Drivers of fog and low stratus - a satellite-based evaluation with machine learning

Eva Pauli, Hendrik Andersen, Jörg Bendix, Jan Cermak, and Sebastian Egli

1Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
2Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
3Laboratory for Climatology and Remote Sensing, Faculty of Geography, University of Marburg, Marburg, Germany

Contact: eva.pauli@kit.edu

1. Motivation and Idea

- Knowledge about drivers of fog and low stratus (FLS) development limited but crucial especially for climate models
- Relationship of spatial and temporal FLS patterns to land cover and meteorological conditions have not been studied explicitly, quantitatively and on a continental scale.
- Approach: Link FLS to various meteorological and land surface parameters using machine learning

2. Data and Methods

- FLS dataset from Egli et al. (2017), daily data from 2006-2015 on the basis of Meteosat SEVIRI
- Land surface data from LSA-SAF (Trigo et al., 2011), meteorological data from ERA5 (Copernicus Climate Change Service (C3S), 2017)
- Study area: continental Europe from 46°-53°N and 5°-15°E
- One Gradient Boosting Regression Trees (GBRT) model is trained and tested for each 10x10 or 15x15 SEVIRI pixel model unit spanning the study area
- Predictand: FLS hours day⁻¹
- Predictors: Mean Surface Pressure (MSP), Wind Speed (WS), FLS on the previous day (FLSprev), Fraction of Vegetation Cover (FVC), Land Surface Temperature (LST), Evapotranspiration (ET), Albedo (ALBEDO)
- Analysis of differences between all high pressure situations and differences between model runs including and excluding seasonality

3. Results

- Good model performance ($R^2 > 0.9$) with performance results depending on modeling time frame, grid size, pressure exclusion and seasonality exclusion settings
- Units in the eastern part of the study area have a positive deviation from the mean $R^2$ over all model units, whereas model units in the western part of the study area and units in topographically higher areas show a negative deviation
- Main drivers of FLS occurrence are meteorological (MSP, WS) as well as FLSprev and ET (especially during high pressure situations)
- High mean surface pressure, low wind speed and high FLSprev lead to high modeled FLS values

4. Conclusions

- Results show that FLS occurrence can be accurately modeled using machine-learning techniques based on meteorological and land surface predictors in large spatial domains such as central Europe
- Further studies will integrate the GBRT model into a land surface-based model grid to further analyze regionally different sensitivities and performances and apply the model to fog and low cloud properties such as cloud top height and liquid water path

Fig. 3: Mean 10x10 permutation importance over all model units for all features and all seasons.

Fig. 4: Partial dependence plot showing the mean response in modeled FLS occurrence to changes in all input features over all seasons for the model run using data with seasonality and all pressure situations.

References


For specific references, please consult the original publication or the cited sources.