

RESEARCH ARTICLE

An analysis of fog and low stratus life-cycle regimes over central Europe

Eva Pauli^{1,2}  | Jan Cermak^{1,2}  | Hendrik Andersen^{1,2} | Julia Fuchs^{1,2}

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

Correspondence

Eva Pauli, Institute of Meteorology and Climate Research and Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.
Email: eva.pauli@kit.edu

Funding information

Deutsche Forschungsgemeinschaft (DFG), Grant/Award Number: FU 1290/1-1; German Federal Ministry for Economic Affairs and Climate Action (BMBF), Grant/Award Number: FKZ 03EE1083C; Graduate Funding of the German States; European Union's Horizon 2020 research and innovation program, Grant/Award Number: 821205 (FORCeS)

Abstract

A better understanding of fog and low stratus (FLS) life cycles, in particular the typical onset and dissipation times, can help traffic safety and solar power planning. Besides its high dependence on the prevalent meteorological conditions, the FLS life cycle is further determined by the FLS type as well as the underlying geography and climate. While satellite-based FLS data sets with high temporal and spatial resolution are available, an objective, large-scale, satellite-based classification of FLS life-cycle regions, which would allow for more targeted analyses of the FLS life cycle, is lacking. In this study, a hierarchical clustering algorithm is applied to identify regions of similar relationships of the FLS life cycle to environmental conditions over central Europe. The clustering reveals major FLS life-cycle regimes with distinct differences in FLS occurrence, temporal FLS life-cycle patterns, and climatic characteristics. Based on regime-specific sensitivity analyses, the relationships of the FLS life cycle with near-ground temperature and specific humidity are identified as the most relevant input relationships for the clustering algorithm, especially in the Mediterranean region. FLS life-cycle regimes are further presented at a regional scale, outlining the applicability of the derived FLS life-cycle regimes to future regional process-oriented FLS life-cycle studies.

KEYWORDS

cloud regimes, Europe, fog, fog dissipation, fog formation, hierarchical clustering, life cycle, low stratus

1 | INTRODUCTION

The occurrence and life cycle, that is, formation and dissipation times, of fog influence the natural and anthropogenic environment in many ways. As fog occurs directly at the ground and reduces visibility to less than 1 km (American Meteorological Society, 2012; Glickman, 2000), it poses a threat to traffic on the ground, at sea, and in the air (Dietz et al., 2019; Kneringer et al., 2019; Leigh

et al., 1998; Pagowski et al., 2004). In addition, the occurrence of fog reduces the accuracy of the prediction of solar power production (Köhler et al., 2017). Fog is a crucial water source and transport agent for nutrients for several ecosystems around the world, such as the Namib desert (Gottlieb et al., 2019; Mitchell et al., 2020) or forests (Berry et al., 2015; Dawson, 1998; Pohl et al., 2021). Furthermore, fog is influenced by the underlying land-surface type. An enhancement of fog occurrence has been found over

forests (Pauli et al., 2022a), whereas fog cover is reduced in urban areas (Fuchs et al., 2022; Gautam & Singh, 2018).

Fog can be classified into different fog types, with the classification depending mainly on the processes leading to the formation of fog (Cotton et al., 2011; Gultepe et al., 2007). The most studied fog type is radiation fog, which forms as a result of nocturnal cooling and dissipates due to the absorption of solar radiation (“burn-off”) as well as the increasing magnitude of the sensible heat flux after sunrise (Bergot, 2016; Haeffelin et al., 2010; Roach, 1995; Steeneveld & de Bode, 2018; Wærsted et al., 2019). Another common and well-studied fog type is advection fog, which forms through the advection of a moist air mass over a surface with a different temperature (Cotton et al., 2011; Gultepe et al., 2007; Pérez-Díaz et al., 2017). Thus one of the major differences from radiation fog is that steady winds are required for the formation of advection fog, whereas radiation fog forms on-site and starts to dissipate in high wind-speed conditions (Bergot, 2016).

Studies on the occurrence and the life cycle of fog are often carried out using localized process studies in combination with numerical models and large-eddy simulations (Dupont et al., 2012; Duynkerke, 1991; Haeffelin et al., 2010; Karimi, 2020; Steeneveld & de Bode, 2018; Wærsted et al., 2019). While these studies provide detailed insights on local fog processes, a spatial overview of fog life-cycle processes is not feasible in these setups. Information on the spatial extent and development of fog can be obtained from satellite data, in particular from geostationary satellites. From the satellite perspective, fog is then treated together as one category with low stratus (fog and low stratus: FLS), as a separation of these two is irrelevant for a number of applications in meteorology and traffic safety and is technically challenging (Cermak et al., 2009). Over central Europe, a satellite-based data set on FLS occurrence (Egli et al., 2017) and FLS formation and dissipation time (Pauli et al., 2022b) exists for the years 2006–2015.

The satellite-based data set on FLS occurrence has been successfully used to distill drivers of FLS occurrence and sensitivities to environmental conditions based on a machine-learning setup (Pauli et al., 2020). Nevertheless, these sensitivities are likely highly dependent on the prevailing type of FLS, thus specific FLS regions should be separated in large-scale analyses. Existing FLS-related studies apply clustering algorithms such as self-organizing maps (SOM) to FLS occurrences and frequencies to characterize FLS occurrence in time and space (Egli et al., 2019; Knerr et al., 2021). However, these approaches are based on FLS occurrence alone and do not take into account the influence of environmental conditions on the FLS life cycle during the clustering, which could help in separating FLS types.

In this study, FLS life-cycle regimes are identified based on the relationships of the FLS life cycle to environmental conditions using a novel data set of FLS formation and dissipation time (Pauli et al., 2022b) and ERA5 reanalysis data (Hersbach, 2016) in a hierarchical clustering approach over central Europe. The results are analyzed further using a tree-based machine-learning model with the aim of identifying the most relevant input relationships for the clustering algorithm. The guiding hypothesis of the study is that the extracted FLS life-cycle regimes show dependences on the geographical setting, and climate and regimes can help to distinguish regions of different types of FLS (i.e., advected FLS, radiation fog). This analysis contributes to the understanding of the geographical distribution of FLS life-cycle regimes over central Europe. The identified FLS life-cycle regimes provide a basis for a regionally targeted analysis of mechanisms driving different FLS types.

2 | DATA AND METHODS

2.1 | Fog and low stratus formation and dissipation time data set

The basis of this study is a satellite-based FLS formation and dissipation time data set which contains information on the timing of the formation and dissipation of FLS for the years 2006–2015 over central Europe. This data set has been created by objectively identifying FLS formation and dissipation on a pixel-by-pixel basis and for each identified FLS event using logistic regression (Pauli et al., 2021; Pauli et al., 2022a). The data basis for this approach, described in Pauli et al. (2022b), is an FLS data set derived from the *Meteosat* Spinning Enhanced Visible and InfraRed Imager (SEVIRI: Egli et al., 2017) with a spatial resolution of 5 km over central Europe. This FLS data set has been created using the satellite-based operational fog observation scheme (SOFOS) by Cermak (2006), which detects FLS using a combination of tests on cloud phase, droplet size, stratiformity, and cloud height and has been well validated using visibility observations from Meteorological Aviation Routine Weather Reports (METARs). Using FLS formation and dissipation time from Pauli et al. (2022b) as a basis for this study makes it possible to cluster specifically with respect to the FLS life cycle as opposed to using FLS occurrence (in $\text{h}\cdot\text{day}^{-1}$) as a data basis.

The FLS formation and dissipation times used in this study are given as the fraction of daytime (night-time) length for FLS formation and dissipation occurring during the day (night). To make the times of FLS formation and dissipation comparable across seasons and regions,

the seasonal cycle is removed from the data. In addition, the data set is filtered for FLS events where a pixel is almost continuously covered by FLS after FLS has formed and before FLS dissipation sets in. In the context of the study at hand, this is done by excluding more dynamic situations by only choosing FLS formation and dissipation events where FLS is present for at least 80% of the 15-min time steps between FLS formation and dissipation. This approach also makes it possible to include FLS events in the analysis that are interrupted by no-data instances, for example at twilight (compare Pauli et al., 2022b) but are otherwise uninterrupted. Whenever FLS is not present for at least 10 consecutive time steps, FLS dissipation is detected (Pauli et al., 2022b). Summarized over the complete study area, these “complete events” make up 77% of all FLS events and 64%–96% from pixel to pixel.

2.2 | Reanalysis data set

To determine the relationship of FLS formation and dissipation times to environmental conditions, ERA5 land reanalysis data from the European Centre for Medium-Range Weather Forecasts (ECMWF), with a spatial resolution of 9 km grid size and a temporal resolution of 1 hour are used (Hersbach, 2016). Similar environmental features to those found to influence the mean daily FLS occurrence in Pauli et al. (2020) are used, specifically air temperature ($t2m$) and specific humidity (q) at 2 m, wind speed (ws) at 10 m, and mean surface pressure (mSP). While the FLS life cycle is also influenced by large-scale dynamic and thermodynamic conditions (Andersen et al., 2020; Dione et al., 2023), in this study we focus on the relationships of the FLS life cycle to environmental conditions near the ground and at the local scale.

The ERA5 data set is rescaled to the spatial resolution of the FLS data set using nearest-neighbour interpolation. Further data preparation steps include the removal of the diurnal cycle from $t2m$ and q , and the removal of the seasonal cycle from $t2m$, q , ws , and mSP . By removing diurnal and seasonal fluctuations from the data sets used, their influence on the relationship of the FLS life cycle with environmental conditions is minimized. Lastly, daytime and night-time means are calculated using the pixel-specific timings of sunrise and sunset, which are then used for the calculation of the relationships of the FLS life cycle with the environmental features. By using the specific daytime and night-time means for the calculation of the correlations, both the day-to-day variability and the changes of the environmental conditions during the day/night are accounted for.

