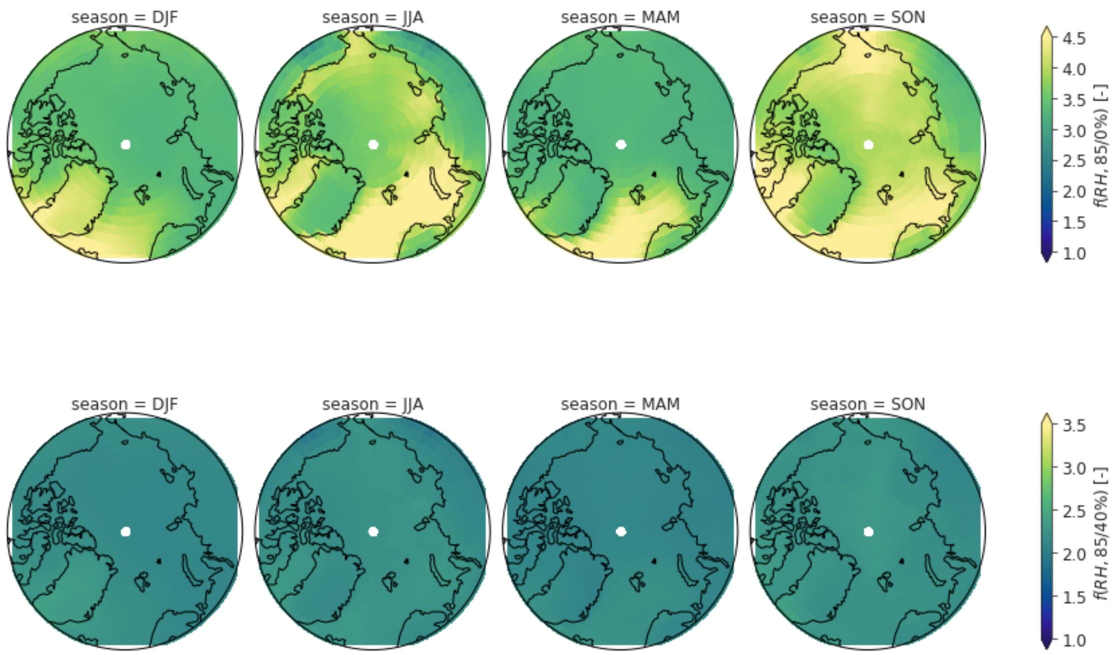
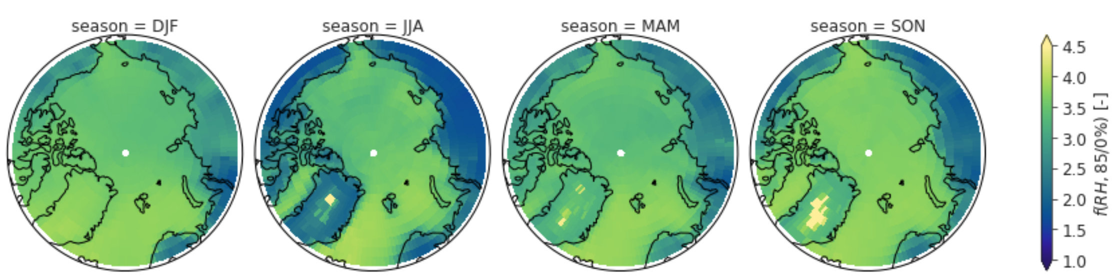


Fig. 9: shows the north pole stereographic mapping for the Chem model output for $f(RH, 85/0\%)$ (above) and $f(RH, 85/40\%)$ (below).

The model output for Chem displays a clear increase in $f(RH, 85/0\%)$ for the areas over the open ocean. There is a slight contrast between the regions above land and regions above the ice sheet, especially during autumn and summer. The enhanced gradient between ice sheets and land masses during autumn and summer may be as a result of the retreating of the Arctic sea ice and/or increased biological activity during periods that experienced solar activity. Notice that there is no variability for the variable $f(RH, 85/40\%)$ suggesting that the differences in $f(RH)$ between different surfaces occur between 85% and 0%.
Considering that Chem for $f(RH, 85/0\%)$ demonstrated the largest correlation coefficient between model outputs and observations, it is interesting to see that the Chem output displays the clearest gradient between oceans and other land types (including the Arctic sea ice).

