NeGIcourse_2019_Jolanta_Rieksta

November 12, 2019

0.1 Model evaluation of the performance of NorESM2-LM vs UKESM1-0-LL in predicting BVOC [Isoprene] in the Arctic

Report for the course:

Climate science at high latitudes: eScience for linking Arctic measurements and modeling

Group 1, Assisstant: Sara M. Blitcher

Jolanta Rieksta

PhD student, University of Copenhagen, Department of Biology, Terrestrial Ecology Section

1 Abstract

Biogenic volatile organic compounds (BVOCs) emitted by plants play a significant role in atmospheric chemistry. The oxidation of BVOCs lead to ozone and aerosol production. Terrestrial ecosystems are the largest conbtributors of BVOCs to the atmosphere and are the largest source of secondary organic aerosols (SOA).

This study aims to investigate the performance of two models available at CMIP6 (NorESM2-LM and UKESM1-0-LL) in predicting most abundant BVOC – isoprene, in the Arctic, as each of these models are using different land components and chemistry part. BVOC estimates associated with high uncertainty especially in the Arctic, that leads to substantial difference in predicted emission rates among different models, because emission factors and land types can differ. This study aims to investigate the performance of two models available at CMIP6 (NorESM2-LM and UKESM1-0-LL) in predicting most abundant BVOC – isoprene, in the Arctic, as each of these models are using different land components and chemistry part.

We found: 1) substantial differences in predicted emission rates between two models; 2) Overall UKESM1-0-LL predicted higher isoprene concentrations compared to NorESM2-LM, except in certain regions in Tundra, where NorESM2_LM had higher concentrations of isoprene; 3) When model data are compared with observation data, UKESM1-0-LL is ovestimating isoprene concentrations.

In general, models are not wery well representing the BVOC emissions in the Arctic. This study showed that NorESM2-LM and UKESM1-0-LL models predict substantial difference in isoprene concentrations. Better understanding and more studies with other BVOCs are needed on asssessing how parameters differ and how they change the emissions and concentrations between both models when predicting BVOCs in the Arctic.

2 Introduction

Biogenic volatile organic compounds (BVOCs) emitted from vegetation are organic atmospheric trace gases other than carbon dioxide and monoxide (Kesselmeier & Staudt, 1999). Plants emit BVOCs in response to biotic (attracting pollinators and herbivore predators, sealing wounds, deterring insect herbivores and pathogens) (Dicke, 2009; Dicke & Baldwin, 2010; Kessler & Baldwin, 2001; Scala, Allmann, Mirabella, Haring, & Schuurink, 2013) or abiotic (increasing thermotolerance) (Peñuelas & Staudt, 2010) factors.

BVOCs play a significant role in atmospheric chemistry and physics (Peñuelas & Staudt, 2010). The oxidation of BVOCs lead to ozone and aerosol production (Goldstein & Galbally, 2007; Holopainen, Kivimäenpää, & Nizkorodov, 2017) (Figure 1). Terrestrial ecosystems are the largest conbtributors of BVOCs to the atmosphere (Lamarque et al., 2010) and are the largest source of secondary organic aerosols (SOA) (Bonn, Kuhlmann, & Lawrence, 2004).

Plants emit high diversity of BVOCs, however the most important volatiles that contribute to the aerosol formation are isoprene, monoterpenes and sesquiterpenes (Kulmala et al., 2013). The importance of these volatiles on atmospheric chemistry has lead BVOCs to be commonly incorporated in various models as essential inputs (Jiang et al., 2019). However, BVOC estimates associated with high uncertainty especially in the Arctic, that leads to substantial difference in predicted emission rates among different models, because emission factors and land types can differ (Jiang et al., 2019).

This study aims to investigate the performance of two models available at CMIP6 (NorESM2-LM and UKESM1-0-LL) in predicting most abundant BVOC – isoprene, in the Arctic, as each of these models are using different land components and chemistry part (Arneth et al., 2007; Guenther et al., 2012; Sporre, Blichner, Karset, Makkonen, & Berntsen, 2019). This study hypothesize that models will substantially differ in their performance to predict isoprene in the Arctic. Here we test it by: (1) to evaluating the seasonal differences between the models; (2) calculating the correlation between two models; (3) comparing the model data with observations from different stations in the Arctic using EBAS database.

3 Methods and data

We used NorESM2-LM and UKESM1-0-LL models in CMIP6. We used output from 'historical' (year 1980 - 2014). We used Isoprene Volume Mixing Ratio e.g. mole fraction of isoprene in the air (mol mol-1) as a variable to compare the performance of the models. We used Total Emission Rate of Isoprene (kg m-2 s-1) as an input in *pyaerocom* package to see the relationship between the two models.

All analysis were performed in Jupyter notebook (Python 3.7.3) (Kluyver et al., 2016).

To test the model performance with the observation data, we obtained observation data of isoprene from three research stations Birkenes (1992-01-01 - 1999-12-30), Birkenes II (2012-04-28 - 2013-02-08) and Zeppelin mountain (Ny-Ålesund) (1992-01-02 - 1999-06-11) which are available on EBAS-NILU database (http://ebas.nilu.no). The region of interest was Arctic, therefore we selected data with latitudes(60–90 N).

To be able to compare the models we collocated data using **pya.colocation.colocate_gridded_gridded** function from *pyaerocom* package (0.8.1.dev4). To be able to compare the models with observation we collocated data using **pya.colocation.colocate_gridded_ungridded** function from *pyaerocom* package (0.8.1.dev4).

To evaluate the seasonal differences between the models we calculated the mean (1980-2014) for each model and assessed the percent difference and graphically represented them using **ccrs** function from *cartopy* package.

To evaluate the seasonal differences between the models and observations, we calculated the mean concentrations for each month and compared the BVOC emissions between the models and observations.

3.1 Other packages used

xarray
pyaerocom numpy
matplotlib
netCDF4
cartopy
sys glob pandas

4 Results

Plot NorthPolarStereo of Isoprene Volume Mixing Ratio (mol mol-1) for seasons for each of the models and obtain the diffrences between the models.

