Fractal Characteristics of Ice-Supersaturated Regions in the Tropopause Region of the Northern Midlatitudes

Helena Schuh[1], Philipp Reutter[1], Stefan Niebler[2,3], and Peter Spichtinger[1]

[1]Institute for Atmospheric Physics, Johannes Gutenberg University Mainz, Mainz, Germany
[2]Institute for Computer Sciences, Johannes Gutenberg University Mainz, Mainz, Germany
[3]now at NVIDIA, Würselen, Germany

Correspondence: Helena Schuh (hschuh01@students.uni-mainz.de)

Abstract. Ice supersaturated regions (ISSRs) are air masses in the upper troposphere and lower stratosphere (UTLS) where saturation ratio over ice (S_i) exceeds ~~1~~one, i.e. regions with enhanced water vapor concentrations. These are potential formation regions of cirrus clouds and contrails. While the impact of cloud free regions of enhanced water vapor on the planetary radiation balance is small to negligible, thin cirrus clouds and aircraft induced contrail cirrus formed within them might have a large radiative impact. Understanding the characteristics of ISSRs, including their geometry and seasonal variability, is essential for ~~improving~~ evaluating atmospheric models in representing ice clouds correctly. While ISSR's pathlength statistics, i.e. 1D characteristics, have already been studied, their geometric properties, particularly fractal properties such as self-similarity, and their seasonal variability remain largely unexplored. We identify ISSRs using ERA5 reanalysis data spanning from 2010 to 2020 at three pressure levels. An area-perimeter method is employed to compute fractal dimensions. The results reveal slopes ~~equaling fractal dimensionswith high coefficients of determination,~~ implying fractal dimensions, strongly suggesting that ISSRs in the UTLS exhibit fractal behavior. A seasonal cycle in total number and area of ISSRs, as well as in the fractal dimension is found, in combination with a strong vertical variation. We hypothesize that this is caused by the seasonal variation of convective and frontal activity. We further analyzed the ~~latitudinal and longitudinal extensions~~ zonal and meridional extents of ISSRs as well as the pathlengths of modeled flights along commercial flight routes. The results of these horizontal ~~extensions~~ extents are consistent with the fractal properties, and suggest distinct formation processes for ISSRs.

1

Ice supersaturated regions (ISSRs) in the upper troposphere and lower stratosphere (UTLS) serve as key environments for the formation and persistence of in situ cirrus clouds. These regions occur when the water vapor partial pressure p_v exceeds the water vapor saturation pressure over ice at temperature T, leading to a saturation ratio

$$S_i = \frac{p_v}{p_{si}(T)} \tag{1}$$

greater than ~~1. Ice supersaturation (first described by Alfred)~~one. Ice supersaturation (first described by Alfred **??**) is a frequent and widespread phenomenon in the UTLS (**??????**), where the cold and dry conditions render even small variations in humidity highly significant. ISSRs not only serve as a key environment for the formation of in situ cirrus clouds (**?**), but also as a prerequisite for persistent contrails ~~; this is nowadays~~ (?????). Currently, this is a major topic in the context of contrail avoidance, since ~~these~~ contrails can persist for over 10 hours and spread into contrail cirrus (?). Cloud free ISSRs ~~themselves~~ have a relatively minor direct influence on the local radiation budget, ~~although~~ but their transition into cirrus clouds significantly alters radiative forcing (?). ISSRs and cirrus clouds are predominantly found in the upper troposphere, but they can also extend above the tropopause (??), affecting the lower stratosphere (LS). The upper troposphere–lower stratosphere (UTLS) region is characterized by extremely cold and dry air masses (??)~~, making outgoing longwave radiation~~. Outgoing longwave radiation is particularly sensitive to small relative perturbations in UTLS water vapor ~~, due to the very low background concentrations~~ (?)~~; gradients~~. Gradients in saturation ratio are closely related to the respective changes in longwave radiation (e.g. ?), potentially resulting into local changes of the thermodynamic structure of the tropopause region (?). Consequently, variations in ISSR occurrence and cirrus cloud formation within this region have important implications for climate dynamics and radiative balance (?). ISSRs are not necessarily cloud free air masses ~~, although former investigations suggest~~ (?), despite former investigations (e.g. ?) suggesting this property. ~~However, the~~ The transition from ice clouds to cloud free air is very smooth and usually there are no sharp boundaries. In our study we will not further distinguish between clear and cloudy air, as we are interested in the thermodynamic state of the air masses. While studies have examined the horizontal extent of ISSRs using pathlength statistics (????), their two-dimensional spatial characteristics, including shape and scaling behavior, remain poorly understood and ~~were not~~ have not been studied in detail yet. Statistical investigations are challenging because of the ISSRs' variability in size and shape. In addition, some metrics must be defined (or even developed) for studying ISSRs as 2D/3D objects, instead of just using point measurements. The radiative effects of cirrus clouds, ~~including those related to~~ related to either natural or anthropogenic origins, are a subject of ongoing research. ~~However, this~~ This study emphasizes the structural and statistical properties of ISSRs. Understanding their morphology and variability is essential for improving microphysical parameterizations and interpreting the atmospheric processes that control cloud evolution in the UTLS.

A potential approach in studying the geometric properties of 2D or 3D objects is to examine their potential self-similarity, i.e. the fractal properties of these macroscopic objects. In general, fractals are geometric structures characterized by self-similarity. These objects exhibit similar patterns on different scales~~, and they~~. They differ from traditional Euclidean shapes by possessing non-integer (fractional) dimensions that describe the way their complexity changes with scale (?). ~~Although the whole~~ The research field of fractal geometry was defined by Mandelbrot in 1975, and became ~~famous~~ prominent in the 1980s (e.g., ?)~~, investigations~~. Investigations on the fractal nature of natural objects were made before Mandelbrot. Notably, **?** carried out first investigations on the nature of turbulence, ~~addressing already some~~ already addressing fractal properties. Thus, fractal analysis has a longstanding tradition in atmospheric physics, especially in cloud physics. In numerous previous studies, the fractal nature of clouds is investigated, ranging from cumulus cloud statistics and properties (?????), rain and hail clouds (???), tropical clouds (??), noctilucent clouds (??), and cloud fields in high resolution models (?). Additionally, fractal analysis is a key technique for turbulence theories and data evaluation (e.g., ??). In some studies, clouds and turbulence investigations are combined in order to explain fractal properties of clouds by the underlying turbulence (or diffusion) characteristics (e.g., ??). Overall, there is strong evidence that clouds at different vertical levels show fractal behavior. Thus, it is reasonable to question whether their potential formation regions, i.e. the ISSRs ~~in case of~~ for ice clouds, ~~also~~ exhibit similar properties. If this holds true and ISSRs also display self-similar characteristics, determining their fractal dimension can provide insight into their spatial attributes and characteristics. ~~In addition, these properties should be represented in models on different scales; thus, such properties might serve as indication for correct representation of ISSRs and cirrus in model parameterizations.~~

To identify and study ISSRs, we analyze ERA5 reanalysis data that incorporates absolute humidity, temperature and pressure~~in terms of ice supersaturation. For the investigation we distinguish water vapor in air masses in the state of sub or super saturation with respect to ice by applying a threshold of $S_{i,\mathrm{threshold}} = 0.9$, already taking into account the dry bias of ERA5 in the tropopause region . A grid point with values $S_i < S_{i,\mathrm{threshold}}$ is marked as subsaturated (index=0), and grid points with values $S_i \geq S_{i,\mathrm{threshold}}$ are counted as supersaturated (i.e. ISSR, index=1).~~ . ~~As a result a~~ Working with a simplified binary representation of ~~these regions on the grid is obtained and~~ ISSRs, connected areas are identified and ~~determined as ISSRs. Based on this consideration the actual~~ their areas and perimeters ~~of ISSRs are calculatedto accurately describe their shapes. For a gridded data set, the spatial resolution is inherently limited; therefore, a rigorous derivation of a fractal dimension is not feasible. However, to assess the the~~ are calculated. To assess the fractal characteristics, ~~and~~ especially the self-similarity of ISSRs, we examine the area-perimeter relation. ~~A scatter plot of measured perimeters against the areas of individual ISSRs on logarithmic scales reveals a linear relationship with slope α, which can be interpreted as the fractal (or self-similar) dimension $D = 2\alpha$. This~~ This relation and variations of it have been used in many investigations of clouds (**?????**). Such an analysis provides a quantitative measure of the macroscopic structures and has important implications for further investigations. ~~The derived fractal properties of ISSRs might be helpful for the evaluation of atmospheric models, since for a correct representation of ISSRs in weather and climate models the self-similar properties should be comparable. For instance, forecasts of ISSRs in terms of contrail avoidance must represent the fractal nature of these regions. Thus, our results could serve as a meaningful test of models regarding the representation of ice supersaturation in the UTLS.~~

The study is structured as follows: In Sec. 2 the data set and the methods are introduced. In Sec. 3 the fractal characteristics of ISSRs as found in the data are investigated. This is followed in Sec. ~~??~~ **??** by the investigation on the horizontal ~~extension~~ extent of ISSRs. The results are discussed ~~in Sec. ??~~ and conclusions are drawn in Sec. **??**. Additional results are presented in the supplement.