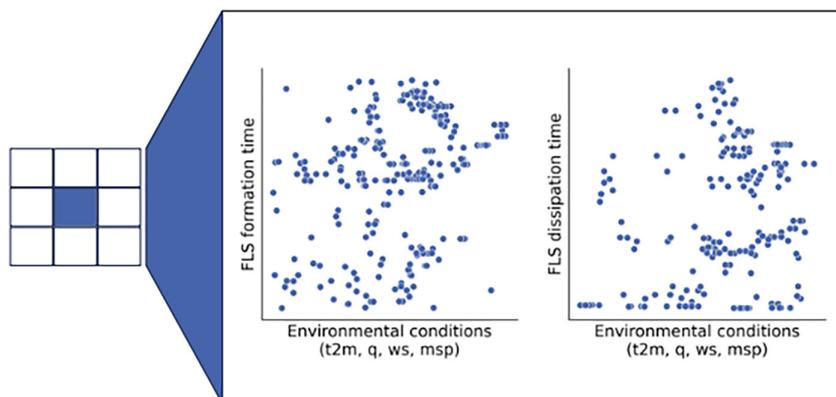
2.3 | Clustering algorithm

2.3.1 | Calculation of pixel-based correlations

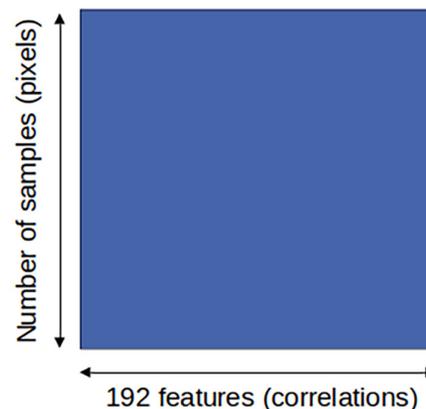
A schematic overview of the methodological approach used in this study is shown in Figure 1. The basis for the clustering algorithm consists of pixel-based Spearman's rank correlations of FLS formation and dissipation time with the respective day- and night-time means of the environmental features from ERA5. In this study, Spearman's rank correlation coefficient is calculated, as it reflects the strength of a monotone relationship and is more robust to outliers compared with Pearson's correlation coefficient, which relies on the linearity of the relationship considered (Wilks, 2006). While the Spearman's rho obtained describes the magnitude and direction of the relationship of the FLS life cycle with each environmental feature and for each pixel separately, these correlations do not necessarily provide evidence of a physical link between the FLS life cycle and the chosen environmental conditions. Nevertheless, temperature, humidity, surface pressure, and wind speed belong to the major environmental drivers of FLS occurrence (Pauli et al., 2020), thus strong correlations in the context of this study likely also indicate a physical link. Therefore, these correlations provide a suitable data basis to identify regions where temporal FLS formation and dissipation patterns can be attributed to similar changes in environmental conditions.

As FLS formation and dissipation do not occur daily over each pixel and occurrences depend strongly on the season (Pauli et al., 2022b), the adjacent eight pixels of the FLS data set are included in the correlation calculation, leading to a larger number of observations available. This is especially beneficial over regions with low FLS occurrence in specific months, that is, over the Mediterranean in summer. Using only data available for one pixel would likely lead to nonreliable correlation values there. Over each of these 3×3 moving pixel windows, the Spearman's rho correlation values are calculated for daytime and night-time FLS formation and dissipation but also for each month separately, as the relationships might vary for day- and night-time and for separate months. As the correlations are calculated for four environmental features, this results in $4 \times 2 \times 2$ (four features, formation and dissipation, daytime and night-time) correlations per month ($4 \times 2 \times 2 \times 12$), thus 192 correlations for each pixel overall. If no FLS formation or dissipation occurs over a certain pixel for a complete month, for example, in the Mediterranean in summer, the Spearman's rho value is set to zero, which is the case for 0.14% of pixels. No significance threshold is applied to the data, as removing a nonsignificant correlation (i.e., $p \leq 0.05$) from the input

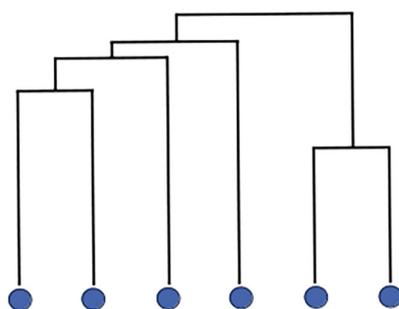
1) Pixel-wise correlation calculation



2) Data set allocation



3) Agglomerative hierarchical clustering



4) Cluster feature importance

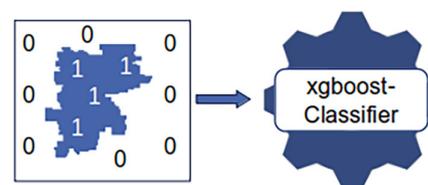


FIGURE 1 Schematic overview of the methodological approach used in this study. The pixelwise correlation is performed for each month and both daytime and night-time formation and dissipation events separately. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

data or setting it to NaN (Not a Number) would lead to an exclusion of the pixel from the clustering. In general, the magnitude of nonsignificant correlations, summarized over all pixels and all input features, is lower than that of significant correlations (Figure S1). Thus, the influence of the nonsignificant correlations on the clustering patterns is most likely marginal, as low correlation values have little influence on the clustering procedure performed.

An overview of the correlations used as input for the clustering algorithm is shown in Table 1. The scatter plot shown in Figure 1 schematically shows the relationship of a chosen environmental feature with either formation or dissipation time for a specific month and for either day- or night-time.

2.3.2 | Hierarchical clustering

After the allocation of the correlation values into a table with the pixels as rows and the correlations as columns (Figure 1, step 2), a hierarchical, agglomerative clustering algorithm is applied to the data. This algorithm first

separates the data set into singleton nodes (pixels) and then merges those nodes that are most similar (depending on a similarity measure) until only one node is left (Everitt, 2011; Müllner, 2011). This approach groups pixels with similar correlations, that is, similar relationships of the FLS life cycle to environmental conditions, into specific FLS life-cycle regimes. Agglomerative hierarchical clustering is an unsupervised machine-learning technique using a specified distance or similarity measure to group the data. Here, the ward linkage is used, which aims to minimize the increase in the error sum of squares (i.e., the sum of the squared differences between each correlation value and the cluster mean) within a cluster summed over all input features (Everitt, 2011). This means that two clusters are merged if the increase in the error sum of squares, summed over all input features, is smallest compared with merging any other two clusters together. Using this approach, the error sum of squares is zero for each cluster at the start of the clustering procedure (as each pixel is its own cluster) and then increases whenever two clusters are merged.

TABLE 1 Overview of the variable combinations resulting in the correlations that serve as an input for the clustering algorithm.

Environmental feature	Day / night	Formation / dissipation	Abbreviation
<i>t2m</i>	day	form	<i>t2m</i> _{day, form}
		diss	<i>t2m</i> _{day, diss}
	night	form	<i>t2m</i> _{night, form}
		diss	<i>t2m</i> _{night, diss}
<i>q</i>	day	form	<i>q</i> _{day, form}
		diss	<i>q</i> _{day, diss}
	night	form	<i>q</i> _{night, form}
		diss	<i>q</i> _{night, diss}
<i>mSP</i>	day	form	<i>mSP</i> _{day, form}
		diss	<i>mSP</i> _{day, diss}
	night	form	<i>mSP</i> _{night, form}
		diss	<i>mSP</i> _{night, diss}
<i>WS</i>	day	form	<i>WS</i> _{day, form}
		diss	<i>WS</i> _{day, diss}
	night	form	<i>WS</i> _{night, form}
		diss	<i>WS</i> _{night, diss}

Note: The combinations are calculated for each month and pixel.

For the application of the clustering algorithm, the fastcluster Python package by Müllner (2011, 2013) is used, as it clusters a large input data set efficiently and has been used successfully in Egli et al. (2019) to cluster daily averages of fog occurrences. The result of the clustering is a dendrogram, which displays how the data set is merged into clusters as a rooted tree. The dendrogram is usually displayed upside down, with the leaves at the bottom, depicting the singleton nodes, and the internal nodes showing where two clusters are joined together. The dendrogram further shows the chosen distance measure on the y-axis, which increases with increasing cluster size, that is, the number of pixels contained in a cluster. One of the main advantages of hierarchical clustering is that the level of hierarchy at which the clusters are extracted can be chosen using the dendrogram and the relationships of clusters on different hierarchy levels can be analyzed. To describe this relationship between different hierarchy levels, we use the word “parent cluster” in the results section to describe multiple clusters that were merged to form one single cluster. In addition, hierarchical clustering always returns the same clusters when run with the same input data, unlike other frequently used clustering methods such as *k*-means, where the random shuffling of the input data has to be set explicitly to receive the same clusters in multiple runs (Hartigan & Wong, 1979). To show further the suitability of

hierarchical clustering over the *k*-means algorithm in the context of the study, the clustering results using the *k*-means algorithm can be found in the Supporting Information (Figures S2 and S3). The obtained dendrogram of the hierarchical clustering approach applied in this study is shown in Figure 2 and its characteristics are described in Section 3.1.