[15]:
```
ds_seas_u = ukesm_isop.groupby('time.season').mean() #Group data by the season␣
 ↪and obtain the mean UKESM model
ds_seas_n = noresm_isop.groupby('time.season').mean() #Group data by the season␣
 ↪and obtain the mean NorESM model

ds_seas_rat = (ds_seas_n-ds_seas_u)/ds_seas_u*100
ds_seas_rat = ds_seas_rat.where(ds_seas_u.isop>ds_seas_u.quantile(.75))

fig, axs = sp_map(2,2, projection=ccrs.NorthPolarStereo(), figsize=[10,10] )
lat_lims = [60,90]
season_l = ds_seas['season'].values
print(season_l)
plt_ds = ds_seas_rat
for ax, seas in zip(axs.flatten(),season_l):
    _ds= plt_ds.sel(season=seas)['isop'].where(ds_seas['lat']>lat_lims[0])
    _ds.attrs['units']='%'; _ds.attrs['long_name'] = 'Isoprene'
    _ds.plot(ax=ax, robust=True,transform=ccrs.PlateCarree())
    ax.set_title(seas)
    polarCentral_set_latlim(lat_lims, ax)
    add_map_features(ax)
plt.show()
```

```
['DJF' 'JJA' 'MAM' 'SON']
```

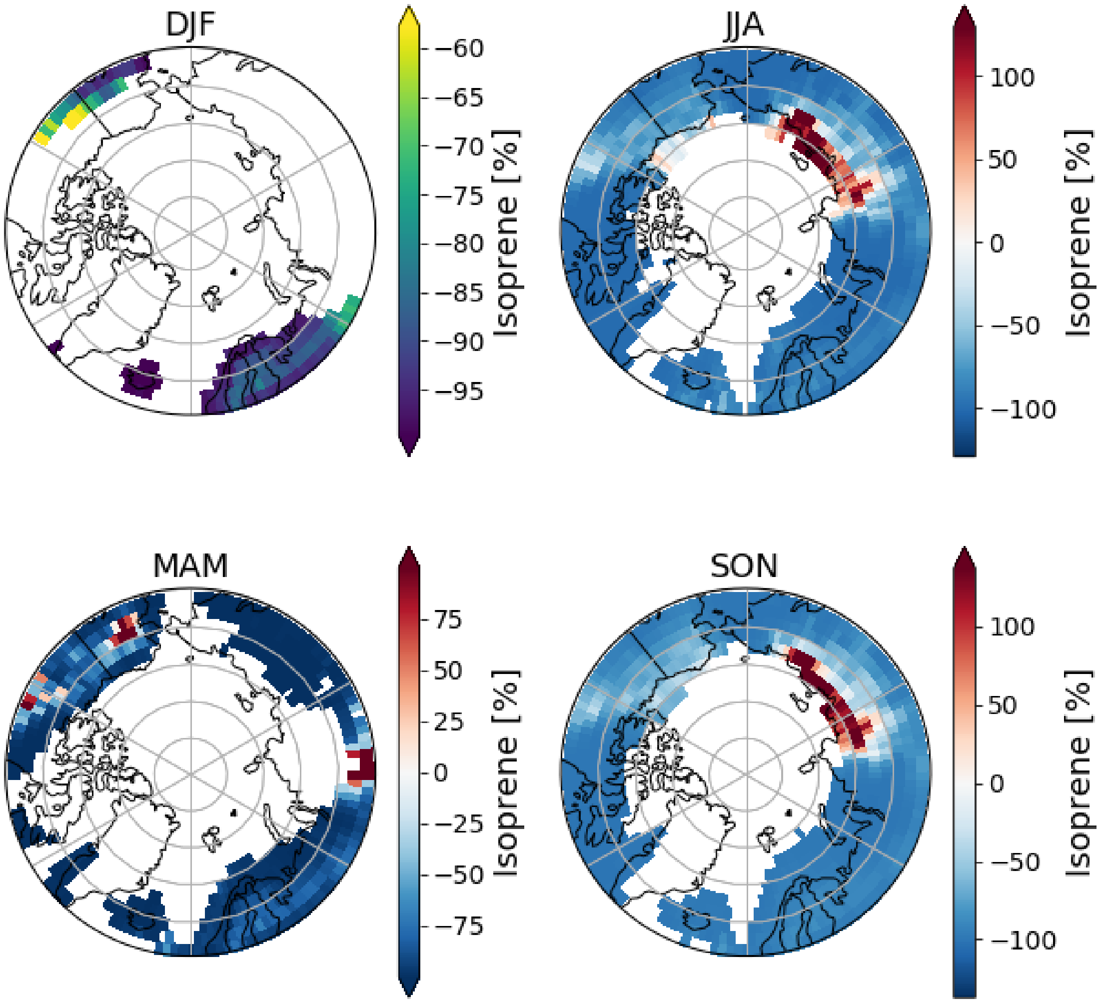


Figure 1. Percent difference between the models (1980-2014) of isoprene concentrations between the two models. With blue color indicating UKESM model having higher isoprene concentrations than NorESM model, and red color indicating vice vera.**DJF**= December, January, February; **MAM** = March, April, May; **JJA** = June, July, August and **SON**= September, Oktober, November.

Table 1 Percent difference in isoprene concentrations between models: NorESM2-LM and UKESM1-0-LL.

[112]:
```
#means of Isoprene [mol/mol] for each of the models
lab_iso = 'Isoprene [mol/mol]'
means=pd.DataFrame(index=['NorESM', 'UKESM', 'UKESM-NorESM %'],␣
 ↪columns=['Isoprene [mol/mol]'])
means.loc['NorESM', lab_iso] = 7.8192895e-12
means.loc['UKESM', lab_iso] = 1.13133856e-10

means.loc['UKESM-NorESM %', :]= (means.loc['UKESM', :]- means.loc['NorESM', :])/
 ↪means.loc['NorESM',:]*100
```

```
means
```

[112]:

	Isoprene [mol/mol]
NorESM	7.81929e-12
UKESM	1.13134e-10
UKESM-NorESM %	1346.86

***Pyaerocom* for collocation of the data and obtaining statistics for model comparison**

[128]:
```
coldata1.plot_scatter(marker='o', color='blue', alpha=0.1)
plt.xlim(xlim)
plt.ylim(ylim)
plt.tight_layout()
```