2 Data and Methods

In this section we present the data set and describe the methods used for the investigation of the data.

ISSRs are naturally occurring physical phenomena whose boundaries are inherently gradual rather than sharply defined. Their identification using gridded reanalysis fields necessarily imposes discrete representations of these continuous structures. Although these representations are not identical to the physical structures in nature, they provide a consistent basis for quantifying the geometry and occurrence of physically meaningful supersaturated regions.

2.1 Data set ERA5

The study analyzes eleven years of ERA5 reanalysis data (**?**) covering the period from January 1, 2010, to December 31, 2020. Data is available at the standard synoptic times each day (0, 6, 12, 18 UTC), resulting in a total of 16,060 datasets for analysis. ERA5 provides data with a horizontal resolution of 0.25 degrees in longitude and latitude, and a vertical resolution of 137 model levels with an approximate resolution in the tropopause region of about $300\,\mathrm{m}$. The data set contains variables of 3D winds (u, v, w), temperature T, and specific humidity q. For the investigation we distinguish water vapor in air masses in the state of sub or super saturation with respect to ice. The geographical area used for the analysis is narrowed down to range from 80 to 30 degrees latitude in the Northern hemisphere in order to exclude the tropics and polar regions. In our investigation we focus on pressure levels of 200 hPa, 250 hPa, and 300 hPa, representing the upper troposphere and tropopause region (similarly to the investigation by **?**)~~; the~~. The data on adjacent model levels were interpolated on the respective isobaric surfaces.

2.2 Methods

Here, we briefly describe the processing of the data; for details we refer to appendix **??**.

2.2.1 Relevant variable

The thermodynamic order parameter for ISSRs is the saturation ratio

$$S_i = \frac{p \cdot q}{\epsilon \cdot p_{si}(T)} = \frac{p_v}{p_{si}(T)} \tag{2}$$

as calculated from the ERA5 variables p, T, q, using the ratio of molar masses of water and air, i.e. $\epsilon = \frac{M_v}{M_d}$ and the saturation vapour pressure over (hexagonal) ice p_{si} in the formulation by **?**, i.e.,

$$p_{si} = \exp\left(9.5504 - \frac{5723.265}{T} + 3.53068\log(T) - 0.00728332\,T\right), \tag{3}$$

Here, T is given in Kelvin and p_{si} is obtained in Pascal~~; the~~. The formulation is valid for temperatures $T > 100$ K. The thermodynamic equilibrium between vapor and ice is given by $S_i = 1$.

~~However, since~~ ERA5 data ~~underestimate~~ underestimates the humidity in the UTLS region (**?**)~~, we choose~~. We counteract this by choosing a threshold value of $S_{i,\mathrm{threshold}} = 0.9$. Note that we use no further information about possible clouds inside the ISSRs. A data point is assigned to be supersaturated (i.e. belonging to an ISSR) if $S_i \geq S_{i,\mathrm{threshold}} = 0.9$. Using this criterion we create a two dimensional binary field with subsaturated data points (index = 0) and data points inside an ISSR (index = 1) for each pressure level. For a macroscopic object we have to specify a measure for the connectivity of single data points (with index = 1) on the gridded isobaric surface. We choose the 8-neighborhood (or Moore neighborhood), i.e. a central pixel is connected to the neighboring pixels via the edges and the diagonals. This definition, originating from investigations with cellular automaton (**?**), is essential, since the connectivity of a physical object should not (or as little as possible) depend on the orientation of a grid. A visualization of ~~the Moore neighborhood~~ different neighborhoods is represented in appendix **??**. For a substantial analysis, we consider only ISSRs consisting of at least $N = 24$ pixels, i.e. exceeding a minimum size~~; otherwise~~.

Otherwise, a possible self-similar structure would not be reliably resolved. Here, we follow the recommendation ~~by~~ of **?**. Furthermore, we ~~concentrate~~ focus on the midlatitudes in the Northern Hemisphere ($30-80°$N). As ISSRs may extend into polar or tropical regions, we exclude those objects crossing the boundaries of 30°N and 80°N completely (setting the index = 0) to avoid ~~artificialyl~~ artificially smoothing perimeter and cropping area. For the further evaluation, the resulting connected binary islands (i.e the ISSRs) are numbered.

2.2.2 Measuring Island Area and Perimeter

The ISSRs are located on the geographical grid of the Earth. Since we use a latitude-longitude description we have to take into account the distortion caused by the curvature of the Earth. Each grid point can be seen as the center of an almost quadratic pixel with a respective area (in square meters, see appendix **??**), which can be assigned by using the Python package `pyproj.Geod` (**?**). For the calculation of the perimeter, we had to develop an additional tool, taking into account the length of the respective boundaries of the pixel without a neighboring pixel. The overall perimeter of the labeled island is calculated iteratively; here, we can determine both ~~,~~ internal and external borders. ~~It should be noted that this method leads to a minor overestimation of perimeter values for gridded objects as compared to their real (potentially smooth) counterparts, as represented on a grid. This issue is discussed in appendix **??**. Further details of the method can be found in appendix **??**.~~

2.2.3 Calculation of the Fractal Dimension

A key characteristic of fractals is self-similarity, the defining property upon which we base our approach for the dimension of objects. The fractal dimension most commonly refers to the Hausdorff dimension~~; however~~. However, its rigorous determination lies beyond the scope of this work. ~~The~~ As the general concept involves progressively increasing the spatial resolution of the data ~~, i.e. examining smaller length scales and measuring geometric properties (such as length, area, or volume) using successively smaller measuring elements (e.g. circles or spheres). In the limit of this sequence the so-called Hausdorff dimension is obtained, which is a positive real number $D_H \in \mathbb{R}_+$~~ (**?**)~~. As a consequence, the perimeter of fractal objects diverges and tends to infinity in the limit. This procedure is often~~ , this procedure is not feasible for real-world data due to the finite resolution~~, in particular for~~ . Specifically, gridded data sets such as ERA5 ~~, where~~ are incompatible with this method, as the resolution is fixed and cannot be refined arbitrarily.

Consequently, a different method ~~is~~ must be employed to calculate the dimension for data in ~~two-dimensions~~two dimensions. For this purpose, the area-perimeter relation is investigated. This type of relationship was introduced by **?** to study fractal properties of clouds and rain areas, i.e. the self-similarity of such objects. It is defined as ~~follows:~~ a power law relationship between the perimeter P and the area A of a two dimensional object is assumed, such that

$$P = C \cdot A^{\frac{D}{2}} \iff \log(P) = \log(C) + \frac{D}{2}\log(A) \tag{4}$$

with C a scaling parameter. ~~Thus, using~~ Using scatter plots of data points (A, P) in a double logarithmic scale, the perimeter shows a linear behavior with respect to the area with a slope α~~, which~~ . α is not necessarily an integer number but rather a non-negative real number $\alpha \in \mathbb{R}$. This can be used for calculating the "fractal" dimension $D = 2 \cdot \alpha$. For regular geometric objects we ~~would~~ obtain a value $D = 1$ (see also the discussion in appendix **??**)~~; however, for~~. D should be understood as the dimension of the perimeter. For planar, fractal objects ~~as e.g. ISSRs on a~~ like ISSRs on an isobaric surface we assume $1 < D < 2$. The value of D can be determined by a linear fit with slope α.

