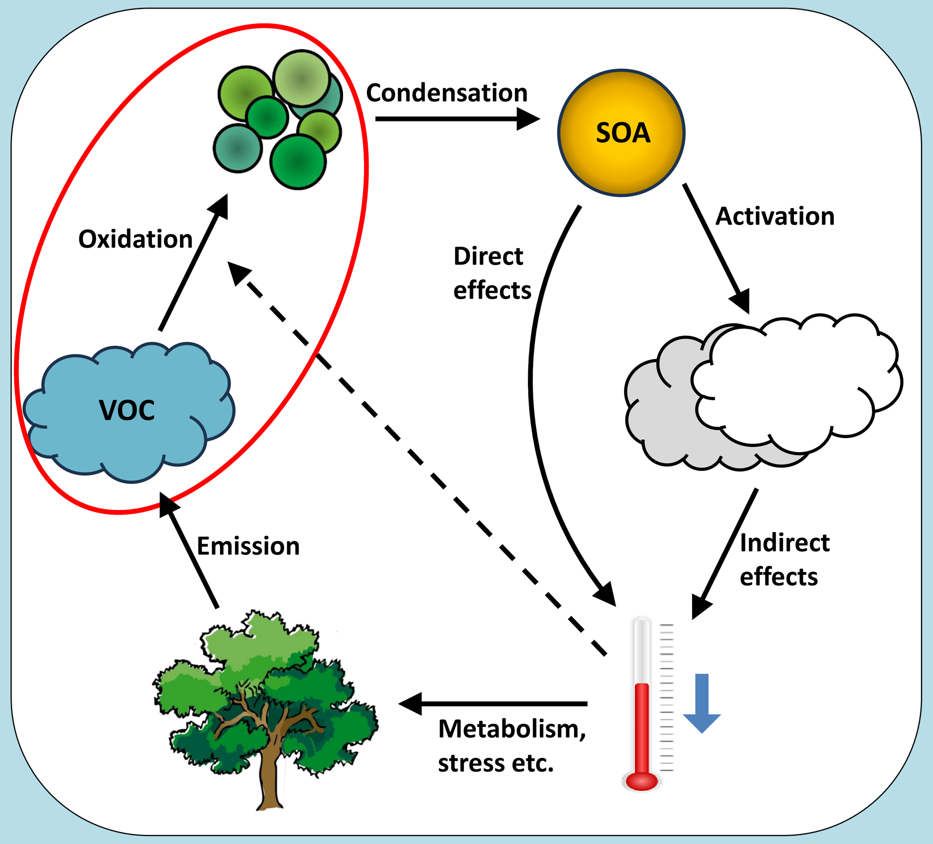
Improving the monoterpene oxidation scheme in a global-scale model through neural network-based bias correction

Antti Vartiainen[1,2] // Pontus Roldin[3] // Muhammed Irfan[1] // August Thomasson[3] // Harri Kokkola[1,4] // Taina Yli-Juuti[1]

[1]University of Eastern Finland, Department of Technical Physics, Kuopio, Finland // [2]CSC – IT Center for Science Ltd., Espoo, Finland // [3]Lund University, Department of Physics, Lund, Sweden // [4]Finnish Meteorological Institute, Kuopio, Finland

1. Background and motivation

- Various volatile organic compounds (**VOC**) are emitted by vegetation. These emissions are **temperature-dependent.**
- VOCs undergo oxidation, forming lower-volatility products that may condense into secondary organic aerosols (SOA).
    - The **volatility** of the products, and thus their ability to form SOA, is affected by **oxidation conditions.**
- Monoterpenes such as **α-pinene** among the main VOCs influencing SOA formation.
- SOA generally induces a negative radiative forcing through scattering and aerosol-cloud interactions → a cooling effect on climate.
    - This leads to a **negative climate feedback** due to further changes in VOC emissions. [1]
- Representing oxidation chemistry accurately in climate models is critical to capture the feedback effect. However, global-scale models heavily simplify chemical processes due to computational costs.