[45]:
```
frh_85_40, frh_85_0 = CAM_frh
frh_85_40_season, frh_85_0_season = fu.generate_season(frh_85_40, frh_85_0)
p = fu.seasonal_plot(frh_85_0_season, 'CAM', 'frh_85_0')

p = fu.seasonal_plot(frh_85_40_season, 'CAM', 'frh_85_40')
```



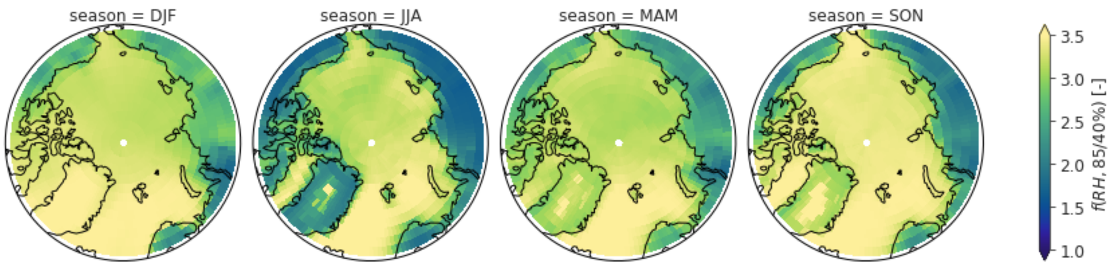


Fig. 10: shows the north pole stereographic mapping for the CAM model output for $f(RH, 85/0\%)$ (above) and $f(RH, 85/40\%)$ (below).

It is clear from the CAM model outputs that the enhancement factors are too large. The CAM model has been shown to have overestimated the emissions from SSA particles. CAM displayed the worst correlation coefficient for the variable $f(RH, 85/0\%)$.

[46]:
```
frh_85_40, frh_85_0 = Oslo_frh
frh_85_40_season, frh_85_0_season = fu.generate_season(frh_85_40, frh_85_0)
p = fu.seasonal_plot(frh_85_0_season, 'Oslo', 'frh_85_0')

p = fu.seasonal_plot(frh_85_40_season, 'Oslo', 'frh_85_40')
```

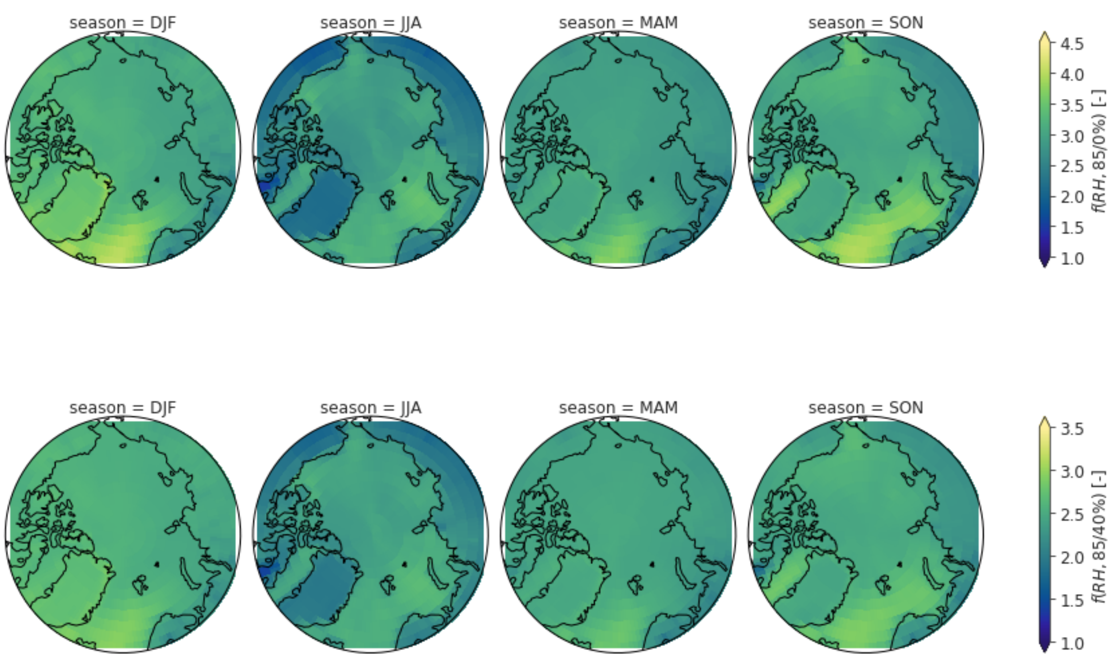
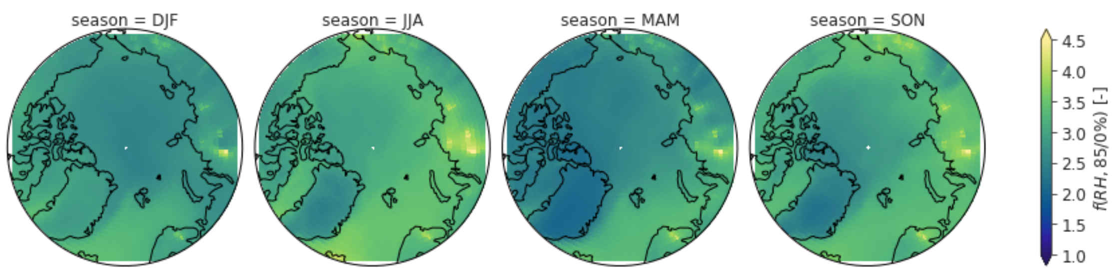


