EGU General Assembly

Towards a new surrogate model for predicting short term NO_x-O_3 effects from aviation using Gaussian processes

Pratik Rao, Richard Dwight, Deepali Singh, Jin Maruhashi, Irene Dedoussi, Volker Grewe, and Christine Frömming

23^{rd} - 28^{th} April 2023

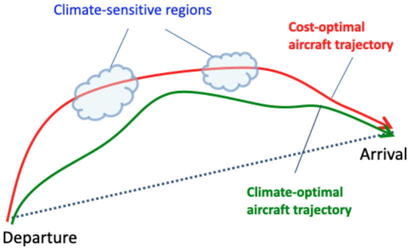
This presentation participates in OSPP
Outstanding Student & PhD candidate Presentation contest

Sharing is encouraged

Why?

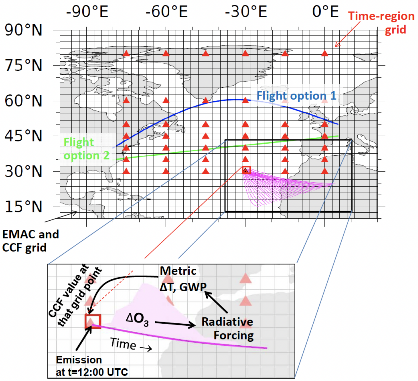
Non-CO_2 aviation effects contribute to $\sim$ **2/3** of the climate impact and are characterised by **high uncertainties** [1]

Since the impact depends strongly on emission location, what if we could get flights to avoid **climate sensitive regions**?



Climate change functions (CCFs)

- CCFs [2] → global climate impact due to emission at $(\boldsymbol{x}, t)$ → expensive and restrictive



Surrogate model (aCCFs)

Surrogate model (aCCFs)

- Reproduce CCF predictions by other means?
- Linearly regress CCFs against atmospheric variables $\rightarrow$ aCCFs:

$$\mathrm{aCCF}_{\mathrm{O}_3} = \theta^{\mathrm{T}}\mathbf{w}, \;\; \theta = < \mathrm{T}, \phi, \mathrm{T}\phi >$$

Surrogate model (aCCFs)

Reproduce CCF predictions by other means?

Linearly regress CCFs against atmospheric variables → aCCFs:

$$\mathrm{aCCF}_{\mathrm{O}_3} = \theta^{\mathrm{T}}\mathbf{w}, \quad \theta = < \mathrm{T}, \phi, \mathrm{T}\phi >$$

Regional flight planning on arbitrary days

Reasonable first estimate [3] but improvements are desirable

Towards a new surrogate model

Expand geographical coverage of training data [4]

↓

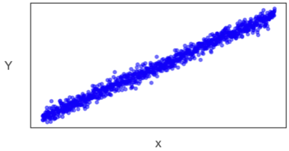
Objective approach to feature selection

↓

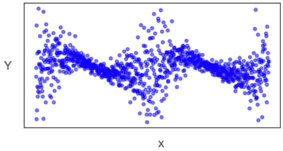
Gaussian process regression model

↙ ↘

Standard [5] | Chained [6]

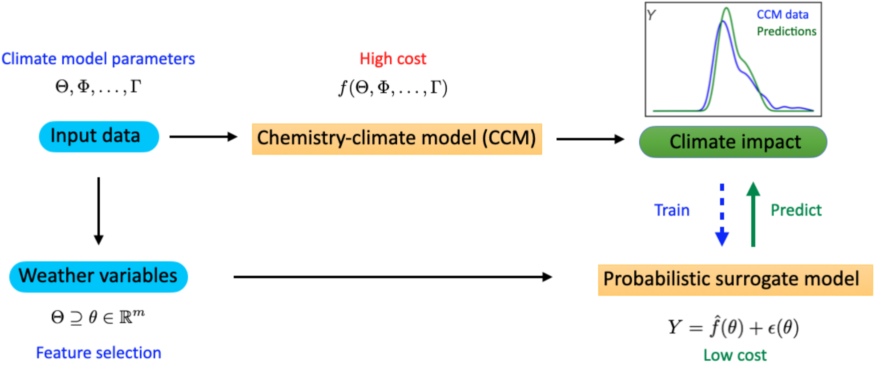


Constant variance



Variance is a function of x

Towards a new surrogate model



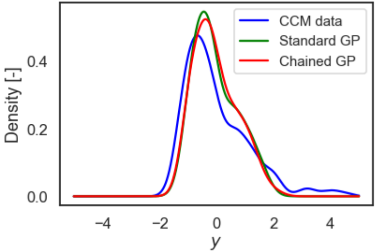
Gaussian process regression is a Bayesian nonparametric approach that can capture more information about $\mathcal{D} = \{\theta, y\}$ with error bars as data grows: $p(\hat{f}|\mathcal{D}) \propto p(\hat{f})\, p(\mathcal{D}|\hat{f})$