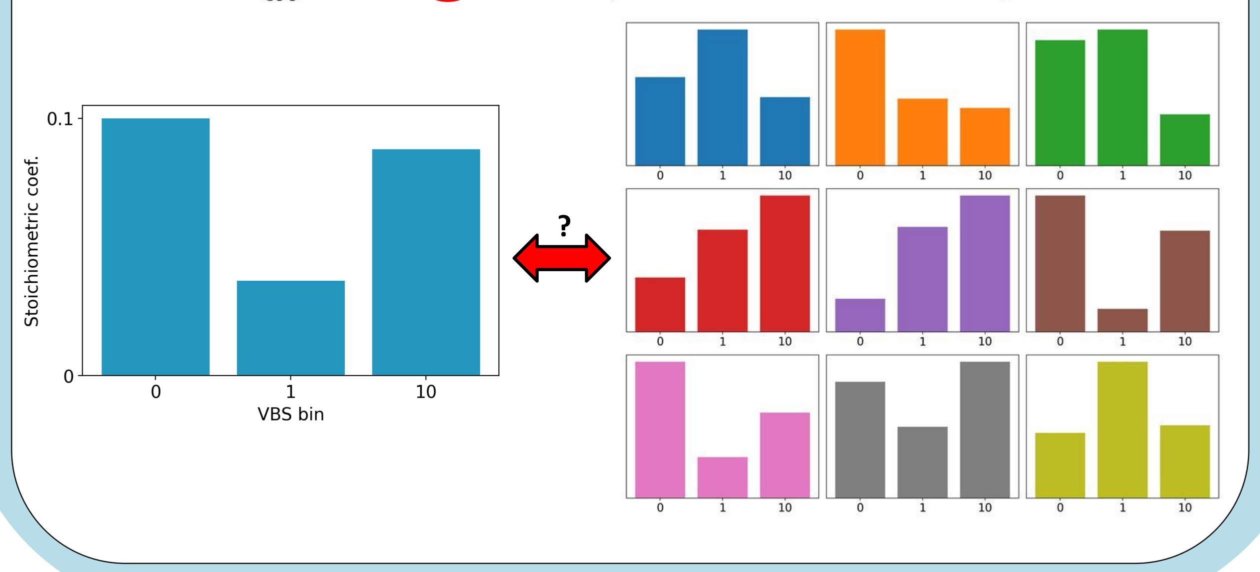


2. Models and objective

- In the **SALSA** aerosol module for climate models [2], volatility of oxidation products is represented by the volatility basis set (**VBS**) → **3 volatility bins** based on saturation concentration (e.g., bin 10: 10 μg/m³)
    - A fixed fraction ("stoichiometric coefficient") of reacted α-pinene is transferred into each bin, independent of factors like temperature, NO_x, or the main oxidant.
- **ADCHAM**, a state-of-the-art chamber chemistry model, simulates the oxidation process using thousands of reactions and compounds [3,4]. However, it is too heavy to implement in global-scale models.