Fig. 11: shows the north pole stereographic mapping for the Oslo model output for $f(RH, 85/0\%)$ (above) and $f(RH, 85/40\%)$ (below).

[47]:
```python
\
frh_85_40, frh_85_0 = Globase_frh
frh_85_40_season, frh_85_0_season = fu.generate_season(frh_85_40, frh_85_0)
p = fu.seasonal_plot(frh_85_0_season, 'Globase', 'frh_85_0')

p = fu.seasonal_plot(frh_85_40_season, 'Globase', 'frh_85_40')
```



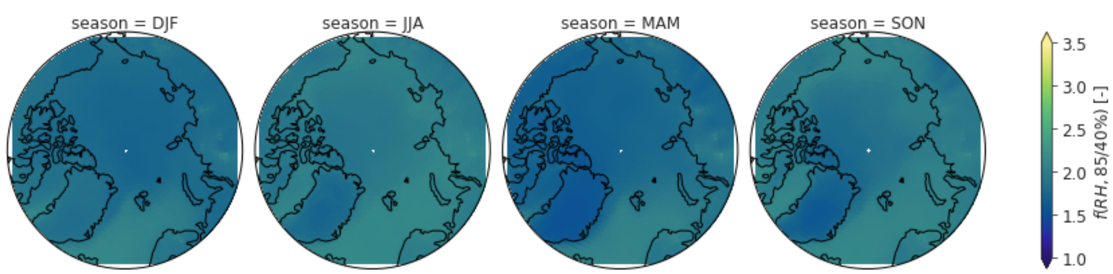
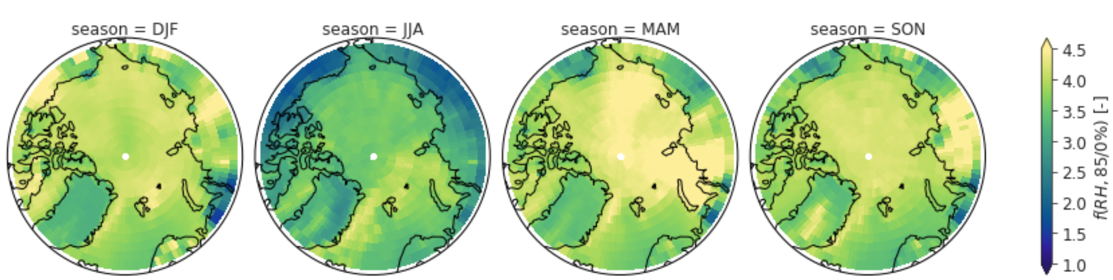


Fig. 12: shows the north pole stereographic mapping for the Globase model output for $f(RH, 85/0\%)$ (above) and $f(RH, 85/40\%)$ (below).