As observed by **?**, the "dimension" D as derived by the area-perimeter relation does not agree with the Hausdorff dimension in general. ~~This raises the question if the use of this relation and thus the obtained value D is meaningful at all. However, this method is frequently employed to ascertain fractal dimensions and analyze the respective objects. Already found that the dimension D obtained by area-perimeter relations agrees with fractal dimensions obtained by spectral analysis. also discussed this issue, and concluded that the method is "the most stable and least ambiguous method" for obtaining fractal dimensions. In other methods for determining fractal dimensions are reviewed. While the dimension D calculated using the area-perimeter relation (Eq. (4)) does not satisfy the mathematical definition of the fractal dimension,~~ However, it is appropriate to use this approach in order to study the fractal characteristics and properties of phenomena ~~. As these naturally occurring phenomena are not perfect fractals, a pragmatic approach such as the area-perimeter relation is sufficient for a relevant evaluation. Finally, showed that for island-type objects, such as our ISSRs, the area-perimeter approach is meaningful and results into dimensions close to the classical fractal dimensions.The area-perimeter relation and variations of it have been used in many applications, as e.g. to study the fractal characteristics of rain and cloud areas , to determine the fractal perimeter dimension of noctilucent clouds , to measure the fractal structure of interstellar clouds , and to classify tropical clouds by their fractal dimension~~ as shown in other studies (e.g., **????**, for details see supplement).

For the ERA5 data in this study (as for all gridded data), the perimeter cannot approach infinity as the horizontal resolution limits the frayed nature of the ISSRs' perimeter, further justifying the employment of the area-perimeter method. The dimension calculated with the area-perimeter method will hereafter be referred to as the fractal dimension. As ISSRs have irregular boundaries, the dimension is expected to be in the range $1 < D < 2$. In appendix **??** we exemplarily show how the fractal dimension might change by either changing the shape of rectangulars systematically or fraying the boundary of squares.
For gridded data, area and perimeter can only be approximated. Possible errors are underestimation of the area and overestimation of the perimeter, as discussed in **?** and **?**. In appendix **??** we investigate these issues for spheres. The analysis indicates that the quality of the evaluation is not compromised. The derived fractal dimensions (i.e. exponents in Eq. (4)) can therefore be regarded as robust estimates and are suitable for further, including quantitative, analysis.

3 Fractal properties of ISSRs

In this section we investigate the ~~fractal~~ properties of ISSRs, ~~i.e. the derived fractal dimension D. Subsequently, more standard properties,~~ such as the number and area of ISSRs, ~~as well as~~ and the derived fractal dimension D. Subsequently, their seasonal and annual evolution are investigated.

3.1 ISSR Number Count and Area

Figure **??** shows the number of observed ISSRs by month on the 200 hPa, 250 hPa and 300 hPa pressure levels during the period from 2010 to 2020. The total number is a sum of all ISSRs appearing at each point in time. This causes the number to be significantly higher than the ISSRs that appeared in a given month, as ISSRs are counted multiple times when having a lifespan longer than 6 hours. A clear dependence of the number of ISSRs with altitude is evident. At lower altitudes (i.e. higher pressures) we find significantly more ISSRs than at higher altitudes. This can be explained by the fact that in the extratropics the (thermal) tropopause is located in the pressure range between 300 and 200 hPa, and ISSRs are almost exclusively located in the troposphere~~; only~~ . Only few ISSRs seem to reach into the stratosphere (**?**). Generally, ISSRs are most frequently located close to the tropopause.

The tropopause also shows a seasonal cycle with high altitudes during summer and lower altitudes during winter (e.g., **?**). Nevertheless, the probability to measure dry and warm air in the stratosphere at a certain pressure level in the extratropics is much higher for pressure level $p = 200\,\mathrm{hPa}$, than for $p = 250\,\mathrm{hPa}$ or $p = 300\,\mathrm{hPa}$. This might explain the strong differences in the amount of ISSRs between the three pressure levels. The strong seasonal cycle in the number of ISSRs is a prominent feature, which is most pronounced at pressure level $p = 300\,\mathrm{hPa}$, i.e. in the upper troposphere. We find maxima in summer and minima in winter for all 3 pressure levels. The month with the highest observed number at pressure level $p = 300\,\mathrm{hPa}$ is July 2013 with 9511 detected ISSRs, the minimum number of 3290 ISSRs at the same pressure level can be observed in February 2018. Respectively, the highest number of ISSRs at pressure level $p = 200\,\mathrm{hPa}$ was recorded in July 2020 with 3462 ISSRs, whereas the minimum was observed in February 2012 with 954 ISSRs. Here, we see the substantial spread in the amount of ISSRs.

The seasonal cycle compares well with former investigations. **?** also found this signature with highest values of ISSR numbers during summer and lowest counts in the winter's season. Furthermore, the strong variation of numbers with the vertical levels agrees ~~also~~ well with our findings.

In addition to the total number of ISSRs we investigate the total area of the detected ISSRs. ~~As for~~ Analogous the number of ISSRs, in Figure **??** the total area of observed ISSRs by month on the 200 hPa, 250 hPa and 300 hPa pressure levels in the timespan from 2010 to 2020 is shown. As for the evaluation of the total number, the area might be enhanced by counting long-lived ISSRs multiple times. The total area is largest at the lowest altitude (i.e. pressure level $p = 300\,\mathrm{hPa}$), and is decreasing with increasing altitude. The highest total area in one month was observed in May 2019 with a summed area of 2.65 $\cdot 10^9\,\mathrm{km}^2$ at pressure level $p = 300\,\mathrm{hPa}$. On the same pressure level the minimum of 1.32 $\cdot 10^9\,\mathrm{km}^2$ is found in February 2018. At the highest altitude level (pressure level $p = 200\,\mathrm{hPa}$) the area is strongly reduced ~~as~~ compared to the other two levels, which have comparable area values. At $p = 200\,\mathrm{hPa}$ the month with the highest total area of supersaturation in respect to ice is July 2020, summing up to 1.12 $\cdot 10^9\,\mathrm{km}^2$. February 2015 at this pressure level constitutes the absolute minimum observed on all pressure levels with 0.41 $\cdot 10^9\,\mathrm{km}^2$. ~~Analogous to~~ Like the total number, the area also shows a seasonal cycle~~, although it is not as strongly pronounced~~. We find maximum values in the summer months, and minimum values in the winter months, similarly to the investigations of the total number of ISSRs. The cycle is most pronounced in the two lower levels ($p = 300\,\mathrm{hPa}$ and

$p = 250\,\mathrm{hPa}$), while for the highest level ($p = 200\,\mathrm{hPa}$) the variation is weak but still apparent. In summer, the total area of ISSRs at $p = 250\,\mathrm{hPa}$ is approaching values comparable to ($p = 300\,\mathrm{hPa}$. This comparability only holds for the area maximum in summer, whereas the area decreases more strongly at $p = 250\,\mathrm{hPa}$ compared to $p = 300\,\mathrm{hPa}$ in winter.

In Figure **??** examples of ISSRs at the three different pressure levels over the North Atlantic region are presented for three times in January 2010, giving a snapshot. A similar example is shown in Figure **??** for one snapshot in July 2019 (3 times). In winter the ISSRs are discernibly more compact and cover a limited area (Fig. **??**). At the highest altitudes (i.e. at pressure level $p = 200\,\mathrm{hPa}$) there are very few ISSRs ~~although~~ while at lower altitudes (higher pressures) more ISSRs can be seen~~; in~~. In winter the tropopause is located at lower altitudes, thus the pressure level $p = 200\,\mathrm{hPa}$ is (most likely) completely contained in the stratosphere. In contrast, in July we find widespread and frayed ISSRs at all three pressure levels (Fig. **??**). Even at $p = 200\,\mathrm{hPa}$ ISSRs can be found, since during summer the tropopause is located at higher altitudes and ice supersaturation is present there.

~~The different~~ To gain a better understanding of the correlation between ISSR count and area in relation to seasons, we calculated the mean ISSR area from the corresponding number and area statistics. The results are shown as timeseries in Figure **??**.