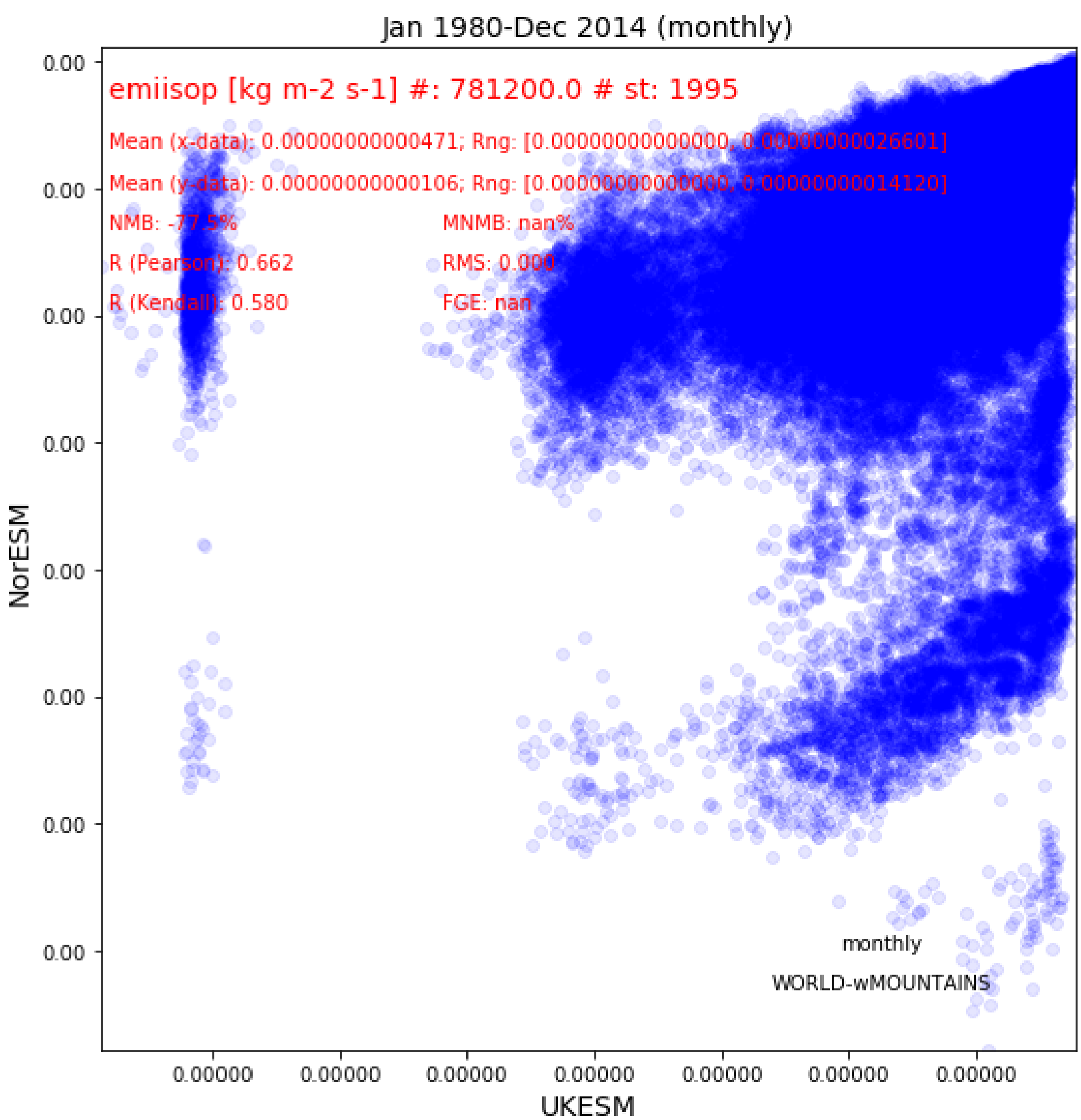


Figure 2 The relationship between collocated NorESM and UKESM models for Total Emission Rate of Isoprene (kg m-2 s-1).

Temporal variations of isoprene (pmol mol-1) for the models (NorESM2-LM and UKESM1-0-LL)

[114]:
```
#Get the median and SD for plotting error for UKESM model
_m = ukesm_isop.groupby('time.month').median()
_s = ukesm_isop.groupby('time.month').std()
median = _m.sel(lat=slat, lon=slon, method='nearest')
std = _s.sel(lat=slat, lon=slon,method='nearest')
#Get the median and SD for plotting error for NorESM model
_m1 =noresm_isop.groupby('time.month').median()
_s1 = noresm_isop.groupby('time.month').std()
median1 = _m1.sel(lat=slat, lon=slon, method='nearest')
std1 = _s1.sel(lat=slat, lon=slon,method='nearest')



fig, axs = plt.subplots(2, sharex=True, figsize=[15,10])
## Plot observation data from Birkenes station and model data.
station_name, slon, slat =get_station_coords('Birkenes II')
plot_model_ds_station(slat, slon, 1e12*ukesm_isop,'Isoprene, pmol/mol, UKESM',␣
 ↪ax=axs[0])
plot_model_ds_station(slat, slon, 1e12*noresm_isop,'Isoprene, pmol/mol,␣
 ↪NorESM', ax=axs[0])

plot_station_data(Birkenes1_station, 'Isoprene, pmol/mol, Birkenes', ax=axs[0])
plot_station_data(Birkenes2_station, 'Isoprene, pmol/mol, Birkenes II',␣
 ↪ax=axs[0])
axs[0].set_title('Station: Birkenes')
ax.fill_between(median['month'], median['isop']-std['isop'], median['isop']+␣
 ↪std['isop'], alpha=0.3, facecolor='g')
ax.fill_between(median1['month'], median1['isop']-std['isop'], median1['isop']+␣
 ↪std1['isop'], alpha=0.3, facecolor='g')

## Plot observation data from Zeppelin mountain station and model data.

station_name, slon, slat =get_station_coords('Zeppelin', unique=False)
plot_model_ds_station(slat, slon, 1e12*ukesm_isop,'Isoprene, pmol/mol, UKESM',␣
 ↪ax=axs[1])
plot_model_ds_station(slat, slon, 1e12*noresm_isop,'Isoprene, pmol/mol,␣
 ↪NorESM', ax=axs[1])
plot_station_data(Zeppelin_station, 'Isoprene, pmol/mol, Zeppelin', ax=axs[1])
axs[1].set_title('Station: Zeppelin')

for ax in axs:
    ax.set_ylabel('pmol/mol')
```

```
ax.legend()
```

```
found station name: Birkenes II
Returning (<station name>, <lon>, <lat>)
('Birkenes II', 8.252, 58.38853)
found station name: Zeppelin mountain (Ny-Ålesund)
Returning (<station name>, <lon>, <lat>)
('Zeppelin mountain (Ny-Ålesund)', 11.88668, 78.90715)
```