2.4 | Clustering feature importance

To investigate which input features are important for the clustering procedure and if these importances vary from one cluster to another, cluster-specific binary classification models using the input variables used for clustering (Badih et al., 2019; Breiman, 2001) are created. This workaround is necessary, as no feature importance measure exists for clustering methods which is comparable to tree-based machine-learning models (Breiman, 2001; Molnar, 2019; Strobl et al., 2007). Here, extreme gradient boosting (XGB) classification is used, which is a tree-based machine-learning technique with short run times and built-in regularization techniques (Chen & Guestrin, 2016). A classification model is created for each major cluster, which predicts if a pixel belongs to the selected cluster (pixel value = 1) or not (pixel value = 0) (Figure 1, step 4).

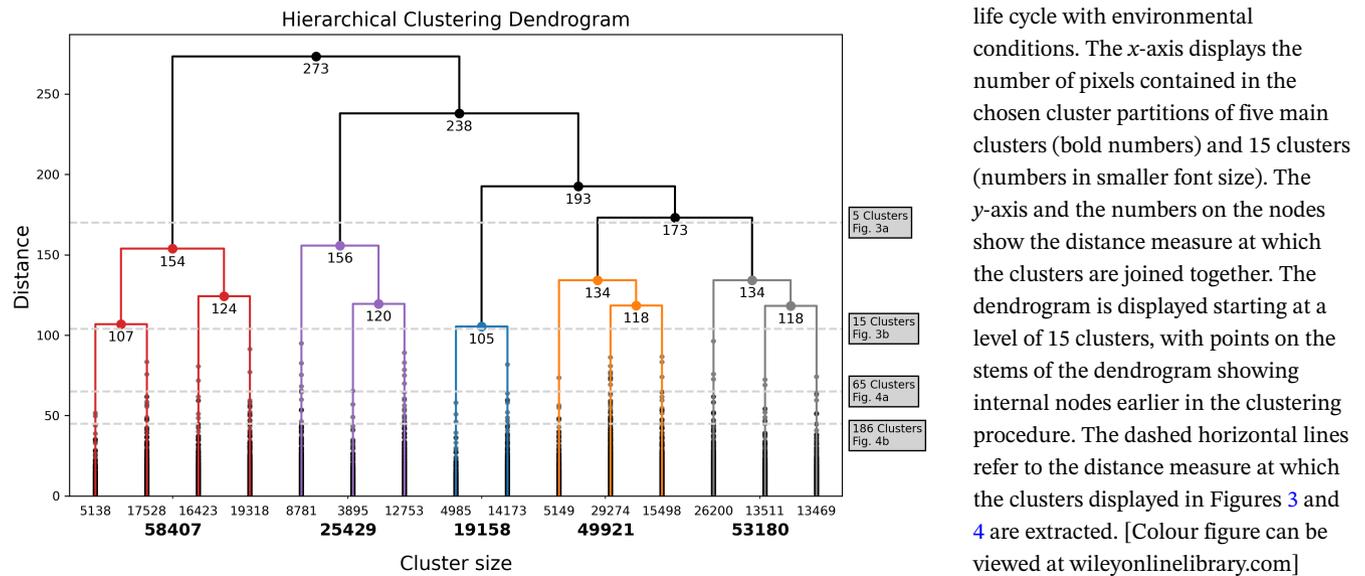


FIGURE 2 Dendrogram of the hierarchical clustering algorithm applied to the correlations of the FLS life cycle with environmental conditions. The x-axis displays the number of pixels contained in the chosen cluster partitions of five main clusters (bold numbers) and 15 clusters (numbers in smaller font size). The y-axis and the numbers on the nodes show the distance measure at which the clusters are joined together. The dendrogram is displayed starting at a level of 15 clusters, with points on the stems of the dendrogram showing internal nodes earlier in the clustering procedure. The dashed horizontal lines refer to the distance measure at which the clusters displayed in Figures 3 and 4 are extracted. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4714)]

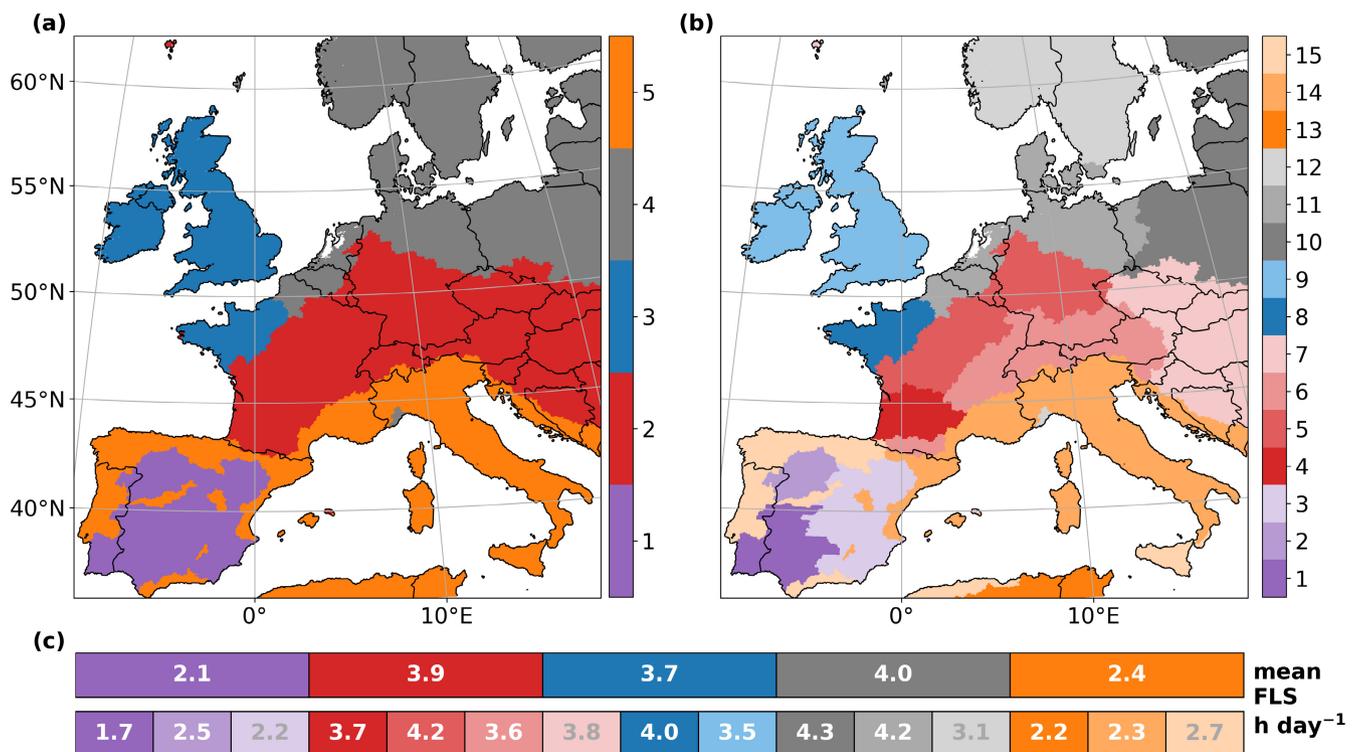


FIGURE 3 Spatial overview of (a) five and (b) 15 FLS life-cycle regimes over the study area. The average FLS occurrence in h day^{-1} over the complete period from 2006–2015 for all five (top) and 15 (bottom) regimes can be found in (c). Country borders are plotted as a black line. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/doi/10.1002/qj.4714)]

The resulting feature importances for each of these classification models indicate the relevant relationships for each specific cluster. As the number of input features used in this study is relatively high (192, cf. Section 2.3.1), separate models for formation and

dissipation and for each month are created, resulting in eight features for each cluster-specific model, with 120 models overall (five major clusters, 12 months, formation and dissipation). The default hyperparameters for model training are used (number of estimators 100,