We see a pronounced seasonal cycle in the mean area of ISSRs with a very distinct maximum in winter (mean values $\bar{A} \sim 5 - 6 \cdot 10^{11}\mathrm{m}^2$) and a clear minimum in summer (mean values $\bar{A} \sim 2.5 - 3 \cdot 10^{11}\mathrm{m}^2$). Assuming a purely spherical shape, this would translate into a factor of 2 for the radii. This behavior hints at two different formation mechanisms dominating in summer and winter respectively.

While the average area per ISSR on the 200 hPa and 300 hPa levels show good agreement across all seasons, this ratio is noticeably larger at 250 hPa during summer. This feature does not represent an outlier but can be explained by jointly considering the number count and total area. In winter, both the number count and total area at 250 hPa are substantially lower than at 300 hPa. During summer, the total area at both levels is comparable, despite a lower number count at 250 hPa. Consequently, the ratio of total area to number count is increased at 250 hPa in summer. The underlying cause of this feature is the 250 hPa level being recurrently located within both the upper troposphere and lower stratosphere depending on the season. In contrast, 200 hPa remains predominantly above and 300 hPa predominantly below the tropopause throughout the year (**?**).

The different shape of ISSRs in winter and summer, as observed in Figures **??** and **??**, can also be seen in the (statistical) investigations of the fractal dimension, as described in the next section.

3.2 ISSR Dimension

~~Next, the fractal dimension~~ The fractal dimensions for all ISSRs in the data set are derived using the methods explained in section 2.2 and appendix **??**. We calculate area and perimeter for each detected ISSR (i.e. each connected "island") and determine the fractal dimension by using all data of a certain pressure level for a whole month in a double logarithmic scatter plot. In Figure **??** an example is shown for January 2017~~(in appendix ?? another example is~~, in the supplement another example provided for July 2017).

In the scatter plots the data ~~are~~ is confined to a linear point cloud. There is a perceptible ~~scale break at perimeters of $P \sim 10^7\,\mathrm{m}$, separating two regimes. However,~~ change in slope at areas of $A \sim 2 \cdot 10^{12}\,\mathrm{m}^2$. As this signature is weak, ~~so~~ we do not separate these two regimes in this evaluation. It might be related to the fact that the horizontal ~~extension~~ extent of objects on isobaric surfaces in the Earth's atmosphere is limited by the finite surface of the Earth. This can also be seen in the investigations of the horizontal ~~extensions~~ extents in Sec. ~~??.~~ ??. A linear regression is applied to the complete dataset of each respective month, and the resulting slope is used to determine the respective fractal or self-similar dimension, as indicated in Eq. (4). We find values in the range $D \sim 1.18 - 1.375$, i.e. values well above the value of $D = 1$ for a geometric object.

As illustrated in the simple examples presented in appendix **??**, the ~~reduction~~ increase of fractal dimension for larger objects might be driven either by more elongated objects (~~as in case of~~ for example the systematically stretched rectangular) or by more frayed objects. Both reasons seem to be plausible from the presented snapshots in Figs. **??** and **??**~~; however,~~, but a clear attribution is not possible.

Finally, we determine the annual evolution of the fractal dimension as obtained from monthly data at different pressure levels; the results are represented in Figure **??**. We see a clear separation by vertical levels. For each month, the fractal dimensions are ordered, with the smallest value assigned to pressure level $p = 200\,\mathrm{hPa}$, followed by the medium value at pressure level $p = 250\,\mathrm{hPa}$ and the highest value at pressure level $p = 300\,\mathrm{hPa}$. The fractal dimensions of the observed ISSRs are in the ranges $D \sim 1.18 - 1.29$ ($p = 200\,\mathrm{hPa}$), $D \sim 1.22 - 1.36$ ($p = 250\,\mathrm{hPa}$), and $D \sim 1.26 - 1.375$ ($p = 300\,\mathrm{hPa}$). All fractal dimensions D show a strong seasonal cycle with maxima during summer and minima during winter. The amplitude for the cycles is largest for the pressure level $p = 250\,\mathrm{hPa}$, followed by a slightly reduced amplitude for the pressure level $p = 300\,\mathrm{hPa}$~~; the~~. The variation of the monthly fractal dimension is smallest for the pressure level ~~$p = 300\,\mathrm{hPa}$~~ $p = 200\,\mathrm{hPa}$, but it is still well pronounced. The correlation coefficient for the linear fit indicating the quality of the fits by the linear regression are relatively high ($r^2 \geq 0.95$). Thus, we can conclude that these results are robust and constitute a new and pronounced feature of ISSRs. The influence of the grid resolution (i.e. grid coarsening) on the derived fractal dimension is discussed in ~~Appendix ??.~~ the supplement. Unless otherwise indicated, all plots and results are based on ERA5 data at a horizontal resolution of 0.25° in both longitude and latitude.

The large seasonal variation of D might point to different physical formation mechanisms, as discussed below.

Although this is the first study investigating the fractal nature of ISSRs, we can try to compare our results to similar studies on fractal dimensions, e.g., for high clouds observed in satellite data and models with global storm resolving resolutions. Since ice clouds form in ISSRs, there should be a comparable relationship. **?** evaluated satellite data (~~spacial~~ spatial resolution of $1\,\mathrm{km}$) of (ice) clouds and found a fractal dimension of $D = 1.37 \pm 0.02$ for high cirrus clouds, and a fractal dimension of $D = 1.18 \pm 0.05$ for cumulonimbus clouds. These values are in the same range as values of D found in this study, showing clear fractal properties of clouds. **?** investigated satellite data ($5\,\mathrm{km}$ grid) and high resolution model data (initialized using the ECMWF $9\,\mathrm{km}$ analysis data, linear resolution of 2.5 to $7.8\,\mathrm{km}$ for the core ensemble) and found fractal dimensions in the range $D \sim 1.35 - 1.45$ for the satellite data, and $D \sim 1.27 - 1.53$ for the different (high-resolution) models, which are also comparable to our findings. Since we are not distinguishing between cloudy and clear air inside ISSRs, there might be a mixture of different signatures of clouds and cloud free air massses (constituting all ISSRs). Nevertheless, these results show that ISSRs exhibit a clear signature of self-similarity.

3.3 ~~Changes due to filled objects~~

4 Horizontal extent of ISSRs

~~In our investigations so far, we used the derived islands of ISSRs as we found them. Thus, these regions can have holes, as can be seen clearly in Figures **??** and **??**. These holes increase the perimeter of the objects in our calculations, changing the fractal dimension, as discussed by . This aspect can be seen from two different point of views.~~

1. ~~We are investigating objects with an assumed self-similarity. As a consequence of this assumption we would expect the same structural properties on every scale, i.e. repeating the main features when we zoom into smaller scales. As known from theory there are objects like e.g. the Sierpinski triangle or the Sierpinski carpet with holes, leading to the general structure of the object, and also to the respective fractal dimension. For clouds as macroscopic objects on different scales, we might assume a similar property. If we zoom into a smaller scale, we still find some holes in the cloudy object, thus we could assume that the "holes" inside the ISSRs should be an important feature. A similar feature can be seen in case of high resolution measurements of water vapor, leading to pathlengths of ISSRs on a much smaller scale . From this point of view, the use of unfilled ISSRs would be meaningful, or even mandatory, since their structure with holes is a key feature.~~

2. ~~Following the discussion by , holes in the observed objects (clouds, ISSRs, etc.) might artificially result from the detection procedure. Since it is not really clear which features are real and which are of artificial nature, it is more meaningful to fill the holes and to evaluate the filled objects. This procedure was suggested by and was also applied by .~~

~~These two aspects are, of course, relevant for our investigation. So far, we have used the first approach, using the ISSRs as they are, assuming that the holes are relevant for the inherent structure of these objects. However, for investigating the impact of a possible erroneous description of islands, we repeat the evaluation of the whole data set again using a filling of occurring holes in the islands rather that eliminating the objects with holes.~~

~~In Figure **??** we show the results of the monthly timeseries of fractal dimension on the three pressure levels. We find that the qualitative behavior of the monthly derived fractal dimension remains the same, even for the filled ISSRs. We still see the vertical layering of the dimension at different pressure levels with the same seasonal cycle as the unfilled ISSRs. However, there is, as expected, a change in the values of the fractal dimensions. The values of D for filled objects are reduced by an absolute value of about $\Delta D \sim 0.03 - 0.05$, with smallest deviations for the pressure level $p = 200\,\mathrm{hPa}$ and largest deviations for the pressure level $p = 300\,\mathrm{hPa}$. Nonetheless, the filled objects show the feature of a self-similarity with dimensions $D > 1$.~~