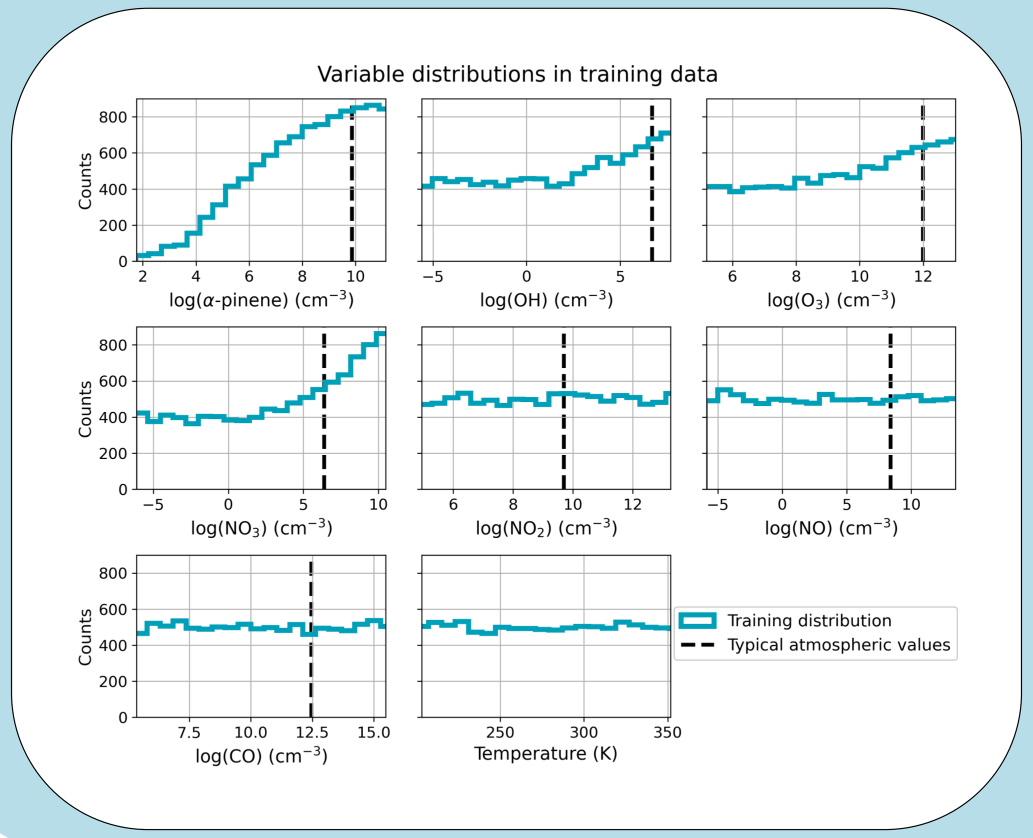
→ We have trained a simple **neural network (NN)** on ADCHAM simulations **to correct the VBS distributions produced by SALSA.**