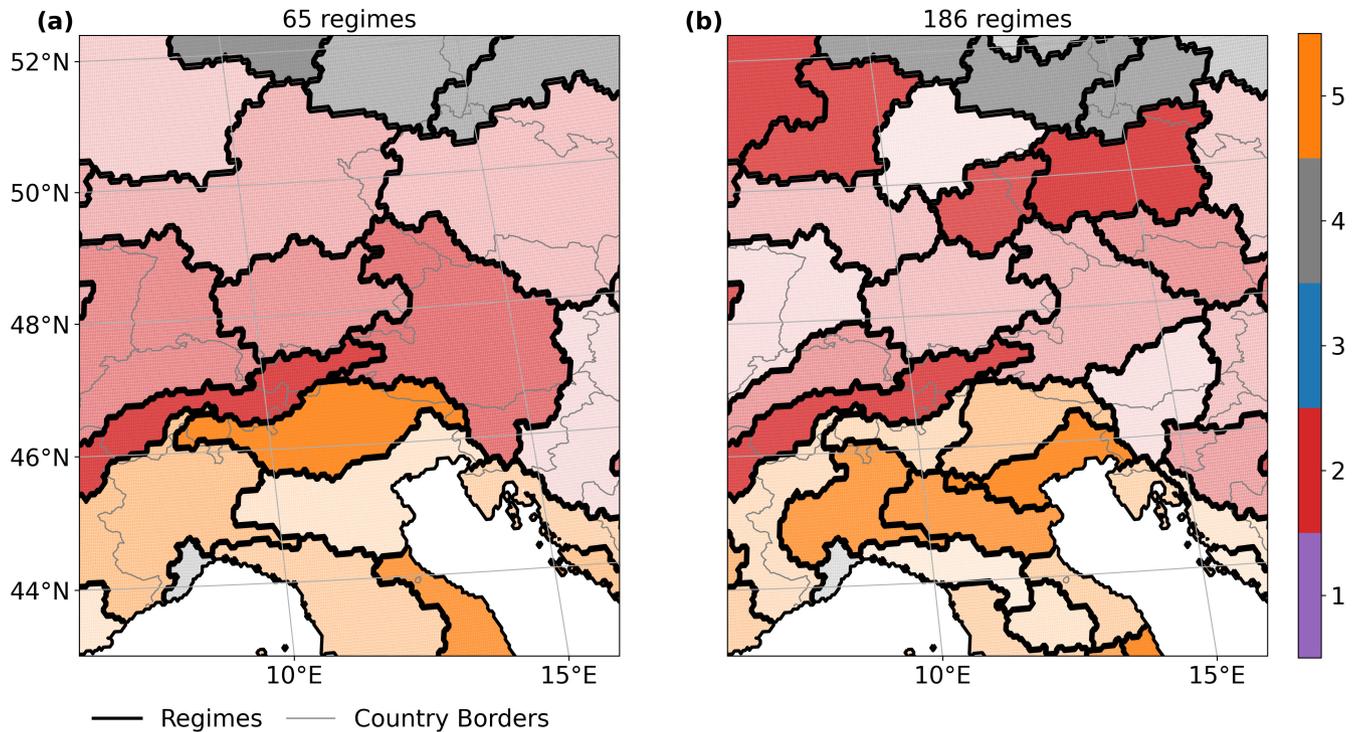


FIGURE 4 Regional and subregional regimes from a clustering hierarchy level of (a) 65 and (b) 186 regimes. Regime borders are plotted as a double black line, country borders as a grey line. The colours of the regimes are chosen based on the five main regimes in Figure 3 and are brightened up for better visual delineation of the regimes. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

learning rate 0.3, maximum depth 6) and model performance is evaluated by calculating the accuracy score (fraction of correct predicted labels) on a held-back test data set (test–train split: 30–70). Then, the permutation feature importance is calculated, which measures the increase in the prediction error after randomly shuffling one feature (Breiman, 2001). The permutation feature importance of each feature is then averaged over all monthly models to give a general overview of the feature importances in each cluster-specific formation and dissipation model.

3 | RESULTS AND DISCUSSION

3.1 | Spatial patterns of identified FLS regimes

The dendrogram in Figure 2 shows the last steps of the hierarchical clustering algorithm, that is, how the largest clusters are merged into one final cluster and how these clusters are related. The main analysis is limited to the five clusters shown in Figure 3a. For additional spatial detail, a finer subcontinental delineation into 15 clusters, that is, precursors of the five clusters, is shown in Figure 3b and a regional and subregional cluster overview is shown in Figure 4. To link the dendrogram to the cluster maps

in Figures 3 and 4, the distance measure at which these clusters are extracted and the resulting number of clusters, as well as the corresponding figures, are marked in the dendrogram. Since the clustering is based on a grouping of parameters of the FLS life cycle, the clusters are referred to as “regimes” in the following, in analogy to other studies that have created cloud regimes by applying *k*-means clustering to cloud properties based on satellite data (Oreopoulos et al., 2014; Tselioudis et al., 2013; Tzallas et al., 2022).

Figure 3 shows a spatial map of (a) five continental and (b) 15 subcontinental FLS regimes over the study area, with (c) the corresponding mean FLS occurrence over the complete period from 2006–2015. Hence, the mean FLS occurrence also includes time steps with FLS that are not assigned to an FLS event, for example, due to frequent changes between FLS and no-FLS. The clear regional delineation of the FLS life-cycle regimes shows that the hierarchical clustering algorithm is able to group the underlying pixel-based correlations into regimes. These match known spatial patterns of FLS occurrence (i.e., Bendix, 1994, 2002; Cermak et al., 2009; Egli et al., 2017; Román-Cascoón et al., 2016; Scherrer & Appenzeller, 2014) and also represent different climate regions of the study area. Figure S2 shows the result of the *k*-means clustering, showing similar spatial features but a lower spatial homogeneity of the regimes, with all

regimes being present at all latitudes. A possible reason for the lower spatial homogeneity of the k -means clustering results could be that the k -means algorithm moves the pixels from one cluster to another, until a local optimum of the within-cluster sum of squares is found (Hartigan & Wong, 1979). In hierarchical clustering, on the other hand, a pixel that is assigned to a cluster cannot be moved to a different one (Everitt, 2011), which in our case likely leads to spatially homogeneous clusters due to similar correlations of neighbouring pixels.

Due to missing ground-truth data of FLS life-cycle regimes over the area and time period investigated in this study, no traditional validation can be performed here. Nevertheless, to support the spatial patterns of the FLS life-cycle regimes based on the hierarchical clustering approach, the regimes are linked to the occurrence of radiation and advection fog in particular, as these are the most studied and common fog types in central Europe and can be detected with the spatial resolution of about $5 \times 5 \text{ km}^2$ used here. Other fog types, such as mountain fog or coastal fog, are also likely to occur in most of the regimes and are to some extent included in the discussion.

Regime 1 (purple) covers the central region of the Iberian Peninsula and parts of southern Portugal and is the smallest of the five main regimes. As the climate in this region has continental characteristics with a reduced maritime influence (Royé et al., 2019), radiation fog is likely the dominating FLS type in regime 1 (Romaán-Cascoón et al., 2016). For the subcontinental clustering level shown in Figure 3b, Regime 1 is based on three different regimes, one covering the more continental plateau of central Spain and the other two each covering the east and west of Regime 1 respectively (Figure 3b). While FLS occurrence in regime 1 is lowest of all five main regimes ($2.1 \text{ hr}\cdot\text{day}^{-1}$), it is even lower for its subcontinental parent cluster in the southeast ($1.7 \text{ hr}\cdot\text{day}^{-1}$).

A wide region from the Atlantic coast of southern France, over Germany to the Balkans, is covered by regime 2 (red). Besides the part of the regime on the Atlantic coast, this regime can be described as continental, with pronounced topographic variability spanning several mountain ranges of Central Europe such as the Massif Central, parts of the Alps, and the Carpathian mountains. FLS at inland locations of this regime is likely dominated by radiative processes, as FLS can form during calm wind situations in topographic low-lying areas, that is, in the Swiss Plateau or the Danube valley (Bendix, 2002; Cermak et al., 2009; Egli et al., 2017; Scherrer & Appenzeller, 2014). Advection fog also likely plays a role, especially at higher elevations, where low stratus is advected onto mountainous sites in the regime, for example in the Czech Republic (Hůnová, 2020), leading to the formation of mountain fog.

Regime 2 also covers the Landes forest on the Atlantic coast of southern France, where higher FLS occurrence over the forest compared with the surrounding agricultural land has been detected (Pauli et al., 2022b). Average FLS occurrence in regime 2 is the second highest of the main regimes considered ($3.9 \text{ hr}\cdot\text{day}^{-1}$), and the highest for the northeastern subregime ($4.2 \text{ hr}\cdot\text{day}^{-1}$).

Regime 3 (blue) covers the British Isles and the northern Atlantic coast of France. Regime 3 is likely dominated by advection fog (Bendix, 2002; Mayes, 2013), with coastal fog being prevalent in the coastal regions (Fallmann et al., 2019), and has high average FLS occurrence ($3.7 \text{ hr}\cdot\text{day}^{-1}$). Concerning the two parent regimes of regime 3 (namely regimes 8 and 9 in Figure 3b), the average FLS occurrence is higher in the parent regime covering the Atlantic coast of northern France.

The Baltic–Scandinavian region, the Benelux countries, and northern Germany are covered by regime 4 (grey), with the highest mean FLS occurrence of $4.0 \text{ hr}\cdot\text{day}^{-1}$. In this regime, FLS of both advective and radiative origin are present: advection fog can form in this region through the advection of warm air from the Baltic sea over cold land, whereas radiation fog especially forms in anticyclonic conditions (Avotniece et al., 2015; Bendix, 2002). Three subcontinental regimes form regime 4, of which the Baltic regime has the highest FLS occurrence of all regimes considered in Figure 3 ($4.3 \text{ hr}\cdot\text{day}^{-1}$). A prominent feature of regime 4 is that it also contains a region on the Mediterranean coast of Italy. While specific case studies are necessary to investigate this pattern further, it should be noted that the hierarchical clustering does not reassign a pixel once it is assigned to a cluster, which is likely the reason why this area is not reassigned to the Mediterranean regime in a later step (Everitt, 2011).

Regime 5 (orange) covers most of the Mediterranean and the Atlantic coast of the Iberian Peninsula, with a comparably low average FLS occurrence of $2.4 \text{ hr}\cdot\text{day}^{-1}$. The subcontinental regimes that form regime 5 describe to some extent the FLS types that are prevalent in this regime. On the Mediterranean coast of Spain, the advection of clouds onto the mountain ranges and orographic lifting of maritime air masses are contributing to FLS occurrence in this region (Azorin-Molina et al., 2014; Estrela et al., 2008; Valiente et al., 2011). Advective processes are also most likely the dominating contributor to FLS formation on the Atlantic coast of Portugal and Spain where moist air and FLS formed over the Atlantic are advected onto the land (Egli et al., 2017; Guerreiro et al., 2020). Regime 5 also contains the Po valley in northern Italy where radiation fog is a frequent phenomenon, especially in fall and winter (Bendix, 1994, 2002; Fuzzi et al., 1992; Wobrock et al., 1992). Further delineation of the Po valley is visible in Figure 4.