~~Finally, we want to remark that for the upcoming evaluation of the ISSRs in terms of horizontal extensions or pathlengths (see section **??**), we will stay with the unfilled ISSRs, since the local structure is more relevant for these investigations.~~

5 ~~Horizontal extension of ISSRs~~

Although ISSRs are quasi 2D objects on isobaric surfaces, a common measure for ISSRs is the so-called pathlength. ~~If a (commercial) aircraft crosses an ISSR, the track inside the ISSR is termed as pathlength~~The pathlegth is the distance traveled by an aircraft within an ISSR. **?** first investigated pathlength statistics of ISSRs using data from the MOZAIC project (**?**), this study was later extended by **?** for the IAGOS data set (**?**). The pathlengths, ~~although constituting just~~ constituting 1D tracks through 3D objects, give a first idea about the horizontal ~~extension and even~~ extent and the shape of ISSRs. Since we use quasi 2D data (at pressure levels) in this study, we can generate artificial pathlengths in order to simulate the distribution of realistic pathlengths. These can be used in two different ways. First, using zonal and meridional directions (i.e. East-West, North-South, respectively), we can statistically investigate the horizontal ~~extension~~ extent (or span) of ISSRs in a systematic but less realistic way. Second, we can generate more realistic pathlengths in using simulated flight tracks along geodesics between frequently used airports. These ~~pathlengths then can be compared to distributions as derived from aircraft data .~~ tracks are chosen specifically to cover routes also included in the IAGOS project (**?**).

~~For the determination of longitudinal and latitudinal extensions or spans of ISSRs, we proceed as follows. These spans~~ The zonal and meridional extents are not measured as a means to obtain the largest extent in one direction, but rather to generate a full set of spans along latitudes/longitudes touched by the respective ISSR. Thus, the ISSR is ~~therefore~~ measured in "slices" with a spacing of 0.25 degrees. This procedure is carried out for both ~~directions,~~ zonal and meridional ~~, respectively~~directions. The resulting data is ~~divided~~ separated by seasons, i.e. summer (JJA) and winter (DJF), and is used ~~for deriving~~ to derive statistical distributions of the horizontal ~~extension/spans~~ extent of ISSRs. ~~Note that we use the "unfilled" ISSRs as a basis for the evaluation, since we are interested in the real distances of ISSRs; filling the holes of ISSRs would lead to artificially enhanced pathlengths. Since one application of this evaluation could be the issue of avoiding ISSRs for reducing the non-CO~~$_2$ ~~effects of aviation , the real horizontal extension of ISSRs including interruptions inside the ISSRare of importance. We report the mean and median values for the filled~~

The data used deliberately does not include measurements shorter than 27 km, as this is approximately the longest single-pixel length that can be measured in the target region. Including distances below that threshold is not meaningful, since this would introduce artificial biases stemming from different pixel sizes at different locations, leading to artificial breaks of regimes. Note that we still use the data set with the restriction of a minimum of 24 pixels per ISSR. This might also crucially reduce the number of small ISSRs in the ~~respective tables~~evaluation, but this is consistent with our investigations in section 3.

Former investigations of pathlength statistics showed Weibull-type distributions (**??**), or at least the distributions could be well represented using Weibull-type fits. ~~However, we~~ We emphasize that the application of the Weibull distribution in this context is purely empirical. In contrast to classical extreme-value statistics, we do not consider maxima of underlying distributions, but rather the full population of pathlengths. Therefore, a strict interpretation in the framework of extreme-value theory is not appropriate here. Instead, the Weibull distribution is used as an empirical fitting function that provides a good representation of the observed distributions and allows for a compact characterization via its parameters, which is consistent with above mentioned studies studying ISSR distributions.

The probability density function of a Weibull distributed random variable x is given by $f(x) = \gamma p x^{p-1}\exp(-\gamma x^p)$ with shape parameter $p \in \mathbb{R}_+$ and scale parameter $\gamma \in \mathbb{R}_+$. The parameter p determines the overall shape of the distribution~~; it~~. It can be obtained by using a Weibull plot (and the cumulative distribution $F(x) = 1 - \exp(-\gamma x^p)$) with linear regression. The slope of the regression line is the parameter p. For more details about Weibull plots we refer to appendix **??**.

4.1 Span along Longitude

First, we investigate the ~~longitudinal~~ meridional slices, i.e. pathlengths along longitudes~~in North-South direction~~. Since the major variation of horizontal ~~extensions~~ extents is between winter and summer, we exclusively show these ~~distributions~~ results in all respective figures. The ~~mean values for all four seasons are then reported in the respective Tables. The normalized statistical distributions for~~ probability density functions for winter and summer, differentiated into three pressure levels are shown in Figure **??**. ~~Here, we use the full data set for the period between 2010 and 2020 using the data with unfilled holes. The corresponding figures based on the data set with filled holes are provided in Appendix~~ ~~**??**, the~~ The mean and median values are reported in ~~table~~Table **??**.

For all pressure levels $p = 200/250/300\,\mathrm{hPa}$, the mean ~~North-South span~~ meridional extent of ISSRs is significantly higher in winter ~~as~~ compared to the summer months. While the density does not differ ~~much between the~~ substantially between seasons for ISSRs up to a size of approximately $10^3\mathrm{km}$, the disparity between summer and winter ~~gets larger with increasing the~~ inflates with increasing measured span. The absolute mean span increases with decreasing pressure (i.e. with rising altitude levels)~~; the~~. The relative difference comparing summer and winter slightly increases with higher pressure (i.e. lower altitudes). All values are reported in Table **??**. The strong seasonal cycle is present in both ~~,~~ the mean values and ~~also in~~ the extreme tails of the distributions. ~~For~~ In winter, the season with maximum mean values, a clear enhanced probability of finding extremely large pathlengths (i.e. enhanced values in the tail of the distribution) can be seen, whereas ~~during summer there is a reduced~~ the probability of extremely large pathlengths is reduced in summer. The histograms are not symmetric but right-skewed, so the mean is pulled upward by few large ISSRs. The median is more robust and thus more suitable to assess the typical 2D features of ISSRs. The median behaves similarly to the mean in that it generally decreases with increasing pressure~~. The~~, the exception being a minor increase from 200 hPa to 250 hPa in the summer months. ~~This behaviour is not observable in winter. As the distribution presents right-skewed, the median is significantly smaller than the mean across pressure levels and seasons, thus the few large events in the tail of the distribution shift the mean towards larger values. This suggests possibly two regimes and/or formation mechanisms (e.g., synoptic-scale vs localized), which is discussed later in more details.~~

~~All histograms show bars representing small ISSRs. The distributions have a strong similarity to Weibull distributions as found in former investigations . Therefore, we use Weibull plots (see Figure **??**) for further evaluations, i.e. to determine the approximate shape parameter of the distribution.~~

~~The data used deliberately does not include measurements shorter than 27 km, as this is approximately the longest single-pixel length that can be measured in the target region. Including distances below that threshold is not meaningful, since this would introduce artificial biases stemming from different pixel sizes at different locations, leading to artificial breaks of regimes.~~

~~Note, that we still use the data set with the restriction of a minimum of 24 pixels per ISSR. This might also crucially reduce the number of small ISSRs in the evaluation.~~

The data shows two regimes with roughly linear correlation but different slopes in the Weibull plots (i.e. in the respective values of parameter p). Thus, there are two distinct distributions superimposed in the histograms. The transition between these regimes is roughly located at ~~$\sim 2 \cdot 10^2$km; for~~ $L \sim 2 \cdot 10^2$km. For the smaller ISSRs we find values of $p \sim 1.5$, whereas ~~for the larger ISSRs the~~ the distribution is close to an exponential distribution ($p \sim 1$~~. This alludes to the existences of two different physical formation processes. Neither season nor pressure level influences this observation.~~) for the larger ISSRs .