In Figure 12, there are small regions above Siberia that produce large $f(RH, 85/0\%)$ values. It would be intereting to explore the exact nature of these seemingly unusual areas of high $f(RH, 85/0\%)$.

[48]:
```python
frh_85_40, frh_85_0 = Atras_frh
frh_85_40_season, frh_85_0_season = fu.generate_season(frh_85_40, frh_85_0)
p = fu.seasonal_plot(frh_85_0_season, 'Atras', 'frh_85_0')

p = fu.seasonal_plot(frh_85_40_season, 'Atras', 'frh_85_40')
```



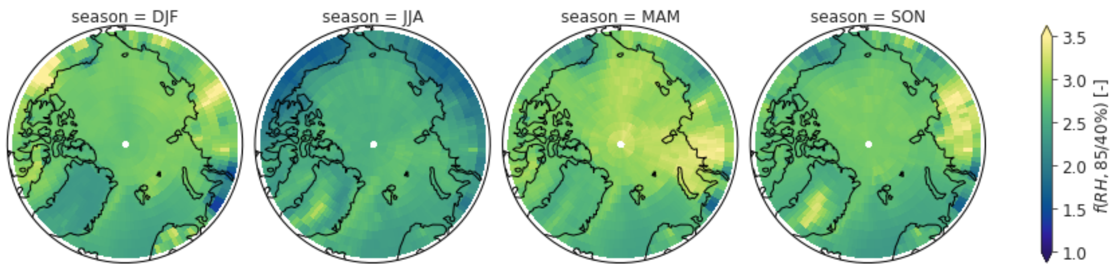


Fig. 13: shows the north pole stereographic mapping for the Atras model output for $f(RH, 85/0\%)$ (above) and $f(RH, 85/40\%)$ (below).

4 Conclusion

The study presented here focuses on measurement and model comparsions. The relative humidity enhancement factor is examined. The site Barrow in Alaska was used for the comparsions.

Overall, the models analysed do not correlate with in situ measurements. Measurements show higher variability while models present a narrower distribution of values.

The models tend to displays higher $f(RH)$ values over the ocean, which could be an indication of the increased presence from sea salt aerosol (SSA) particles. The continents tend to show lower $f(RH)$ values possibly as a result of black carbon (BC), organics, or mineral dust.

Summer and autumn are typically underestimated by models, which could potentially be a result of summer emissions from BC and organics being weighted too much in the models. Alternatively, SSA emissions may be weighted too little in the models.

$f(RH)$ values for the seasons winter and spring tend to be overestimated by the models, which could be due to SSA emissions being weighted too much or sources from the continent being overestimated in models.

However, chemical composition is just one factor that influences the $f(RH)$ values. For example, particles which are less hygroscopic, but smaller could produce similar $f(RH)$ values to particles which are more hygroscopic, but larger. In order to properly examine the model outputs and discrepancies between in situ measurements and models, size distributions will also need to be analysed.

5 References

[1] Burgos MA, Andrews E, Titos G, Buchard V, Randles C, Kirkevag A, Zieger P (2018, October). Comparison between in-situ surface measurements and global climate model outputs of particle light scattering coefficient as a function of relative humidity, 17th AeroCom meeting, College Park, Maryland, USA.

[2] Burgos MA, Andrews E, Titos G, Alados-Arboledas L, Baltensperger U, Day D, Jefferson A, Kalivitis N, Mihalopoulos N, Sherman J, Sun J. A global view on the effect of water uptake on aerosol particle light scattering. Scientific data. 2019 Aug 22;6(1):1-9.

[3] IPCC. The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. 2013. There is no corresponding record for this reference. 2013:33-118.

[4] Zieger P, Fierz-Schmidhauser R, Gysel M, Ström J, Henne S, Yttri KE, Baltensperger U, Weingartner E. Effects of relative humidity on aerosol light scattering in the Arctic. Atmospheric Chemistry and Physics. 2010 Apr 27;10(8):3875-90.