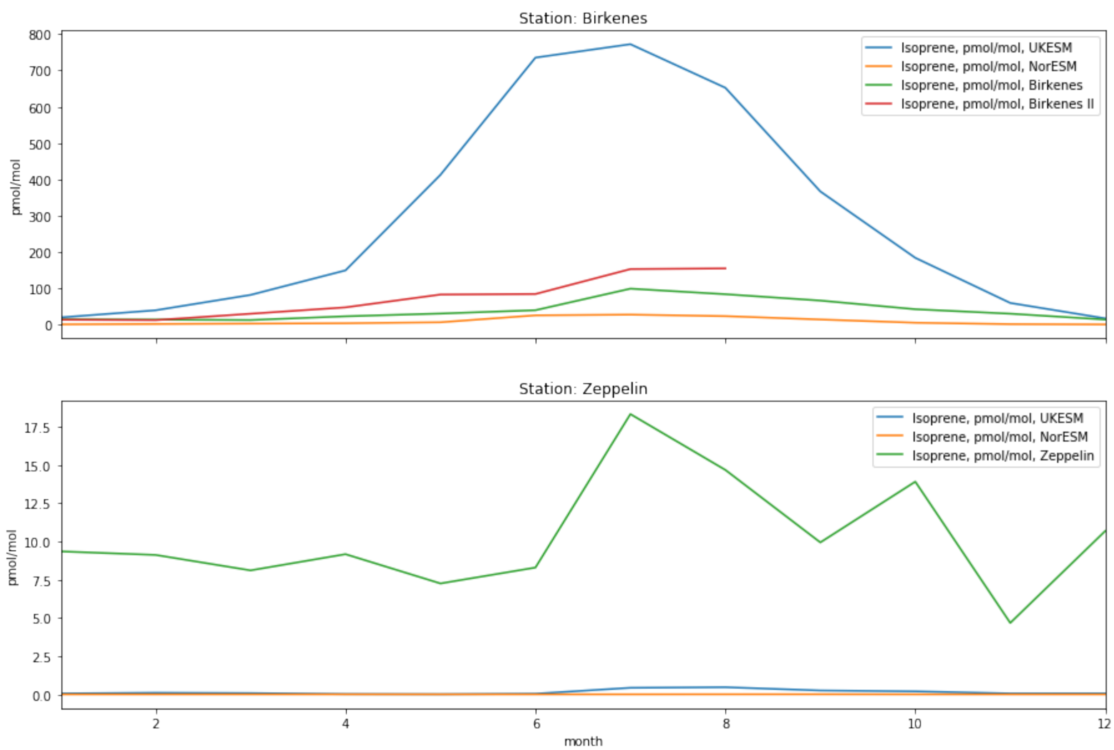


Figure 3 Monthly variations of average (1980-2014) grid-scale isoprene(pmol mol-1) for the models (NorESM2-LM and UKESM1-0-LL) and data from the observations. Upper graph: observation and model averages from Birkenes and Birkenes II stations; Bottom graph: observation and model averages from Zeppelin mountain (Ny-Ålesund). Model data are taken using 'nearest' function in that way obtaining the nearest available value to the observation collection site.

5 Discussion

Isoprene concentrations estimated by the NorESM2-LM model and UKESM1-0-LL showed substantial differences in both spatial and temporal variations (Fig. S1 and Fig.4). Overall NorESM2-LM had lower isoprene concentrations compare to UKESM1-0-LL model (Fig. 1). There was 1346.86% difference in isoprene concentrations between the models (Table 1). However, NorESM2-LM had higher isoprene concentrations in certain regions in summer (JJA) and autumn (SON): Krasnoyarsk region and Sakha Republic, which are Eurasian Taiga regions (Fig.1). These regions are associated with substantian sources of isoprene from Norway spruce, European aspen, willows and open wetland ecosystems (Rinne, Bäck, & Hakola, 2009). During spring period NorESM2-LM has higher concentrations UKESM1-0-LL in Taiga in North America (Fig. 1).

We used Total Emission Rate of Isoprene as an input variable in *pyaerocom* to be able to assess the relationship between two models. Covariance between two models was measured using Pearson's R, which was 0.662, indicating that the fit between two models is not perfect (Fig.2). A limitation of this study is that *pyaerocom* does not have isoprene concentrations, which would be helpful in validating the model and better understanding differences by assesing the relationship of each model against the observation data. Besides that investingating these relationships with other BVOC groups like monoterpenes, sesquiterpenes etc. should be tested to see whether with these compound groups we also see the same patterns.

Comparing isoprene concentrations temporarily, we saw a peak during the summer months, which agrees with previous studies. Interestingy, we found differences in isoprene observations between the stations, where Birkenes and Birkenes II stations had much lower isoprene concentrations compare to Zeppelin station (Fig.3). We also found that NorESM2-LM model data are better predicted for Birkenes and Birkenes II observation data, compare to UKESM1-0-LL which is overestimating the isoprene concentrations. For Zeppelin station, both models have much lower isoprene concentrations in comparison with observation data. We speculate, that the emissions from vegetation in this region are very low compare to transport of volatiles.