The presence of two clearly separated size regimes ~~suggest potentially~~ suggests different formation mechanisms of ISSRs operating at distinct spatial scales, i.e. ~~processes acting on distinct but different scales~~synoptic-scale vs localized on mesoscales, leading to characteristic horizontal extents. ~~However, this~~ This hypothesis is subject ~~of further investigations~~to further investigation, since an in depth analysis is out of scope of this study.

4.2 Span along Latitude

Alike investigating the ~~North-South~~ meridional slices along longitudes, we now investigate horizontal ~~extensiosn or spans~~ extents along latitudes, i.e. slices in zonal (East-West) direction within ISSRs. The ~~normalized statistical distributions~~ probability density functions for the two seasons (summer/winter), ~~differentiated~~ separated into three pressure levels, are shown in Figure **??**. ~~Again, we use the full data set for the period between 2010 and 2020.~~

The distribution of measured ~~latitudinal~~ zonal distances shows almost ~~the same~~ identical qualitative characteristics as the ~~longitudinal~~ meridional counterpart. We see a strong decrease in frequency of occurrence for spans larger than the mean values. ~~However, in~~ In contrast to the ~~longitudinal~~ meridional distributions, there is an additional very small mode at very large spans (~~$\sim 10,000$km~~$L \sim 10,000$km), indicating the occurrence of very few very large ISSRs in East-West direction. This difference can be explained by pathlengths in meridional direction being limited to the $30-80°$N range we study, while zonal pathlengths can theoretically extend to the length of the latitude. For all pressure levels, the mean span is significantly higher in the winter months compared to the summer months. Differences in the distribution of spans by season are most clear for larger spans. The mean generally decreases slightly with a higher pressure level. The decrease in mean span with higher pressure is more evident in the summer months. Analogous to the meridional observations, the median is consistently lower than the mean. The median increases ~~slightly~~ marginally from 200 hPa to 250 hPa ~~with the exception of summer using unfilled data~~in winter but behaves analogous to the mean in all other instances. It is lowest at the highest pressure level. As with the meridional span, ~~this~~ the distribution is right-skewed~~distribution strengthens the case for using Weibull distributions to further investigate the extensions~~. All values are reported in Table **??**~~, together with the mean/median values for the filled ISSRs. We also~~. We see the same feature of enhanced probability for extremely large pathlengths during winter and reduced probability during summer.

In Figure **??** we investigate the Weibull plots for the distributions. For the ~~East-West~~ zonal spans, three distinct regimes ~~appear to be present~~are visible. There is a regime of very small ~~latitudinal~~ spans (up to ~~$\sim 10^2$, km~~$L \sim 10^2$, km), without a clear Weibull signature in the Weibull plot. In addition, there are two further regimes, separated at values ~~$\sim 2 \cdot 10^3$ km, which clearly~~ $L \sim 2 \cdot 10^3$ km, which show linear behavior in the Weibull plots, thus indicating two different regimes ~~and~~

clearly distinguishable by a different Weibull parameter p. The regime of intermediate spans of ISSRs results ~~into~~ in an almost exponential distribution (Weibull parameter $p \sim 1$), whereas the regime of very large spans ~~clearly show a Weibull charactersisties~~ shows a Weibull characteristic with parameter $p \sim 0.2$~~; this characteristics remains almost the same over the~~. This characteristic remains virtually identical across seasons.

The regime of very small pathlengths might be an artifact of the evaluation. The real area of a pixel (or grid point) depends crucially on the latitude. ~~For East-West spans we regularly scan through at all latitude, thus~~ Scanning high latitude circles will systematically produce small horizontal sizes for ISSRs consisting of few pixels.

The tiny mode of very large spans of ISSRs ~~up to few ten thousands of kilometers~~ of orders up to almost latitudinal circumference might indicate the impact of synoptic and large scale weather systems, triggering atmospheric flows in the zonal direction with a persistent vertical upward motion, thus leading to a persistence of ice supersaturation in the UTLS. The upper bound of such synoptically driven ISSRs is ~~, of course,~~ given by the circumference of the Earth at certain latitudes.

4.3 Span along modeled aircraft tracks (pathlengths)

In addition to the rather artificial longitudinal and latitudinal measurements, realistic pathlengths are investigated. In lieu of real in situ measurements, ~~some tracks of~~ tracks along commercial aircraft routes are investigated. ~~For this purpose we exemplarily chose four typical flight routes between the US and Europe,~~ We exemplarily chose three transatlantic flight routes based on the flight tracks frequently found in the IAGOS data (**?**). The three tracks ~~,~~ (i.e. direct flight routes ~~are~~ Berlin–Chicago, Frankfurt–Atlanta, and Paris–Miami~~, placed on the~~) are assumed as geodesics on a sphere (Figure **??**). The pathlengths are obtained by measuring the distance along the sections of ice supersaturation on the flight track from airport to airport. The histograms (Figure **??**) show numerical artifacts for small ISSRs. ~~However, we~~ We find the same features of pathlength statistics for mean values and distributions as in Secs. ~~?? and ??~~ **??** and **??**. The winter months yield the highest mean pathlengths while summer ~~brings forth~~ returns the shortest mean distances. For all seasons there is an increase of mean and median values with increasing ~~heights:~~ altitude. The highest mean/median pathlengths can be found at the highest level ($p = 200\,\mathrm{hPa}$), and the lowest values at $p = 300\,\mathrm{hPa}$ with intermediate values at the level $p = 250\,\mathrm{hPa}$. The distributions mostly agree for the two seasons, with a larger deviation for the largest pathlengths~~, explaining also the seasonal difference~~. During winter, there are more large ISSRs than during summer. The mean and median values of the pathlength distributions are represented in Table **??**. The mean exceeds the median in all instances.

In the evaluation of the Weibull plots (Fig. **??**), we find two different regimes of distinct Weibull distributions with different values of p, separated at pathlengths ~~$\sim 10^2\,\mathrm{km}$, as for the longitudinal extensions (North-South). However, the~~ $L \sim 10^2\,\mathrm{km}$, analogous to the meridional extents. The difference in the regimes is more pronounced ~~than in~~ compared to the former evaluation, and the ~~clear~~ signal of the very large zonal ~~extensions (East-West direction)~~ extents is missing in this evaluation. This is expected, as the largest possible extent is the distance between airports. As before, we ~~skip~~ overlook the very small pathlengths in the linear regression~~, since the probabilities are probably biased by the selection process (leaving out regions smaller than 24 pixels).~~. For both regimes, the seasonal cycle is not strongly pronounced. In the small pathlengths regime, we find values $p \sim 1.95$, whereas in the regime of large pathlengths, the values are close to an exponential distribution, i.e. $p \sim 1$. This signal is very robust, with almost no seasonal variation in the parameters or with respect to the different levels.

4.4 Agreement of pathlength distributions

The results of the longitudinal, latitudinal and pathlength measurements agree well with another. The observation of the summer months yielding the shortest and winter the longest mean distances generally holds true across methods. This agreement applies to most results in terms of dependence on pressure level as well. The mean span is generally larger at lower pressures, i.e. higher up in the atmosphere~~; while~~ . While this is always true for the comparison between levels $p = 200\,\mathrm{hPa}$ and $p = 300\,\mathrm{hPa}$, in the intermediate level $p = 250\,\mathrm{hPa}$ some variations can be seen. The mean values for ~~East-West spans~~ zonal extents are generally larger than ~~for the North-South extensions~~ those for the meridional extents and the flight track pathlengths. This feature ~~is obvious from the example of ISSRs~~ likely stems from ISSRs usually extending more in East-West direction (as shown in Figs. **??** and **??**~~. The ISSRs usually extend more in East-West direction, as~~), driven by the mean large scale flow.

The results of the simulated IAGOS flights have a larger qualitative agreement with the meridional ~~extensions~~extents. In both cases, there is a scale break at $L \sim 10^2\,\mathrm{km}$, separating two different modes of Weibull distributions. The tiny mode of exceptionally large spans as found in the zonal slices cannot be detected in the pathlengths of realistic flight tracks~~, since these~~ . These very few, very large ISSRs are mostly elongated into zonal direction ~~. Usual flight pattern along geodesics must certainly cross them for reaching the respective destination.~~ and the maximum of possible pathlength is limited by flight track length.