A more detailed view on the regional and subregional FLS regimes composing the major continental regimes 2, 4, and 5 is shown in Figure 4 for the hierarchy levels of (a) 65 and (b) 186 regimes. These FLS regimes divide southern Germany, Switzerland, and the northern part of Italy into regional FLS regimes, which have been partly discussed above: for example, the Po valley between 44°–46°N and 9°–12°E. These regimes are known for frequent FLS occurrences, in particular radiation fog, and can be used to study FLS occurrence, life cycle, and processes on a regional scale in the future. Such a setup makes it possible to investigate FLS-regime or FLS-type specific processes on a regional to subregional scale. Using regional or subregional FLS regimes to investigate FLS life-cycle processes also avoids the (potentially misleading) summary of processes when using large, continental regimes (i.e., as in Figure 3a). The regional and subregional FLS life-cycle regimes based on *k*-means clustering are shown in Figure S3 and are similar to the hierarchical clustering regimes, but with a slightly higher spatial variability.

3.2 | Most frequent FLS formation and dissipation time of major FLS life-cycle regimes

As discussed in the context of Figure 3, the clustering procedure leads to FLS regimes with distinct differences in the mean FLS occurrence, which is likely a result of the geographical setting leading to increasing FLS occurrence in continental regions in the northeast of the study area

(Cermak et al., 2009; Egli et al., 2017). The most frequent formation and dissipation times summarized over all pixels of each of the five major FLS regimes are given in Figure 5 (for the complete time period) and Figure 6a,b (for each season). A definition of the formation and dissipation times can be found in Pauli et al. (2022b). The duration of all FLS events considered in the study, for each of the five major regimes, is shown in Figure S4.

Over all regimes and when considering the complete time period (full year), FLS forms most frequently in the morning or around sunset and dissipates most frequently at sunrise, in the morning, or in the afternoon. As the underlying FLS data set is not able to detect FLS during sunrise and sunset (Egli et al., 2017), FLS events that extend over sunrise or sunset are occasionally interrupted, which leads to higher uncertainties in the detection of FLS formation and dissipation time (Pauli et al., 2022b).

The Mediterranean regime 1 shows a distinct pattern of mostly morning FLS formation and afternoon dissipation, with a shift to earlier dissipation in spring and summer. This is likely a radiation fog pattern, as FLS typically forms around two hours after sunrise, since more time is needed for saturation to be reached in the dry climate of this region (Romaán-Cascoón et al., 2016). The duration of FLS events in this regime is generally lowest of all regimes and FLS events occur most frequently in winter (Figures 6c and S4).

In regime 2, FLS forms most frequently around midnight and during sunset and dissipates around midday and in the afternoon (Figure 5). These patterns are strongly dependent on the season considered, with more FLS formation around sunrise in spring and summer and more

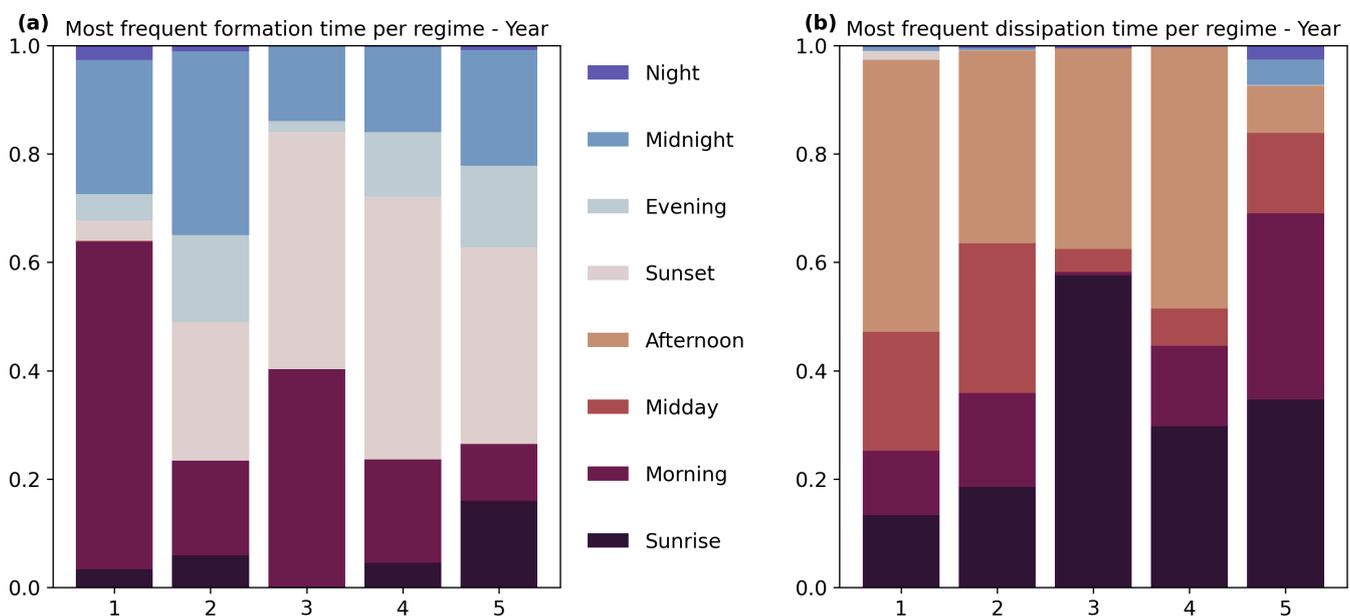


FIGURE 5 Most frequent (a) FLS formation time and (b) FLS dissipation time for each regime over the complete time period (Year). [Colour figure can be viewed at wileyonlinelibrary.com]

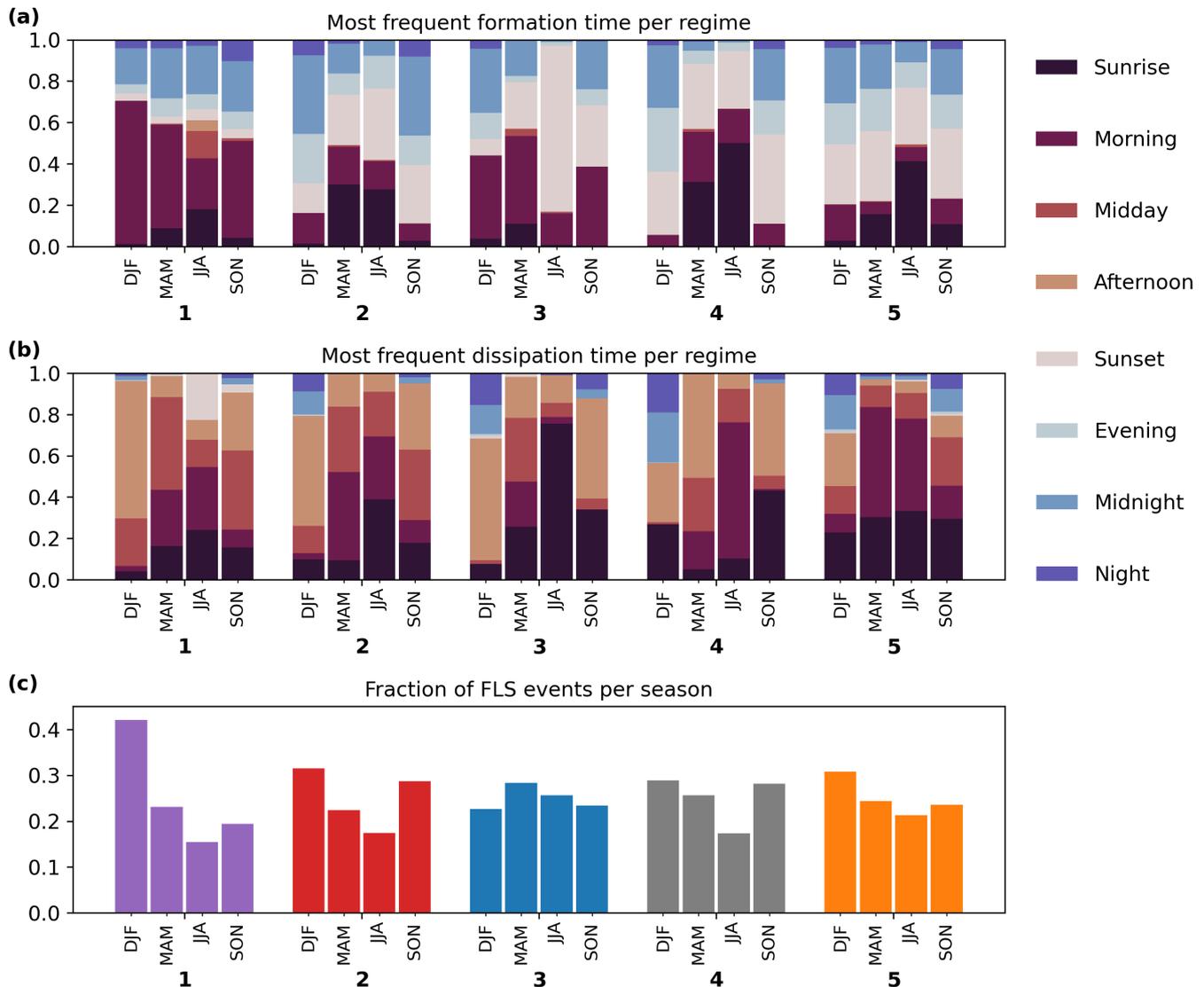


FIGURE 6 Most frequent (a) FLS formation, (b) FLS dissipation times, and (c) fraction of FLS events per season for each regime over separate seasons. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com/terms-and-conditions)]

night-time formation in fall and winter (Figure 6a,b). Dissipation patterns similarly switch from dissipation in the afternoon in winter and fall to the morning hours in spring and summer. FLS events are more frequent and longer in fall and winter, whereas in spring and summer fewer and shorter FLS events occur (Figures 6c and S4).