$$\frac{d[VBS]_i}{dt} = a_i [VOC] \left(\sum_{n=1}^{3} k_n(T) [Oxid]_n \right)$$



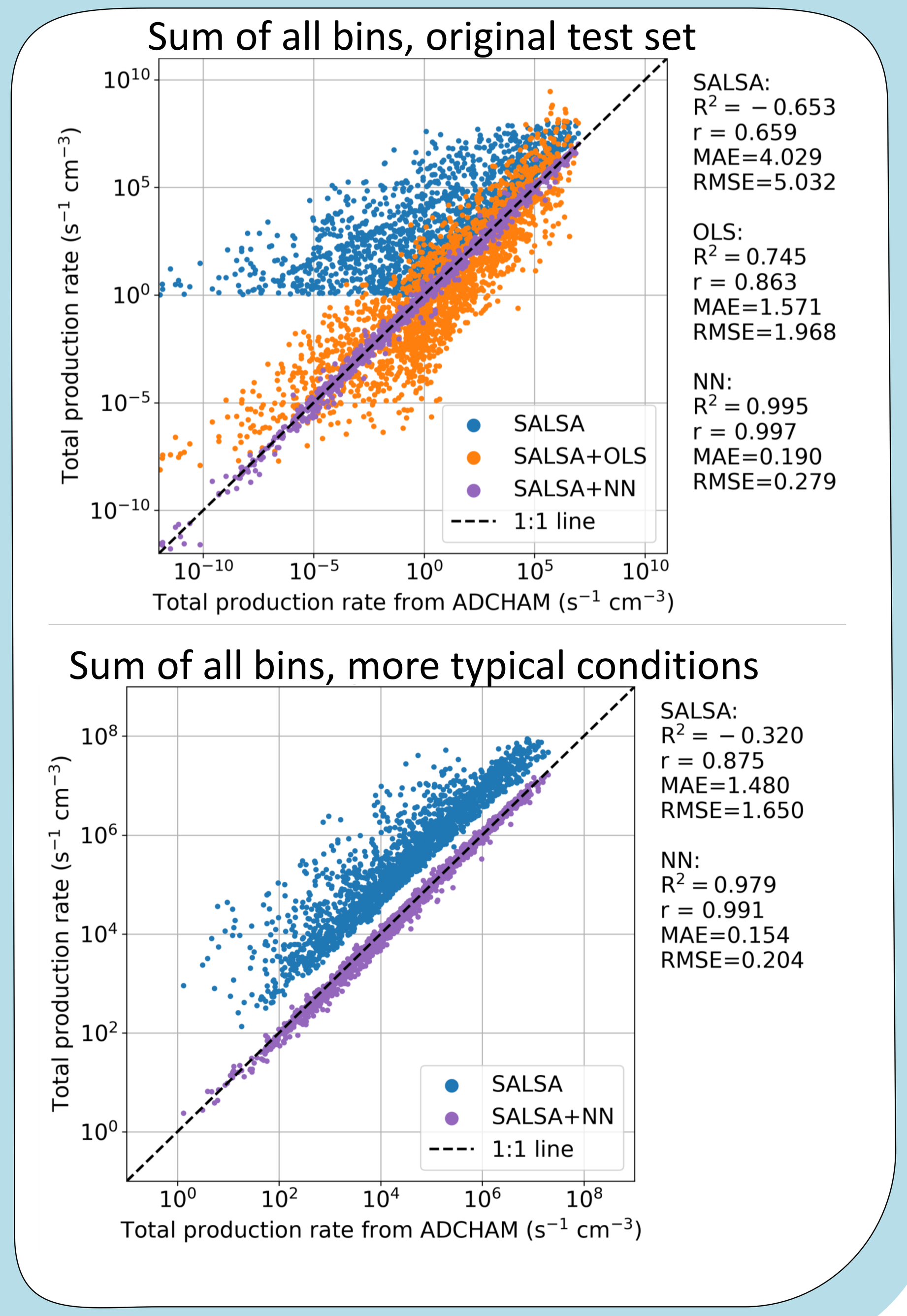
3. Features and targets

- The NN was trained on the bias between the models based on oxidation conditions.
- Differences in **production rates** (1/cm³s) of VBS bins from a 7.5-minute simulation **used as targets** of the NN:
  $\log(VBS_{ADCHAM}) - \log(VBS_{SALSA})$
- **Eight input variables** selected based on an analysis of ADCHAM: α-pinene, O_3, OH, NO_3, CO, NO, NO_2, T
- 10 000 cases sampled from atmospherically realistic ranges of input variables → 80% training, 20% testing
    - Cases with very small SALSA production rates replaced by new samples.