The substantial differences between the models could be because of thow these models convert fluxes into concentrations, which might be different for both of the models. Another factor, comparing the models, NorESM2-LM which used Guenther et al. (2016) algorithm for isoprene calculations, only used vegetation composition for calculations, whereas UKESM1-0-LL which used Arneth et al.(2007) for BVOC calculations used BVOC production as a function of photosynthesis, which could cause differences. Another difference between models is that NorESM2-LM has full energy balance, and thus been fed with proper canopy skin temperature, whereas UKESM1-0-LL used air tempeature at 2 m height and built simple energy balance to get 'sort of' canopy skin tempeture for BVOC model.

In general, models are not wery well representing the BVOC emissions in the Arctic (Tang et al., 2016). First, both models this study investigated, have limited data for charaterizing the emission factors and plant composition (Tang et al.,2016). Second, Tang et al.(2016) suggests that to improve BVOC emissions in the Arctic, in their study discuss that models should adjusted temperature response curve for Arctic plants with much stronger temperature sensitivity than the commonly used algorithms.

6 Conclusions

This study showed that NorESM2-LM and UKESM1-0-LL have substantial differences in predicted isoprene concentrations. Therefore understanding of how eaxactly each of the models are calculating isoprene concentrations and how parameters differ between both models are needed to make conclusions of which of the models is better. In addition, validating the model and better understanding differences by assesing the relationship of each model against the observation data are needed to fully confirm observed resutls. Nonetheless, investingating these relationships with other BVOC groups like monoterpenes, sesquiterpenes etc. should be tested to see whether with these compound groups we also see the same patterns. In general, theory suggests that both models do not represent very well the BVOC emissions in the Arctic due limited data for charaterizing the emission factor and plant composition, suggesting, that more studies and data are needed, for feeding the models with Arctic observations and making better predicions.

7 References

Bonn, B., Kuhlmann, R. von, & Lawrence, M. G. (2004). High contribution of biogenic hydroperoxides to secondary organic aerosol formation. Geophysical Research Letters, 31(10). https://doi.org/10.1029/2003GL019172

Dicke, M. (2009). Behavioural and community ecology of plants that cry for help. Plant Cell and Environment, 32(6), 654–665. https://doi.org/10.1111/j.1365-3040.2008.01913.x

Dicke, M., & Baldwin, I. T. (2010). The evolutionary context for herbivore-induced plant volatiles: Beyond the “cry for help.” Trends in Plant Science, 15(3), 167–175. https://doi.org/10.1016/j.tplants.2009.12.002

Goldstein, A. H., & Galbally, I. E. (2007). Known and unexplored organic constituents in the earth’s atmosphere. ACS Publications. Holopainen, J. K., Kivimäenpää, M., & Nizkorodov, S. A. (2017). Plant-derived Secondary Organic Material in the Air and Ecosystems. Trends in Plant Science, 22(9), 744–753. https://doi.org/10.1016/j.tplants.2017.07.004

Guenther, A. B., Jiang, X., Heald, C. L., Sakulyanontvittaya, T., Duhl, T., Emmons, L. K., & Wang, X. (2012). The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions. Geoscientific Model Development, 5(6), 1471–1492. https://doi.org/10.5194/gmd-5-1471-2012

Jiang, J., https://orcidorg/0000-0002-5356-5633, S. A., Ciarelli, G., Oikonomakis, E., El-Haddad, I., Canonaco, F., … Prévôt, A. S. H. (2019). Effects of two different biogenic emission models on modelled ozone and aerosol concentrations in Europe. Atmospheric Chemistry and Physics;

Katlenburg-Lindau, 19(6), 3747–3768. http://dx.doi.org/10.5194/acp-19-3747-2019 Kesselmeier, J., & Staudt, M. (1999). Biogenic volatile organic compounds (VOC): An overview on emission, physiology and ecology. Journal of Atmospheric Chemistry, 33(1), 23–88.

Kessler, A., & Baldwin, I. T. (2001). Defensive function of herbivore-induced plant volatile emissions in nature. Science, 291(5511), 2141–2144. https://doi.org/10.1126/science.291.5511.2141

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., … Corlay, S. (2016). Jupyter development team. Jupyter notebooks. a publishing format for reproducible computational workflows. In

Kulmala, M., Nieminen, T., Chellapermal, R., Makkonen, R., Bäck, J., & Kerminen, V.-M. (2013). Climate Feedbacks Linking the Increasing Atmospheric CO2 Concentration, BVOC Emissions, Aerosols and Clouds in Forest Ecosystems. In Ü. Niinemets & R. K. Monson (Eds.), Biology, Controls and Models of Tree Volatile Organic Compound Emissions (pp. 489–508). https://doi.org/10.1007/978-94-007-6606-8_17

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., … Vuuren, D. P. van. (2010). Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application. Atmospheric Chemistry and Physics, 10(15), 7017–7039. https://doi.org/10.5194/acp-10-7017-2010

Peñuelas, J., & Staudt, M. (2010). BVOCs and global change. Trends in Plant Science, 15(3), 133–144.

Rinne, J., Bäck, J., & Hakola, H. (2009). Biogenic volatile organic compound emissions from the Eurasian taiga: Current knowledge and future directions.

Scala, A., Allmann, S., Mirabella, R., Haring, M. A., & Schuurink, R. C. (2013). Green Leaf Volatiles: A Plant's Multifunctional Weapon against Herbivores and Pathogens. International Journal of Molecular Sciences, 14(9), 17781–17811. https://doi.org/10.3390/ijms140917781

Tang, J., Schurgers, G., Valolahti, H., Faubert, P., Tiiva, P., Michelsen, A., & Rinnan, R. (2016). Challenges in modelling isoprene and monoterpene emission dynamics of Arctic plants: A case study from a subarctic tundra heath. Biogeosciences, 13(24), 6651–6667. https://doi.org/10.5194/bg-13-6651-2016

8 Acknowledgments

I acknowledge the World Climate Research Program's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

I acknowledge EBAS http://ebas.nilu.no/ (database infrastructureoperated by NILU – Norwegian Institute for Air Research) responsible of open access of the station data used in this study.