When considering the full year, FLS formation in the maritime regime (regime 3) occurs most frequently in the morning and around sunset and dissipation at sunrise and in the afternoon (Figure 5). The patterns of most frequent formation and dissipation time are particularly interesting in summer, where FLS formation occurs almost exclusively around sunset and dissipation mostly around sunrise. This is likely FLS that is formed over the sea and advected onto coastal areas, for example on the east coast of the UK, resulting in coastal fog (Fallmann et al., 2019).

Still, the maritime regime 3 most likely contains both radiative and advective FLS events, with radiation fog especially prevalent in winter (Price et al., 2015).

In the Baltic–Scandinavian regime 4, most of the FLS events form around sunset, or around midnight. Dissipation frequently occurs in the afternoon or at sunrise (Figure 5). The average duration of FLS events occurring in this regime is high and FLS events occur mostly in winter and fall (Figures 6c and S4). The distribution of most frequent formation and dissipation time shows considerable seasonal differences (Figure 6): in winter and fall, FLS events with formation at sunset, in the evening, and during the night dominate, with FLS dissipation during the night (DJF) and in the afternoon (SON). This pattern shifts towards spring and summer, where FLS forms frequently around sunset and sunrise and dissipates

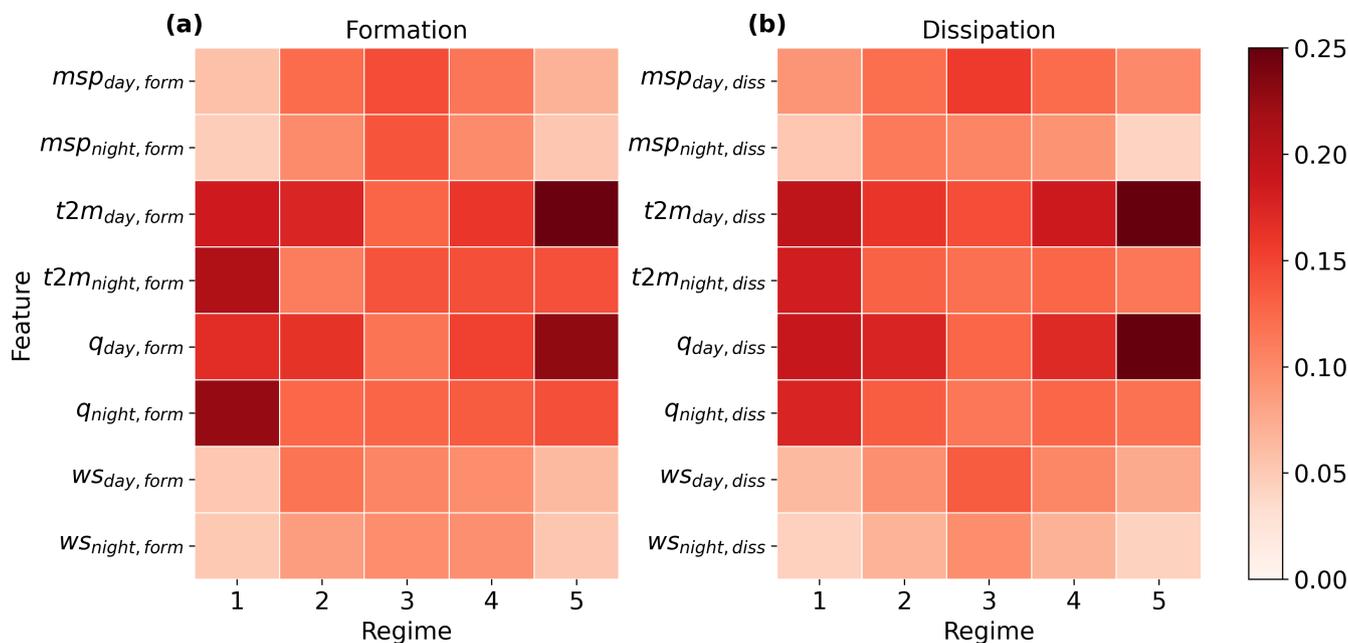


FIGURE 7 Permutation importance for all cluster specific (a) formation and (b) dissipation models averaged over all months. [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

in the morning (JJA) or in the afternoon (MAM). This potentially displays a shift of radiation fog occurrence in winter to more frequent advective fog events in spring (Avotniece et al., 2015). FLS events that dissipate during the night are found to be slightly shorter compared with the duration averaged over all dissipation times (compare Figure S4), a pattern that has also been identified by other studies (Wærsted et al., 2019). Potential reasons for FLS dissipation at night might be increases in wind speed or the advection of higher clouds, leading to the dissipation of FLS.

FLS formation in the Mediterranean regime 5 occurs most frequently around sunset, at sunrise, or around midnight. Dissipation occurs most frequently at sunrise and in the morning. While the most frequent formation time is similar over all seasons, the distribution of the most frequent dissipation time varies strongly. In winter and fall, dissipation occurs mainly at sunrise, around midday, or in the afternoon, whereas in spring and summer dissipation shifts to the morning hours. The formation and dissipation patterns in winter and fall in regime 5 are likely the result of radiation fog events, as nocturnal cooling leads to FLS formation, for example, in the Po valley (Bendix, 1994; Fuzzi et al., 1992). A possible explanation for frequent night-time FLS formation in this regime is FLS formation at mountainous sites during the night. According to Bendix (1994), a night-time peak of FLS occurrence at mountainous sites, for example, close to the gulf of Genoa, likely occurs due to a combination of a valley breeze with a sea breeze, where moist air is transported over the mountains onto the Po valley, leading to condensation and

FLS formation. This also highlights the importance of considering advective processes in FLS life-cycle studies, as these can lead to the occurrence of different FLS types at locations in close proximity and under the same synoptic conditions (Bari et al., 2015).

3.3 | Feature importance

In the following, the most important input correlations for each regime are described and discussed using the permutation importance derived from the cluster-specific classification models described in Section 2.4. This is done by averaging the permutation feature importance obtained from all regime-specific monthly models. This allows us to determine the most important correlations and thus potential important environmental drivers for the FLS life cycle in a specific regime.

The accuracy score of these monthly models (fraction of correct predicted labels) is generally high and ranges from around 0.75–0.94 (Figure S5), showing that the correlations used as an input for the clustering algorithm are well suited to separating one specific regime from all others. The average permutation importance (Figure 7) shows that the correlations of temperature and specific humidity with both formation and dissipation, during day and night, are most important across all regimes. The importance of these correlations is highest for the Mediterranean regimes 1 and 5, indicating that, in a Mediterranean climate, temperature and specific humidity are more important for the FLS life cycle than large-scale

atmospheric dynamics. In all other regimes, the permutation importance is more evenly distributed over the features and wind speed and mean surface pressure gain importance. This may suggest that both large-scale atmospheric dynamics and near-ground temperature and specific humidity variations have a strong influence on FLS in these regimes, as has been shown for a region situated in regime 2 in Pauli et al. (2020).

4 | CONCLUSIONS

In this study, FLS life-cycle regimes have been identified based on a hierarchical clustering approach using the relationships of the FLS life cycle to environmental conditions. The extracted regimes separate coherent FLS regions, which can be associated with different FLS types. The underlying relationships of the FLS life cycle to environmental conditions likely vary due to a varying climate background and potential underlying FLS processes. The geographic extent, dominating FLS types, and FLS formation and dissipation climatologies of five major regimes are described and discussed. Two of the five major FLS life-cycle regimes, covering parts of the Iberian Peninsula (regime 1) and most of the midlevel mountain ranges of the central study area (regime 2) show FLS life-cycle characteristics primarily associated with radiation fog, with formation most frequently in the morning, at sunset, and in the night, and dissipation at midday and in the afternoon. The maritime regime covering the British Isles and parts of the Atlantic coast (regime 3), the regime covering mostly the Baltic–Scandinavian area (regime 4), and the Mediterranean regime (regime 5) have a higher frequency of FLS formation events at sunset as well as more dissipation events at sunrise and in the morning. While radiation fog events are frequent in these regimes as well, there is likely a higher frequency of advection of FLS from the sea and/or onto mountain ranges in these regimes. The most important features for the clustering approach hint towards the importance of temperature and specific humidity in regimes that are likely dominated by radiative FLS events and highlight the role of the large-scale atmospheric dynamics for regimes where advective FLS events are frequent.

The results of the study show the successful application of the cloud regimes concept to fog and low stratus clouds and extend this concept further by running the clustering exclusively on correlations of the FLS formation and dissipation time with environmental conditions. With an objective, satellite-based approach, central Europe is divided into distinct FLS life-cycle regimes on a continental to subregional level and can be linked to FLS types known from the literature. These spatially coherent regimes are

an ideal basis for further regional FLS life-cycle sensitivity studies, for example, by setting up regime-specific machine-learning models to predict the formation and dissipation of FLS. In this approach, a set of meteorological reanalysis variables at the ground and on pressure levels could be used as predictors, in both cases at several time steps before FLS formation and dissipation to include the current and past state of the atmosphere. The relationships of the FLS life cycle to the predictors used can then be explored further by using model-agnostic methods such as SHapley Additive exPlanations (SHAP) values (Lundberg et al., 2020; Lundberg & Lee, 2017). These sensitivities can then also be compared across regimes and used to explore different processes across FLS types.