Results

- Feature selection yields $\theta \in \mathbb{R}^6$: Temperature, geopotential, solar irradiance, specific humidity, zonal velocity, and release location

Results

- Feature selection yields $\theta \in \mathbb{R}^6$: Temperature, geopotential, solar irradiance, specific humidity, zonal velocity, and release location
- Full distribution for (predicted) climate impact Y on test space

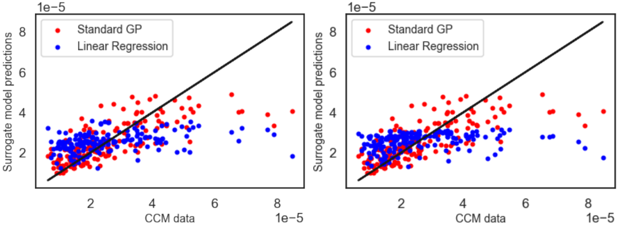


Results

- Performs significantly better than Linear regression ($R^2 = 0.54$)
- Linear Regression model: Using selected features ($R^2 = 0.13$) vs original features ($R^2 = 0.05$)

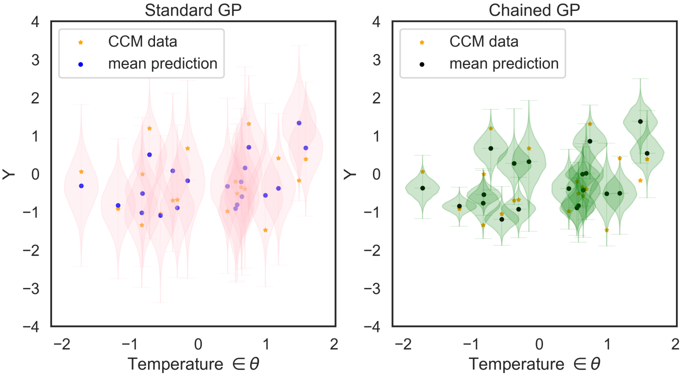
Results

- Performs significantly better than Linear regression ($R^2 = 0.54$)
- Linear Regression model: Using selected features ($R^2 = 0.13$) vs original features ($R^2 = 0.05$)

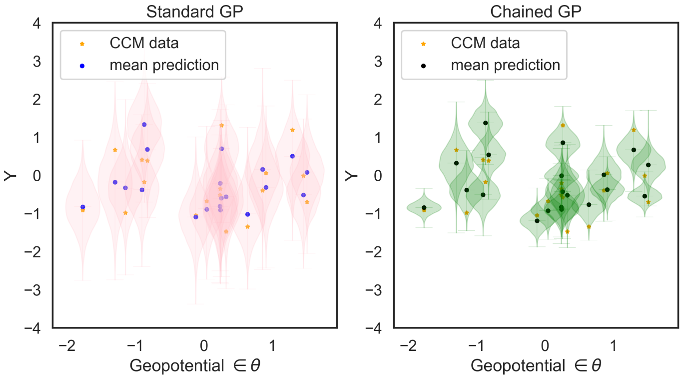


Results

- Comparing test data and predictions $\forall\ \theta$
- Violin plot shows variance of every prediction in the test space

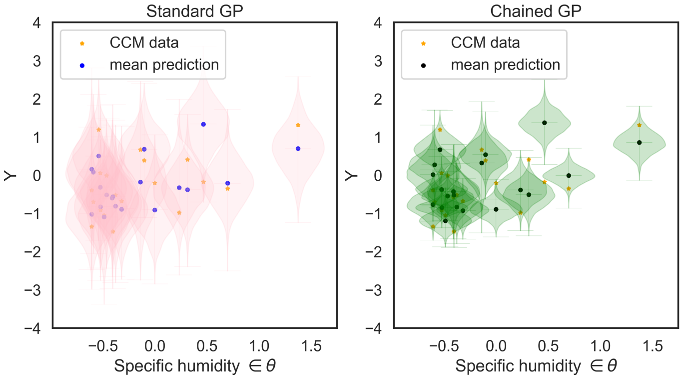


Results

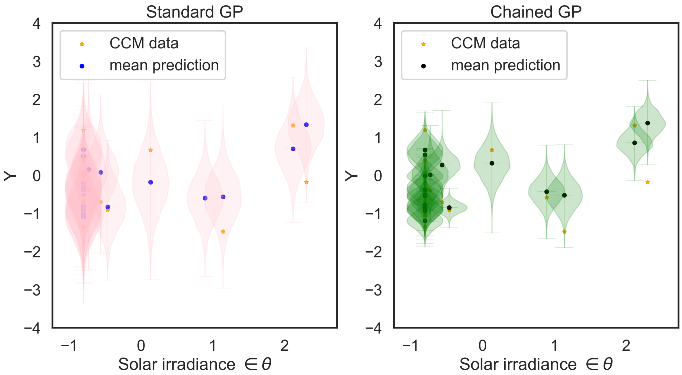


Standard GP
Chained GP
CCM data
mean prediction
CCM data
mean prediction
Y
Y
Geopotential $\in \theta$
Geopotential $\in \theta$
4
3
2
1
0
−1
−2
−3
−4
−2
−1
0
1