Many thanks to Sara and Anne for amazing practical help with Jupyter Notebook <3

9 Supplementary information

Import packages

[2]:
```
from pydap.client import open_dods, open_url
from netCDF4 import num2date
import pandas as pd
import xarray as xr
import numpy as np
import matplotlib.pyplot as plt
import matplotlib as mpl
mpl.rcParams.update(mpl.rcParamsDefault)
%matplotlib inline
import cartopy.crs as ccrs
import cartopy as cy
import matplotlib.path as mpath
from pydap.client import open_dods, open_url
from netCDF4 import num2date
import pandas as pd
import sys
import glob
import pyaerocom as pya
import matplotlib.pyplot as plt
from warnings import filterwarnings
filterwarnings('ignore')
pya.change_verbosity('critical', log=pya.const.print_log) # don't output␣
 ↪warnings
```

```
pya.__version__
```

```
/opt/conda/lib/python3.7/site-packages/geonum/__init__.py:39: UserWarning:
Plotting of maps etc. is deactivated, please install Basemap
  warn('Plotting of maps etc. is deactivated, please install Basemap')
```

[2]: '0.8.1.dev4'

9.0.1 Get CMIP6 data from NorESM and UKESM models

Isoprene Volume Mixing Ratio (mol mol-1)

[ ]:
```
#This function selects the isoprence concentrations from NorESM2-LM. Due to␣
↪computing power, the selected years are from
#1980-2015. The region of interest is Arctic, therefore we slice that data by␣
↪latitudes (60:90). Due to high
#reactivity of the BVOCsThe level of interest is level=0.
path = '/home/bd15084e-2d25b7-2d47db-2dade3-2daea695ce03d8/
↪shared-cmip6-for-ns1000k/historical/NorESM2-LM/r1i1p1f1/'
var = 'isop'
from_y = '1980-01-01'
to_y = '2015-01-01'
files = glob.glob(path+var+'*')
files.sort()
noresm_isoprene = xr.open_mfdataset(files, combine='nested', concat_dim = 'time'
                          ).sel(time=slice(from_y, to_y))
noresm_isoprene=noresm_isoprene.sel(lat=slice(60,90)).isel(lev=0)
noresm_isoprene.to_netcdf('noresm_isoprenelev0.nc')
noresm_isoprene.close()
```

[ ]:
```
#This function selects the isoprence concentrations from UKESM1-0-LL. Due to␣
↪computing power, the selected years are from
#1980-2015. The region of interest is Arctic, therefore we slice that data by␣
↪latitudes (60:90). Due to high
#reactivity of the BVOCsThe level of interest is level=0.
path = '/home/bd15084e-2d25b7-2d47db-2dade3-2daea695ce03d8/
↪shared-cmip6-for-ns1000k/historical/UKESM1-0-LL/r1i1p1f1/'
var = 'isop'
from_y = '1980-01-01'
to_y = '2015-01-01'
files = glob.glob(path+var+'*')
files.sort()
ukesm_isoprene = xr.open_mfdataset(files, combine='nested', concat_dim = 'time'
                          ).sel(time=slice(from_y, to_y))
ukesm_isoprene=ukesm_isoprene.sel(lat=slice(60,90)).isel(lev=0)
ukesm_isoprene.to_netcdf('ukesm_isoprenelev0.nc')
ukesm_isoprene.close()
```

Total Emission Rate of Isoprene (kg m-2 s-1)

[ ]:
```
#This function selects the Total Emission Rate of Isoprene (kg m-2 s-1) from␣
↪NorESM2-LM. Due to computing power, the slected years are from
#1980-2015. The region of interest is Arctic, therefore we slice that data by␣
↪latitudes (60:90). Due to high
#reactivity of the BVOCsThe level of interest is level=0.
path = '/home/bd15084e-2d25b7-2d47db-2dade3-2daea695ce03d8/
↪shared-cmip6-for-ns1000k/historical/NorESM2-LM/r1i1p1f1/'
var = 'emiisop'
from_y = '1980-01-01'
to_y = '2015-01-01'
files = glob.glob(path+var+'*')
files.sort()

noresm_is = xr.open_mfdataset(files, combine='nested', concat_dim = 'time'
                          ).sel(time=slice(from_y, to_y))
noresm_isop=noresm_is.sel(lat=slice(60,90)).isel(lev=0)

noresm_isop.to_netcdf('noresm_emis1.nc')
noresm_isop.close()
```

[ ]:
```
#This function selects the Total Emission Rate of Isoprene (kg m-2 s-1) from␣
↪UKESM1-0-LL. Due to computing power, the slected years are from
#1980-2015. The region of interest is Arctic, therefore we slice that data by␣
↪latitudes (60:90). Due to high
#reactivity of the BVOCsThe level of interest is level=0.
var = 'emiisop'
from_y = '1980-01-01'
to_y = '2015-01-01'
ukesm = xr.open_mfdataset(files, combine='nested', concat_dim = 'time'
                          ).sel(time=slice(from_y, to_y))
path = '/home/bd15084e-2d25b7-2d47db-2dade3-2daea695ce03d8/
↪shared-cmip6-for-ns1000k/historical/UKESM1-0-LL/r1i1p1f2/'
files = glob.glob(path+var+'*')#" get_fl('emibvoc',path )
files.sort()
ukesm_is = xr.open_mfdataset(files, combine='nested', concat_dim = 'time'
                         ).sel(time=slice(from_y, to_y))
ukesm_isop=ukesm_is.sel(lat=slice(60,90)).isel(lev=0)
ukesm_isop.close()
```