AUTHOR CONTRIBUTIONS

Eva Pauli: conceptualization; formal analysis; investigation; methodology; visualization; writing – original draft; writing – review and editing. **Jan Cermak:** conceptualization; methodology; writing – review and editing. **Hendrik Andersen:** methodology; writing – review and editing. **Julia Fuchs:** methodology; writing – review and editing.

ACKNOWLEDGEMENTS

We thank two anonymous reviewers and Stephanie Westerhuis for their careful reviews, which have helped improve the manuscript. E. Pauli has been financially supported by the Graduate Funding of the German States. We also acknowledge support by Deutsche Forschungsgemeinschaft (DFG) in grant FU 1290/1-1, URCLOCK, the German Federal Ministry for Economic Affairs and Climate Action (BMBF) in grant FKZ 03EE1083C, SnowFogS, and the European Union's Horizon 2020 research and innovation program under grant agreement no. 821205 (FORCeS). Open Access funding enabled and organized by Projekt

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

DATA AVAILABILITY STATEMENT

The FLS formation and dissipation time data set is available at <https://doi.org/10.5445/IR/1000141293>. The era5-land data is can be downloaded from <https://doi.org/10.24381/cds.e2161bac>.

ORCID

Eva Pauli  <https://orcid.org/0000-0002-5506-4736>

Jan Cermak  <https://orcid.org/0000-0002-4240-595X>

REFERENCES

- American Meteorological Society. (2012) *Fog*. In: *Glossary of meteorology*. Boston: American Meteorological Society.
- Andersen, H., Cermak, J., Fuchs, J., Knippertz, P., Gaetani, M., Quinting, J. et al. (2020) Synoptic-scale controls of fog and low-cloud variability in the namib desert. *Atmospheric Chemistry and Physics*, 20(6), 3415–3438. Available from: <https://doi.org/10.5194/acp-20-3415-2020>
- Avotniece, Z., Klavins, M. & Lizuma, L. (2015) Fog climatology in Latvia. *Theoretical and Applied Climatology*, 122(1), 97–109. Available from: <https://doi.org/10.1007/s00704-014-1270-4>
- Azarin-Molina, C., Sanchez-Lorenzo, A., Calboó, J., Azarin-Molina, C., Sanchez-Lorenzo, A. & Calbo, J. (2014) A climatological study of sea breeze clouds in the southeast of the Iberian Peninsula (Alicante, Spain). *Atmosfera*, 22(1), 33–49.
- Badih, G., Pierre, M. & Laurent, B. (2019) Assessing variable importance in clustering: a new method based on unsupervised binary decision trees. *Computational Statistics*, 34(1), 301–321. Available from: <https://doi.org/10.1007/s00180-018-0857-0>
- Bari, D., Bergot, T. & El Khilifi, M. (2015) Numerical study of a coastal fog event over Casablanca, Morocco. *Quarterly Journal of the Royal Meteorological Society*, 141(690), 1894–1905. Available from: <https://doi.org/10.1002/qj.2494>
- Bendix, J. (1994) Fog climatology of the Po valley. *Rivista di Meteorologia Aeronautica*, 3–4, 25–36.
- Bendix, J. (2002) A satellite-based climatology of fog and low-level stratus in Germany and adjacent areas. *Atmospheric Research*, 64, 3–18.
- Bergot, T. (2016) Large-eddy simulation study of the dissipation of radiation fog. *Quarterly Journal of the Royal Meteorological Society*, 142(695), 1029–1040. Available from: <https://doi.org/10.1002/qj.2706>
- Berry, Z.C., Johnson, D.M. & Reinhardt, K. (2015) Vegetation-zonation patterns across a temperate mountain cloud forest ecotone are not explained by variation in hydraulic functioning or water relations. *Tree Physiology*, 35(9), 925–935. Available from: <https://doi.org/10.1093/treephys/tpv062>
- Breiman, L. (2001) Random forests. *Machine Learning*, 45, 5–32. Available from: <https://doi.org/10.1109/ICCECE51280.2021.9342376>
- Cermak, J. (2006) *SOFOS-A new satellite-based operational fog observation scheme*, Unpublished doctoral dissertation. Marburg, Germany: Philipps-Universität Marburg.
- Cermak, J., Eastman, R.M., Bendix, J. & Warren, S.G. (2009) European climatology of fog and low stratus based on geostationary satellite observations. *Quarterly Journal of the Royal Meteorological Society*, 135, 2125–2130. Available from: <https://doi.org/10.1002/qj.503>
- Chen, T. & Guestrin, C. (2016) XGBoost: a scalable tree boosting system. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. New York, NY, USA: Association for Computing Machinery, pp. 785–794. Available from: <https://doi.org/10.1145/2939672.2939785>
- Cotton, W.R., Bryan, G. & van den Heever, S.C. (2011) Fogs and stratocumulus clouds. *International Geophysics*, 99, 179–242. Available from: [https://doi.org/10.1016/S0074-6142\(10\)09912-2](https://doi.org/10.1016/S0074-6142(10)09912-2)
- Dawson, T.E. (1998) Fog in the California redwood forest: ecosystem inputs and use by plants. *Oecologia*, 117(4), 476–485. Available from: <https://doi.org/10.1007/s004420050683>
- Dietz, S.J., Kneringer, P., Mayr, G.J. & Zeileis, A. (2019) Low-visibility forecasts for different flight planning horizons using tree-based boosting models. *Advances in Statistical Climatology, Meteorology and Oceanography*, 5(1), 101–114. Available from: <https://doi.org/10.5194/ascmo-5-101-2019>
- Dione, C., Haeffelin, M., Burnet, F., Lac, C., Canut, G., Delanoë, J. et al. (2023) Role of thermodynamic and turbulence processes on the fog life-cycle during sofog3d experiment. *Atmospheric Chemistry and Physics*, 23(24), 15711–15731. Available from: <https://doi.org/10.5194/acp-23-15711-2023>
- Dupont, J.C., Haeffelin, M., Protat, A., Bouniol, D., Boyouk, N. & Morille, Y. (2012) Stratus-fog formation and dissipation: a 6-day case study. *Boundary-Layer Meteorology*, 143(1), 207–225. Available from: <https://doi.org/10.1007/s10546-012-9699-4>
- Duynkerke, P.G. (1991) Radiation fog: a comparison of model simulation with detailed observations. *Monthly Weather Review*, 119, 324–341.
- Egli, S., Thies, B. & Bendix, J. (2019) A spatially explicit and temporally highly-resolved analysis of variations in fog occurrence over Europe. *Quarterly Journal of the Royal Meteorological Society*, 145(721), 1721–1740. Available from: <https://doi.org/10.1002/qj.3522>
- Egli, S., Thies, B., Droöner, J., Cermak, J., Bendix, J., Thies, B. et al. (2017) A 10 year fog and low stratus climatology for Europe based on Meteosat second generation data. *Quarterly Journal of the Royal Meteorological Society*, 143(702), 530–541. Available from: <https://doi.org/10.1002/qj.2941>
- Estrela, M.J., Valiente, J.A., Corell, D. & Millán, M.M. (2008) Fog collection in the western Mediterranean basin (Valencia region, Spain). *Atmospheric Research*, 87(3–4), 324–337. Available from: <https://doi.org/10.1016/j.atmosres.2007.11.013>
- Everitt, B. (Ed.). (2011) *Cluster analysis*, 5th edition. Chichester: Wiley.
- Fallmann, J., Lewis, H., Sanchez, J.C. & Lock, A. (2019) Impact of high-resolution ocean-atmosphere coupling on fog formation over the North Sea. *Quarterly Journal of the Royal Meteorological Society*, 145(720), 1180–1201. Available from: <https://doi.org/10.1002/qj.3488>
- Fuchs, J., Andersen, H., Cermak, J., Pauli, E. & Roebeling, R. (2022) High-resolution satellite-based cloud detection for the analysis of land surface effects on boundary layer clouds. *Atmospheric Measurement Techniques*, 15(14), 4257–4270. Available from: <https://doi.org/10.5194/amt-15-4257-2022>
- Fuzzi, S., Facchini, M.C., Orsi, G., Lind, J.A., Wobrock, W., Kessel, M. et al. (1992) The Po Valley fog experiment 1989 an overview. *Tellus B: Chemical and Physical Meteorology*, 44(5), 448. Available from: <https://doi.org/10.3402/tellusb.v44i5.15561>
- Gautam, R. & Singh, M.K. (2018) Urban Heat Island over Delhi punches holes in widespread fog in the indo-Gangetic Plains. *Geophysical Research Letters*, 45(2), 1114–1121. Available from: <https://doi.org/10.1002/2017GL076794>
- Glickman, T. (2000) *Glossary of meteorology*, 2nd edition. Boston: American Meteorological Society.
- Gottlieb, T.R., Eckardt, F.D., Venter, Z.S. & Cramer, M.D. (2019) The contribution of fog to water and nutrient supply to *Arthroa leubnitziae* in the central Namib Desert, Namibia. *Journal of Arid Environments*, 161, 35–46. Available from: <https://doi.org/10.1016/j.jaridenv.2018.11.002>