Results

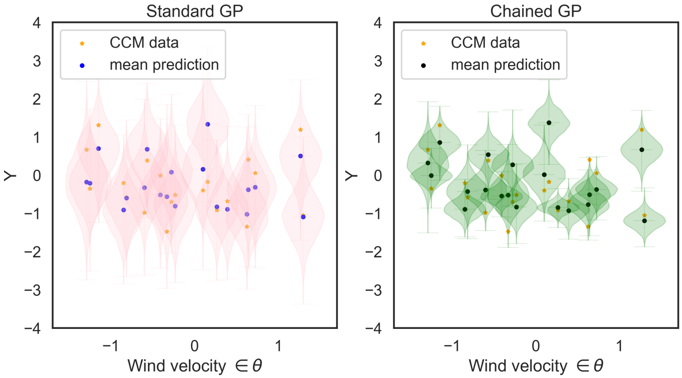

Standard GP
CCM data
mean prediction
4
3
2
1
0
−1
−2
−3
−4
Y
−0.5
0.0
0.5
1.0
1.5
Specific humidity $\in \theta$
Chained GP
CCM data
mean prediction
4
3
2
1
0
−1
−2
−3
−4
Y
−0.5
0.0
0.5
1.0
1.5
Specific humidity $\in \theta$

Results



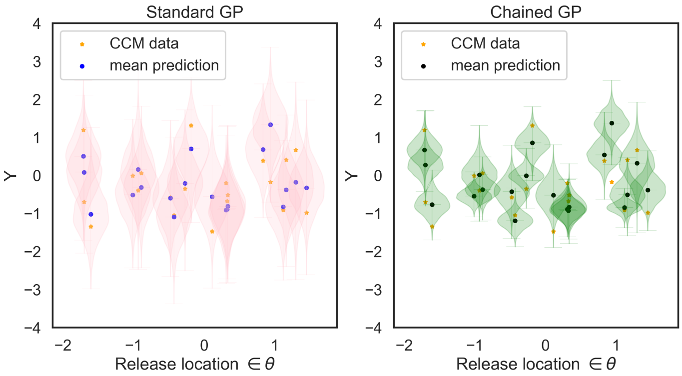
Standard GP
CCM data
mean prediction
4
3
2
1
0
−1
−2
−3
−4
Y
−1
0
1
2
Solar irradiance $\in \theta$
Chained GP
CCM data
mean prediction
4
3
2
1
0
−1
−2
−3
−4
Y
−1
0
1
2
Solar irradiance $\in \theta$

Results



Standard GP
CCM data
mean prediction
Y
Wind velocity $\in \theta$
Chained GP
CCM data
mean prediction
Y
Wind velocity $\in \theta$

Results

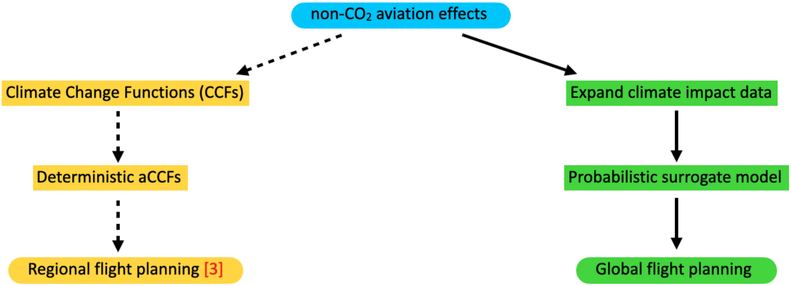


Standard GP
CCM data
mean prediction
Y
Release location $\in \theta$
Chained GP
CCM data
mean prediction
Y
Release location $\in \theta$

Summary

- $R^2 > 0.50$ for GP models, while D_{kl} is lower for chained GP model
- Analyse statistical 'outliers'
- We have a model that predicts the climate impact of aviation NO_x with (varying) confidence levels

Take away?



Improved understanding of climate effects of local aviation NO_x emissions

References

[1] Lee et al., 2020. The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018.

[2] Grewe et al., 2014. Aircraft routing with minimal climate impact: the REACT4C climate cost function modelling approach (V1.0).

[3] Rao et al., 2022. Case Study for Testing the Validity of NO_x-Ozone aCCFs for Optimising Flight Trajectories.

[4] Maruhashi et al., 2022. Transport Patterns of Global Aviation NO_x and their Short-term O_3 Radiative Forcing A Machine Learning Approach.

[5] Rasmussen & Williams, 2006. Gaussian processes for machine learning.

[6] Saul et al., 2016. Chained Gaussian processes.

Thank you for your kind attention!

QR code for abstract