~~As we investigate ISSRs on fixed pressure levels. Thus, the seasonal variation of the tropopause, which is closely correlated with ice supersaturation , might result into fewer ISSRs during winter time at high altitude levels and skew the results.~~

The distributions as well as the mean (and median) values of the simulated pathlengths compare qualitatively well with the former investigations of pathlengths as obtained from aircraft measurements (**??**). The mean (and median) values as obtained from the ERA5 data are generally larger than ~~in case~~ those of aircraft data. This is not surprising, since the aircraft data has a much higher temporal (and thus spatial) resolution~~; even for running means over the data set, mimicking a coarser resolution, the data still has a signature of small scale events.~~ . The Weibull shaped distributions can be seen in all evaluations ~~, with a more or less prominent feature of~~ with a varying degree of prominence of a scale break at pathlengths $L \sim 10^2\,\mathrm{km}$. ~~However, the~~ The fitted Weibull parameter of the distributions differ between aircraft data and ERA5 evaluation~~, although the~~ . The general feature of a smaller slope for pathlengths below the scale break and a higher slope for pathlengths above the scale break is present in all evaluations. The difference in slopes for aircraft vs. ERA5 data is likely due to the different resolutions of the data sets. Finally, we see that the range of the obtained distributions of pathlengths is roughly the same in all investigations (ranging from few $10\,\mathrm{km}$ up to $\sim 5000\,\mathrm{km}$).

5 ~~Conclusion,~~ Discussion~~and Outlook~~

5.1 ~~Conclusion~~

Ice supersaturated regions in the UTLS are important potential formation regions of ice clouds and contrails. For a better understanding of these regions ~~and thus a meaningful representation in models~~ on different scales and resolutions, we investigated ISSRs in ERA5 data as "macroscopic" and quasi two dimensional objects on pressure levels in the UTLS. We can summarize the main findings as follows:

1. The analysis of the area–perimeter relationship demonstrates that ISSRs in the UTLS exhibit self-similar (fractal) behavior, with dimension values between 1 and 2, as expected for planar structures with irregular boundaries. Pathlength statistics are consistent with these fractal properties.

2. There is a seasonal cycle in the number of ISSRs. We find many ISSRs for the summer season, much less in winter. There is also a vertical layering of the amount of ISSRs, with high numbers at lower altitudes, and a decrease in the number for increasing heights. For the total area of ISSRs we find the same type of seasonal cycle with high values during summer and low values during winter. The vertical layering is the same for the total number of ISSRs.

3. For the fractal dimension of ISSRs we find values in the range $1.175 \leq D \leq 1.375$. The fractal dimensions also show a distinct vertical layering with largest values for ~~high~~ low altitudes and lower values for ~~low altitudes. For filled ISSRs the fractal dimension is systematically lowered by about $\Delta D \sim 0.02 - 0.05$, however still showing a fractal characteristics of $D > 1$.~~ high altitudes.

4. There is a clear seasonal cycle for the fractal dimensions with smaller values for summer and larger values for winter. This signal appears on all considered pressure levels.

5. In the evaluation of the horizontal ~~extensions~~ extents of ISSRs, i.e. deriving pathlengths in various ways (artificial East-West and North-South spans vs. realistic flight paths), we find a seasonal cycle as well. Generally, we find smaller pathlengths during summer (smaller mean values, less extremes), whereas during winter the pathlengths are generally larger (larger mean value, more ~~really~~ large pathlengths). This feature is robust for all evaluated pressure levels.

6. In the ~~realistic~~ modeled flight pathlengths statistics, we find different regimes, as also indicated by different values of the shape parameter p in the Weibull plots, indicating a superposition of two different Weibull distributions. The scale break can be estimated at $L \sim 2 \cdot 10^2 \mathrm{km}$, which is ~~comparable to~~ consistent with former investigations.

5.1 ~~Discussion~~

As can be seen in the simple exercise in appendix **??** about changes in fractal dimensions, more elongated or even strongly frayed objects have larger fractal dimensions than their strictly geometric counterparts (i.e. $D > 1$). As indicated by the snapshots in Figs. **??** and **??**, in summer the ISSRs are more elongated and more frayed than in winter. The fractal dimension is larger for ISSRs in summer, thus we expect more frayed/elongated objects. In turn, these objects result in shorter pathlengths, as traversing them generally involves smaller cross sections of ice supersaturation. On the other hand, ISSRs in winter have smaller fractal dimensions, thus suggesting a more compact shape. This is consistent with the pathlength statistics, as traversing more compact objects results in fewer short pathlengths and, on average, larger cross sections. Thus, the seasonal cycles of fractal dimensions and pathlengths are physically consistent.

The strong contrast between fractal properties of ISSRs for different seasons suggest different formation processes. For this we might also take into account the seasonal fluctuation in the number of ISSRs with a peak in summer (in agreement with **?**). A possible explanation is the transportation of water vapor into the UTLS by convection during the summer months. A multitude of convection sites causes high ISSR numbers while the winter months are dominated by low pressure areas and frontal systems. The ascending air flow of the warm conveyor belt causes a connected area of supersaturation. This hypotheses is supported by the shorter mean latitudinal and longitudinal spans in summer compared to winter months.

~~As a further corroboration we calculated the~~ It is further corroborated by the pronounced seasonal cycle in the mean area of ISSRs~~from our statistics of number and area of ISSRs in Section **??**~~, which is ~~represented~~ presented in Figure **??**.

~~We clearly see a pronounced seasonal cycle in the mean area of ISSRs with a very~~

The distinct maximum in winter ~~(mean values $\bar{A} \sim 5-6\cdot 10^{11}\mathrm{m}^2$) and a~~ and clear minimum in summer ~~(mean values $\bar{A} \sim 2.5-3\cdot 10^{11}\mathrm{m}^2$); assuming a purely spherical shape, this would translate into a factor of 2 for the radii. This behavior~~ is consistent with the proposed formation mechanisms: convective processes in summer produce more localized ISSRs, whereas large-scale dynamics in winter lead to more extended structures.~~While the 200 hPa and 300 hPa levels show good agreement across all seasons, the average area per ISSR is noticeably larger at 250 hPa during summer. This feature does not represent an outlier but can be explained by jointly considering the number count and total area. In winter, both the number count and total area at 250 hPa are substantially lower than at 300 hPa. During summer, the total area at both levels becomes comparable, despite a lower number count at 250 hPa. Consequently, the ratio of total area to number count is increased at 250 hPa in summer. Seasonal variations in tropopause altitude primarily affect the 250 hPa level as it is intermittently located within both the upper troposphere and lower stratosphere. In contrast, 200 hPa remains predominantly above and 300 hPa predominantly below the tropopause throughout the year , resulting in less pronounced seasonal variability at those levels.~~

Extending this hypothesis to the seasonal variation in the fractal dimension, particularly during the period with the lowest values in summer, provides two possible interpretations. One perspective suggests that convective sites lead to ISSRs with frayed boundaries, resulting in long perimeters compared to the smoother edges of ISSRs caused by frontal activity, i.e. large scale dynamics. An alternative interpretation posits that multiple convective "bubbles" occurring in close proximity may be collectively labeled as a single ISSR. This would lead to elevated perimeters attributed to their intricate structure and the added internal perimeter created by holes that segregate convective regions.

In addition, the scale break in the Weibull plots of the pathlengths also suggest a possible superposition of two different formation pathways. This can possibly originate in the slopes corresponding to smaller and larger ISSRs. Smaller scale ISSRs form in localized lifting phenomena as e.g. convection. Frontal activity and large scale dynamics create larger areas of supersaturation with respect to ice.

~~However, an~~ An in depth investigation of the formation processes of ISSRs for the full data set is beyond the scope of this study, but is planned as future investigation.