- Guerreiro, P.M., Soares, P.M., Cardoso, R.M. & Ramos, A.M. (2020) An analysis of fog in the mainland Portuguese international airports. *Atmosphere*, 11(1239), 1–19. Available from: <https://doi.org/10.3390/atmos11111239>
- Gultepe, I., Tardif, R., Michaelides, S.C., Cermak, J., Bott, A., Bendix, J. et al. (2007) Fog research: A review of past achievements and future perspectives. *Pure and Applied Geophysics*, 164, 1121–1159. Available from: <https://doi.org/10.1007/s00024-007-0211-x>
- Haeffelin, M., Bergot, T., Elias, T., Tardif, R., Carrer, D., Chazette, P. et al. (2010) PARISFOG: shedding new light on fog physical processes. *Bulletin of the American Meteorological Society*, 91(6), 767–783. Available from: <https://doi.org/10.1175/2009BAMS2671.1>
- Hartigan, J.A. & Wong, M.A. (1979) A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), 100–108.
- Hersbach, H. (2016) The ERA5 atmospheric reanalysis. AGU fall meeting abstracts, 12–16 December 2016, San Francisco.
- Hůňová, I. (2020) Ambient air quality in the Czech Republic: past and present. *Atmosphere*, 11(214), 1–26. Available from: <https://doi.org/10.3390/atmos11020214>
- Karimi, M. (2020) Direct numerical simulation of fog: the sensitivity of a dissipation phase to environmental conditions. *Atmosphere*, 11(12), 1–22. Available from: <https://doi.org/10.3390/ATMOS11010012>
- Kneringer, P., Dietz, S.J., Mayr, G.J. & Zeileis, A. (2019) Probabilistic now casting of low-visibility procedure states at Vienna international airport during cold season. *Pure and Applied Geophysics*, 176(5), 2165–2177. Available from: <https://doi.org/10.1007/s00024-018-1863-4>
- Knerr, I., Trachte, K., Egli, S., Barth, J.A., Bräuning, A., Garel, E. et al. (2021) Fog - low stratus (FLS) regimes on Corsica with wind and PBLH as key drivers. *Atmospheric Research*, 261, 1–12. Available from: <https://doi.org/10.1016/j.atmosres.2021.105731>
- Köhler, C., Steiner, A., Saint-Drenan, Y.-M., Ernst, D., Bergmann-Dick, A., Zirkelbach, M. et al. (2017) Critical weather situations for renewable-energies – part B: low stratus risk for solar power. *Renewable Energy*, 101, 794–803. Available from: <https://doi.org/10.1016/j.renene.2016.09.002>
- Leigh, R.J., Drake, L. & Thampapillai, D.J. (1998) An economic analysis of terminal aerodrome forecasts with special reference to Sydney airport. *Journal of Transport Economics and Policy*, 32(3), 377–392.
- Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B. et al. (2020) From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56–67. Available from: <https://doi.org/10.1038/s42256-019-0138-9>
- Lundberg, S.M. & Lee, S.-I. (2017) A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774. Available from: <https://doi.org/10.48550/arXiv.1705.07874>
- Mayes, J. (2013) Regional weather and climates of the British Isles - part 5: Wales. *Weather*, 68(9), 227–232. Available from: <https://doi.org/10.1002/wea.2149>
- Mitchell, D., Henschel, J.R., Hetem, R.S., Wassenaar, T.D., Strauss, W.M., Hanrahan, S.A. et al. (2020) Fog and fauna of the Namib Desert: past and future. *Ecosphere*, 11(1), 1–40. Available from: <https://doi.org/10.1002/ecs2.2996>
- Molnar, C. (2019) Interpretable machine learning. <https://christophm.github.io/interpretable-ml-book/>
- Müllner, D. (2011) Modern hierarchical, agglomerative clustering algorithms. arXiv <https://doi.org/10.48550/ARXIV.1109.2378>
- Müllner, D. (2013) Fastcluster: fast hierarchical, agglomerative clustering routines for R and python. *Journal of Statistical Software*, 53(9), 1–18. Available from: <https://doi.org/10.18637/jss.v053.i09>
- Oreopoulos, L., Cho, N., Lee, D., Kato, S. & Huffman, G.J. (2014) An examination of the nature of global MODIS cloud regimes. *Journal of Geophysical Research Atmospheres*, 119, 8362–8383. Available from: <https://doi.org/10.1038/175238c0>
- Pagowski, M., Gultepe, I. & King, P. (2004) Analysis and Modeling of an extremely dense fog event in southern Ontario. *Journal of Applied Meteorology*, 43, 3–16.
- Pauli, E., Andersen, H., Bendix, J., Cermak, J. & Egli, S. (2020) Determinants of fog and low stratus occurrence in continental central Europe – a quantitative satellite-based evaluation. *Journal of Hydrology*, 591, 1–15. Available from: <https://doi.org/10.1016/j.jhydrol.2020.125451>
- Pauli, E., Cermak, J. & Andersen, H. (2021) Satellite-based fog and low stratus-cloud formation and dissipation times data set. KIT-Bibliothek, KITopenData <https://doi.org/10.5445/IR/1000141293>
- Pauli, E., Cermak, J. & Andersen, H. (2022a) A satellite-based climatology of fog and low stratus formation and dissipation times in central Europe. *Quarterly Journal of the Royal Meteorological Society*, 148(744), 1439–1454. Available from: <https://doi.org/10.1002/qj.4272>
- Pauli, E., Cermak, J. & Teuling, A.J. (2022b) Enhanced Nighttime fog and low stratus occurrence over the Landes Forest, France. *Geophysical Research Letters*, 49, 1–9. Available from: <https://doi.org/10.1029/2021GL097058>
- Pérez-Díaz, J.L., Ivanov, O., Peshev, Z., Álvarez-Valenzuela, M.A., Valiente-Blanco, I., Evgenieva, T. et al. (2017) Fogs: physical basis, characteristic properties, and impacts on the environment and human health. *Water (Switzerland)*, 9(10), 1–21. Available from: <https://doi.org/10.3390/w9100807>
- Pohl, M.J., Lehnert, L., Bader, M.Y., Gradstein, S.R., Viehweger, J. & Bendix, J. (2021) A new fog and low stratus retrieval for tropical South America reveals widespread fog in lowland forests. *Remote Sensing of Environment*, 264, 1–14. Available from: <https://doi.org/10.1016/j.rse.2021.112620>
- Price, J., Porson, A. & Lock, A. (2015) An observational case study of persistent fog and comparison with an ensemble forecast model. *Boundary-Layer Meteorology*, 155(2), 301–327. Available from: <https://doi.org/10.1007/s10546-014-9995-2>
- Roach, W. (1995) Back to basics: Fog: part 2 - the formation and dissipation of land fog. *Weather*, 50(1), 7–11.
- Román-Cascoón, C., Steeneveld, G.J., Yagüe, C., Sastre, M., Arrilaga, J.A. & Maqueda, G. (2016) Forecasting radiation fog at climatologically contrasting sites: evaluation of statistical methods and WRF. *Quarterly Journal of the Royal Meteorological Society*, 142(695), 1048–1063. Available from: <https://doi.org/10.1002/qj.2708>
- Royé, D., Lorenzo, N., Rasilla, D. & Martí, A. (2019) Spatio-temporal variations of cloud fraction based on circulation types in the Iberian Peninsula. *International Journal of Climatology*, 39(3), 1716–1732. Available from: <https://doi.org/10.1002/joc.5914>

- Scherrer, S.C. & Appenzeller, C. (2014) Fog and low stratus over the Swiss plateau—a climatological study. *International Journal of Climatology*, 34(3), 678–686. Available from: <https://doi.org/10.1002/joc.3714>
- Steenefeld, G.J. & de Bode, M. (2018) Unravelling the relative roles of physical processes in modelling the life cycle of a warm radiation fog. *Quarterly Journal of the Royal Meteorological Society*, 144(714), 1539–1554. Available from: <https://doi.org/10.1002/qj.3300>
- Strobl, C., Boulesteix, A.L., Zeileis, A. & Hothorn, T. (2007) Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics*, 8. Available from: <https://doi.org/10.1186/1471-2105-8-25>
- Tselioudis, G., Rossow, W., Zhang, Y. & Konsta, D. (2013) Global weather states and their properties from passive and active satellite cloud retrievals. *Journal of Climate*, 26(19), 7734–7746. Available from: <https://doi.org/10.1175/JCLI-D-13-00024.1>
- Tzallas, V., Hünerbein, A., Stengel, M., Meirink, J.F., Benas, N., Trentmann, J. et al. (2022) CRAAS: a European cloud regime dAtAset based on the CLAAS-2.1 climate data record. *Remote Sensing*, 14(5548), 1–25. Available from: <https://doi.org/10.3390/rs14215548>
- Valiente, J.A., Estrela, M.J., Corell, D., Fuentes, D., Valdecantos, A. & Baeza, M.J. (2011) Fog water collection and reforestation at a mountain location in a western mediterranean basin region: air-mass origins and synoptic analysis. *Erdkunde*, 65(3), 277–290. Available from: <https://doi.org/10.3112/erdkunde.2011.03.05>
- Wærsted, E.G., Haeffelin, M., Steeneveld, G.-J. & Dupont, J.-C. (2019) Understanding the dissipation of continental fog by analysing the LWP budget using idealized LES and in situ observations. *Quarterly Journal of the Royal Meteorological Society*, 145(719), 784–804. Available from: <https://doi.org/10.1002/qj.3465>
- Wilks, D.S. (2006) *Statistical Methods in the Atmospheric Sciences*. <https://doi.org/10.1016/B978-0-12-823043-5.00015-1>
- Wobrock, W., Schell, D., Maser, R., Kessel, M., Jaeschke, W., Fuzzi, S. et al. (1992) Meteorological characteristics of the Po Valley fog. *Tellus B: Chemical and Physical Meteorology*, 44(5), 469. Available from: <https://doi.org/10.3402/tellusb.v44i5.15562>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pauli, E., Cermak, J., Andersen, H. & Fuchs, J. (2024) An analysis of fog and low stratus life-cycle regimes over central Europe. *Quarterly Journal of the Royal Meteorological Society*, 1–15. Available from: <https://doi.org/10.1002/qj.4714>