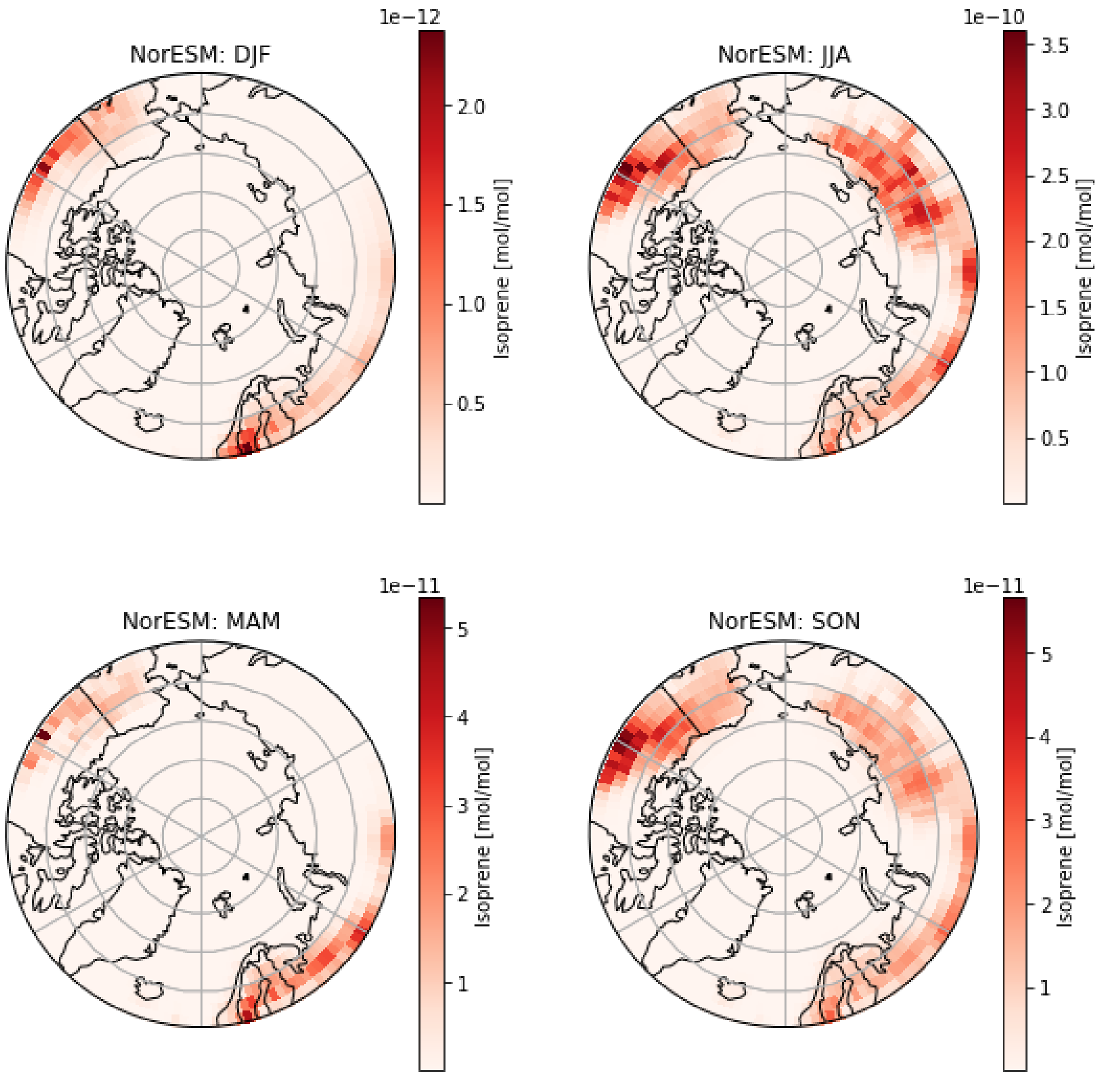
NorESM (upper) and UKESM (bottom) model Isoprene Volume Mixing Ratio (mol mol-1)

[102]:
```
#Group data by the season and obtain the mean from (1980-2014)
ds_seas = noresm_isop.groupby('time.season').mean()
# Make the axis with NorthPolarStereo() projection:
fig, axs = sp_map(2,2, projection=ccrs.NorthPolarStereo(), figsize=[10,10] )
lat_lims = [60,90]
```

```
season_l = ds_seas['season'].values
print(season_l)
for ax, seas in zip(axs.flatten(),season_l):
    _ds = ds_seas.sel(season=seas)['isop'].where(ds_seas['lat']>lat_lims[0])#
    _ds.attrs['units']='mol/mol'; _ds.attrs['long_name'] = 'Isoprene'
    _ds.plot(ax=ax, cmap=plt.get_cmap('Reds'),transform=ccrs.PlateCarree())
    ax.set_title('NorESM: '+ seas )
    polarCentral_set_latlim(lat_lims, ax)
    add_map_features(ax)

plt.show()
#Group data by the season and obtain the mean from (1980-2014)
ds_seas = ukesm_isop.groupby('time.season').mean()
fig, axs = sp_map(2,2, projection=ccrs.NorthPolarStereo(), figsize=[10,10] )
lat_lims = [60,90]
season_l = ds_seas['season'].values
print(season_l)
for ax, seas in zip(axs.flatten(),season_l):
    _ds = ds_seas.sel(season=seas)['isop'].where(ds_seas['lat']>lat_lims[0])#
    _ds.attrs['units']='mol/mol'; _ds.attrs['long_name'] = 'Isoprene'
    _ds.plot(ax=ax, cmap=plt.get_cmap('Reds'),transform=ccrs.PlateCarree())
    ax.set_title('UKESM: '+ seas)
    polarCentral_set_latlim(lat_lims, ax)
    add_map_features(ax)
plt.show()
```

```
['DJF' 'JJA' 'MAM' 'SON']
```




```
['DJF' 'JJA' 'MAM' 'SON']
```