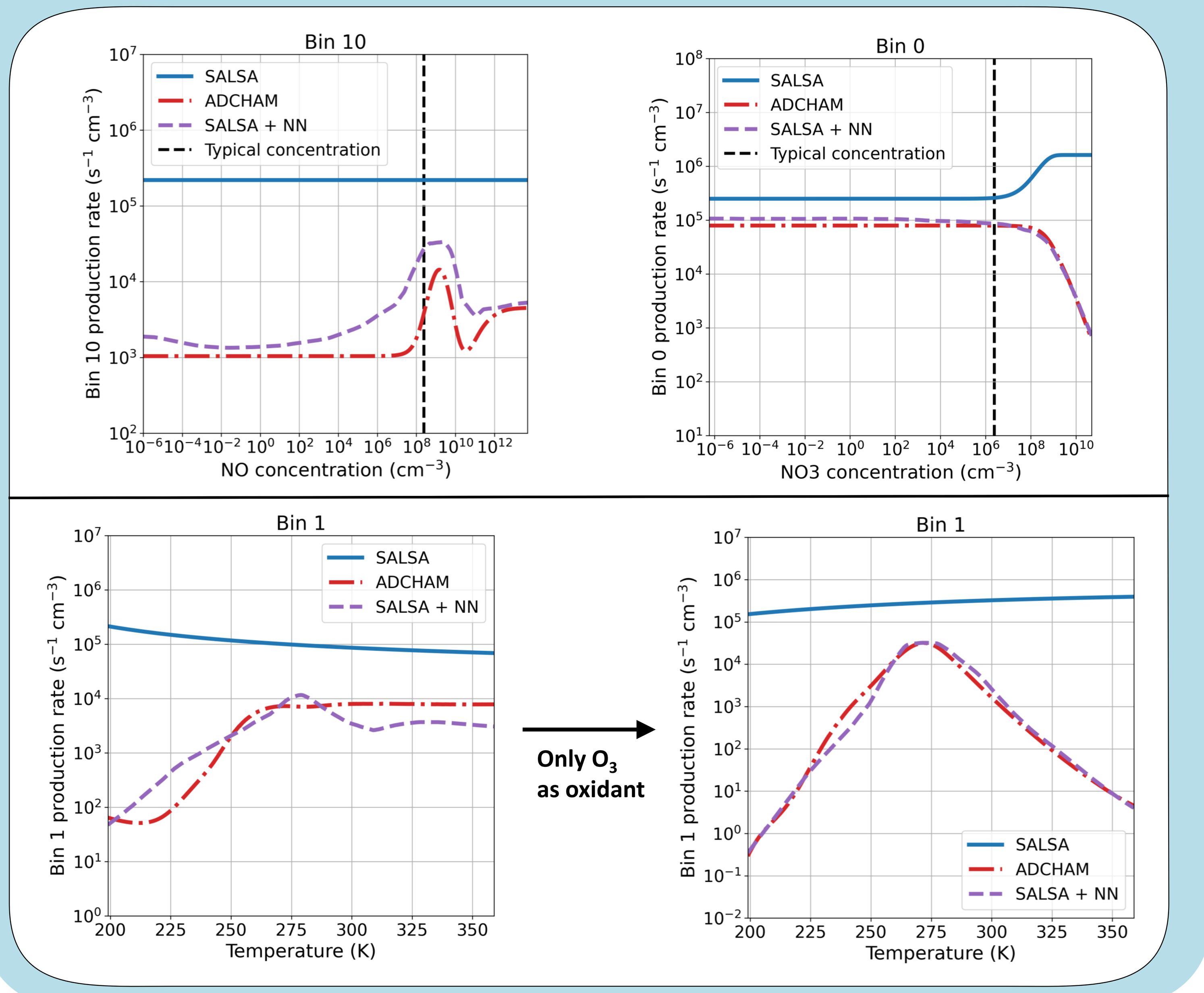
4. Results of bias correction

- **SALSA overestimates** production rates for all VBS bins, in some cases by more than 10 orders of magnitude.
- A linear model (OLS) can already correct most of the systematic overestimation, but high error variance remains.
- The **NN-corrected production rates correspond extremely well** to the ADCHAM results.
    - No major differences in performance between VBS bins; predictions for bin 0 slightly more accurate, likely due to a smaller range of production rates compared to the other two bins.
- Due to the wide sampling ranges of the original test set, another test was conducted on cases sampled from more typical atmospheric conditions → metrics remain high for all bins.



5. Dependence analysis

- The dependence of production rates on input variable values was compared across models by varying the values of one input, keeping all others fixed.
- Generally, the dependences in ADCHAM were much more complex compared to SALSA.
- The NN can emulate the general behavior of ADCHAM even in complex cases; when the dependence is more consistent, almost perfect correspondence is achieved.
- The selection of the fixed values affects the dependence significantly.



6. Conclusions

- The production of condensable organics in SALSA and ADCHAM differs significantly, suggesting that monoterpene oxidation chemistry is more complex than the current representations in global-scale models can capture.
- Correcting simplified chemistry with neural networks can yield performance similar to state-of-the-art chemistry models. Linear regression is not enough to achieve this.
- The study demonstrates a promising approach for representing the effect of varying atmospheric conditions on SOA formation in climate models.
    - We will further evaluate the hybrid SALSA+NN in terms of SOA mass yields produced, and ultimately plan to implement it in a climate model.

References

[1] Yli-Juuti *et al.* (2021). Significance of the organic aerosol driven climate feedback in the boreal area. *Nat. Commun.*, **12**, 5637

[2] Kokkola *et al*. (2018). SALSA2.0: The sectional aerosol module of the aerosol–chemistry–climate model ECHAM6.3.0-HAM2.3-MOZ1.0. *Geosci. Model Dev.*, **11**, 3833–3863.

[3] Roldin *et al.* (2014). Modelling non-equilibrium secondary organic aerosol formation and evaporation with the aerosol dynamics, gas- and particle-phase chemistry kinetic multilayer model ADCHAM. *Atmos. Chem. Phys.*, **14**, 7953–7993.

[4] Roldin *et al.* (2019). The role of highly oxygenated organic molecules in the Boreal aerosol-cloud climate system. *Nat. Commun.*, **10**, 4370.

Link to the abstract:

CONTACT INFORMATION
antti.vartiainen@uef.fi
UNIVERSITY OF EASTERN FINLAND | AEROSOL PHYSICS