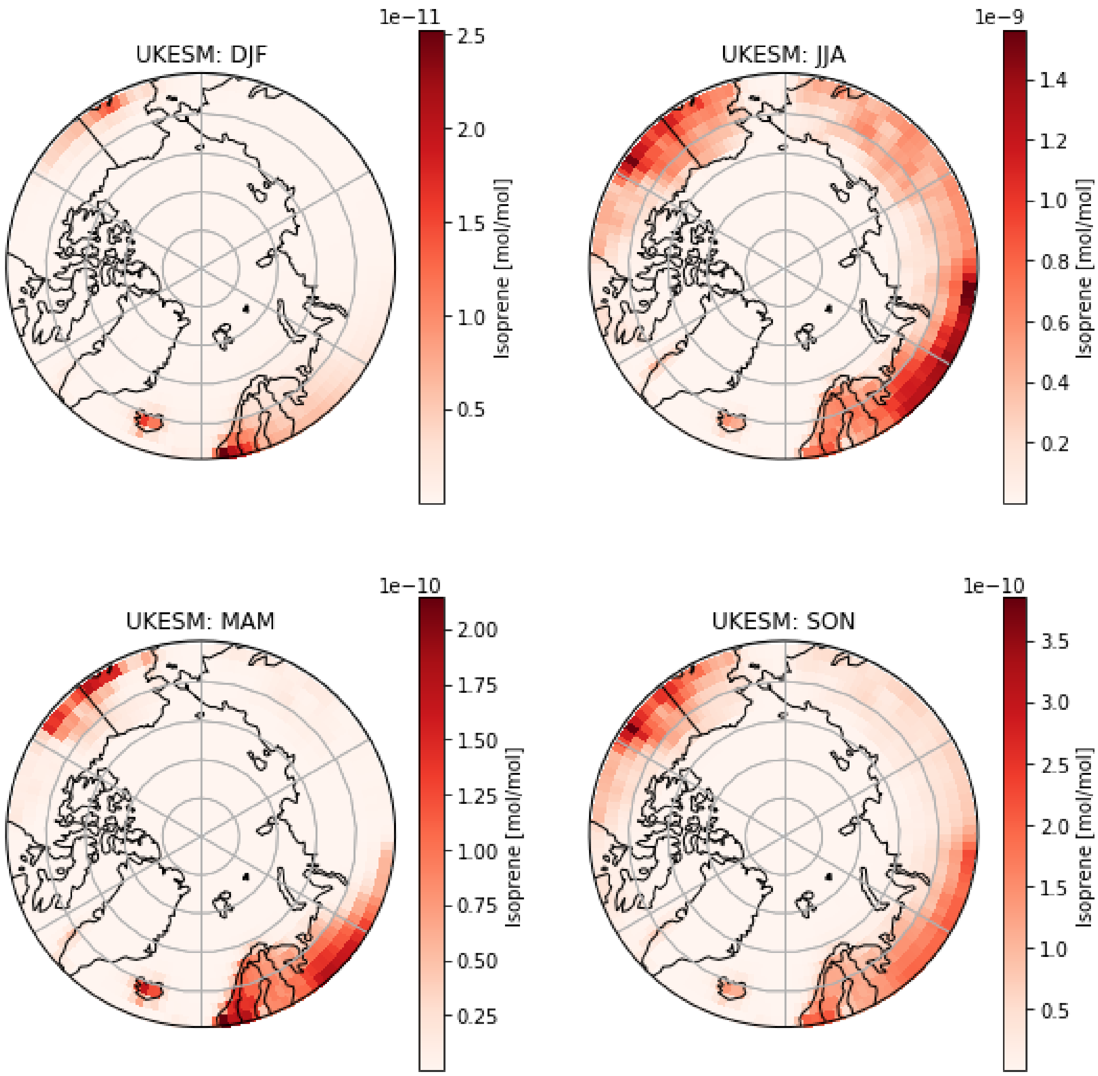


Figure S1. Seasonal average (1980-2014) isoprene concentrations for each of the models (upper graph - NorESM and bottom graph - UKESM). **DJF**= December, January, February; **MAM** = March, April, May; **JJA** = June, July, August and **SON**= September, Oktober, November.

9.0.2 *Pyaerocom* for collocation of the data and obtaining statistics for model comparison

[3]:
```
#Select the NorESM model data
CMIP6_TEST_DIR = '/home/bd15084e-2d25b7-2d47db-2dade3-2daea695ce03d8/'
CMIP6_TEST_FILE = 'noresm_emis1.nc'
path = CMIP6_TEST_DIR + CMIP6_TEST_FILE
modeldata_noresm_conc = pya.GriddedData(path, var_name='emiisop')
```

[4]:
```
#Select the UKESM model data
CMIP6_TEST_DIR = '/home/bd15084e-2d25b7-2d47db-2dade3-2daea695ce03d8/'
CMIP6_TEST_FILE = 'ukesm_emiisop_4.nc'
```

```
path = CMIP6_TEST_DIR + CMIP6_TEST_FILE
modeldata_ukesm_conc = pya.GriddedData(path, var_name='emiisop')
```

[5]:
```
modeldata_noresm_conc.metadata['ts_type'] = 'monthly' # models ts_type monthly,␣
↪will be used to collocate data
modeldata_noresm_conc.ts_type
```

[5]:
```
'monthly'
```

[6]:
```
modeldata_ukesm_conc.metadata['ts_type'] = 'monthly' # models ts_type monthly,␣
↪will be used to collocate data
modeldata_ukesm_conc.ts_type
```

[6]:
```
'monthly'
```

[7]:
```
modeldata_noresm_conc.metadata['data_id']='NorESM' #give names to axis
modeldata_ukesm_conc.metadata['data_id']='UKESM' #give names to axis
```

[8]:
```
modeldata_noresm_conc.start,modeldata_noresm_conc.stop
```

[8]:
```
(numpy.datetime64('1980-01-01T00:00:00.000000'),
 numpy.datetime64('2014-12-31T23:59:59.999999'))
```

[9]:
```
modeldata_ukesm_conc.start,modeldata_ukesm_conc.stop
```

[9]:
```
(numpy.datetime64('1980-01-01T00:00:00.000000'),
 numpy.datetime64('2014-12-31T23:59:59.999999'))
```

[10]:
```
#collocate the both model data:
try:
    coldata1 = pya.colocation.colocate_gridded_gridded(modeldata_noresm_conc,
                                                       modeldata_ukesm_conc,
                                                       start=1850,
                                                       ts_type='monthly')
    stats = coldata1.calc_statistics()
except Exception as e:
    print('Colocating failed. Reason: {}'.format(repr(e)))
```

[11]:
```
#set x and y axis as min-max from the models
_n= coldata1.data.sel(data_source='NorESM')
_u= coldata1.data.sel(data_source='UKESM')
xlim=[np.nanmin(_n), np.nanmax(_n)]
ylim=[np.nanmin(_u), np.nanmax(_u)]
```

[ ]:
```
#plot the covariance of the two models
coldata1.plot_scatter(marker='o', color='blue', alpha=0.1)#, xlabel='blabla');
plt.xlim(xlim)
plt.ylim(ylim)
```