5.1 ~~Outlook~~

6 Outlook

The impact and implications of these findings are broad. ~~The level of self similarity across scales can prove helpful in modelling more accurate representations of ISSRs and can thus impact forecasts. Knowing that ISSRs largely keep their characteristics irrespective of their scale can simplify their modelling. The~~ The identification of fractal structures enriches scientific inquiry by providing new tools for analysis ~~, enhancing modelling accuracy,~~ and offering deeper insights into the complexity of natural phenomena. Future studies using different definitions of fractal dimension (e.g. the Hausdorff dimension) might substantiate our understanding of the fractal properties of ISSRs. Confirming our findings with studies on measured rather than modeled data could act as an endorsement to incorporate self-similar and fractal characteristics in other models. As a more practical aspect of the analysis, the fractal properties of ISSRs must be taken into account for the highly discussed concepts of contrail avoidance (e.g., **?**). Looking at the examples of frayed ISSRs in Figs. **??** and **??**, it seems to be very optimistic that contrail avoidance might work in a meaningful, impactful or ~~feasable way; the~~ feasible way. The large amount of small ISSRs (or short pathlengths) in connection with the huge fractal objects makes it very difficult to manage efficient contrail avoidance in the Northern hemisphere.

Appendix A: Methods for evaluation

In Figure **??** two possible definitions of neighborhoods of points on gridded data are shown, originating from investigations with cellular automaton (**?**). ~~In case of~~ Given a 4-neighborhood (or von Neumann neighborhood), the central pixel ~~has only connection~~ is only connected to pixels via the four edges of the point~~, whereas in case of a~~ . A 8-neighborhood (or Moore neighborhood) is defined as the central pixel ~~is~~ being additionally connected to pixels via the diagonals[1]. In our evaluation, we choose the Moore neighborhood, since the physical properties as connectivity should not depend on a grid representation, i.e. rotation of the object should result into the "same" connected object.

For the derivation of the grid point's area we use the coordinates of the point (x, y) as the center of a rectangular with vertices $(x \pm 0.125, y \pm 0.125)$ (see right panel of figure **??**). The area (in square metres) is then assigned to the grid point, taking into account the latitudinal stretching of the grid. For the derivation of the perimeter, we use the four edges of the introduced rectangular, with the respective length (in metres). If there is no neighboring grid with index=1, the edge is counted as part of the perimeter.

Appendix B: Changes in fractal dimensions

In this appendix we illustrate some effects which might have some influence on the change in the fractal dimension. We use the area-perimeter relation $P = CA^{\frac{D}{2}}$ for the investigations.

For geometric objects, the fractal dimension coincides with the classical integer values dimension, as can be seen directly for simple examples as spheres, squares, rectangulars with fixed ratio of sides.

1. Sphere: For a sphere with radius r we easily calculate the perimeter $P = 2\pi r$ and the area $A = \pi r^2$, thus leading to the derivation $2\pi r = P = CA^{\frac{D}{2}} = C\left(\pi r^2\right)^{\frac{D}{2}} = C\pi^{\frac{D}{2}} r^D$ which immediately gives $D = 1$ and $C = 2\sqrt{\pi}$.

2. Rectangular with fixed ratio of sides: We assume sides a and $b = d \cdot a$ with a fixed value $d > 0$; thus, we calculate $2(1+d)a = P = CA^{\frac{D}{2}} = C\left(d \cdot a^2\right)^{\frac{D}{2}} = Cd^{\frac{D}{2}} a^D$ with $D = 1$ and $C = \frac{2(1+d)}{\sqrt{d}}$. Note that the ~~case~~ example of a square $(d = 1)$ is included here, leading to $C = 4$.

As a first example for a fractal object, we compare the fractal dimension of a square with a systematically stretched rectangular, i.e. we change the ratio between the two sides, while we proceed to larger sizes. In fact, we set $a = \epsilon$ and $b = \epsilon^2$ for $\epsilon \in \mathbb{R}_+$. Thus, we can calculate

$$A = a \cdot b = \epsilon^3, \;\; P = 2a + 2b = 2\epsilon + 2\epsilon^2 = 2\epsilon(1+\epsilon) \tag{B1}$$

leading to the area-perimeter relation

$$2\epsilon(1+\epsilon) = P = CA^{\frac{D}{2}} = C\left(\epsilon^3\right)^{\frac{D}{2}}. \tag{B2}$$

[1] In the description of the used program package, the 4-neighborhood is called 1-connectivity, and the 8-neighborhood is called 2-connectivity. In fact, in literature these terms are not frequently used.

Applying the logarithm, we can derive an approximate value of the dimension D (and also the constant C).

Using $\log(P) = \log(2) + \log(\epsilon) + \log(1+\epsilon)$, and $\log(A) = 3\log(\epsilon)$, we obtain for the dimension

$$\frac{D}{2} = \frac{\log(P) - \log(C)}{\log(A)} = \frac{\log(2) + \log(\epsilon) + \log(1+\epsilon) - \log(C)}{3\log(\epsilon)} \tag{B3}$$

For large ϵ we find $\log(1+\epsilon) \approx \log(\epsilon)$ and $\log(P) - \log(C) \approx \log(P)$, thus we find

$$\frac{D}{2} = \lim_{\epsilon\to\infty} \frac{\log(P) - \log(C)}{\log(A)} = \frac{2}{3} \Leftrightarrow D = \frac{4}{3}, \tag{B4}$$

and we also obtain $C \approx 2$. In the left panel of Figure **??** the area-perimeter relation for a square and a stretched rectangular are shown together with the theoretically derived values for the fits. The systematically stretched rectangular turns into a more line-shaped geometric object, thus changing the fractal dimension to larger values.

As a second example, we investigate ~~the case of~~ a square, ~~which boundary is systematically frayed , thus~~ with a systematically frayed boundary leading to a slightly smaller area but to a strongly enhanced perimeter. ~~Again, for a normal square we would find~~ A regular, unfrayed square would yield a (fractal) dimension of $D = 1$. ~~However, for~~ For the frayed square, we see that the fractal dimension is also enhanced, leading to values of about $D \sim 1.21$. The scatter plot of the normal square and the frayed square, respectively, is shown in the right panel of Figure **??**, together with the linear fit via regression. These two mechanisms (stretching and fraying) can for instance serve for changing fractal dimensions of "two dimensional" geometric objects.

Appendix C: ~~Changes due to coarsening of the grid~~

~~For a rigorous definition of fractal dimensions (in the Hausdorff sense) one would have to send the grid resolution to zero, i.e. approaching an infinitely small length scale for measuring perimeters and areas, respectively. Only in the limit, the fractal dimensions are well defined. We cannot obtain this limit, since we are limited by the highest resolution of the ERA5 data, which is~~ ~~$0.25° \times 0.25°$~~ ~~on an isobar surface. However, we can check the robustness of the fractal properties by artificial coarsening of the data and redoing the evaluation in terms of calculating the fractal dimension for the coarser objects. For this purpose, we introduced two different data sets, i.e. a **coarse** set (half resolution, i.e.~~ ~~$0.5° \times 0.5°$~~~~), and a **very coarse** set (quarter resolution, i.e.~~ ~~$1° \times 1°$~~~~). In both cases, we use the same ERA5 data, i.e. a full set of 11 years (2010-2020).~~

~~In Figure **??** we exemplarily show the results for the pressure level~~ ~~$p = 250\,\mathrm{hPa}$~~~~; for the other two pressure levels we find qualitatively the same features (not shown). For all grid resolution we still find the signature of self-similarity, i.e. a fractal dimension of~~ ~~$D > 1$~~~~. There is a similar qualitative change if we proceed from coarser to finer grid resolution. The highest values of~~ ~~D~~ ~~are found for the very coarse resolution, with a jump in order of~~ ~~$\Delta D \sim 0.04 - 0.06$~~ ~~to the coarse resolution. The next step to the fine (i.e. ERA5 native) resolution is then much smaller, the difference is in order of~~ ~~$\Delta D \sim 0.01 - 0.02$~~~~. More precisely, the dimension seem to "converge" to the values of the fine resolution, although this statement is only qualitatively, based on 3 available grid resolution. For the other pressure levels, we find the same qualitative behavior, and also the values of the difference in~~ ~~D~~ ~~are almost the same.~~

~~Thus, we can conclude that the fractal properties seem to be robust over all tested grid resolutions. Surprisingly, the highest values of the fractal dimension can be obtained for the coarsest resolution of~~ ~~$1° \times 1°$~~~~, whereas the self-similarity is still~~

~~prominent for the ERA5 resolution. Finally, we can state that the behavior is exactly the same for the filled versions of ISSR, as described in the previous section **??**.~~

Appendix C: Potential errors for gridded data: Overestimation of perimeters

It is well known that the calculation of perimeter on gridded data lead to overestimation of this quantity (**?**). However, it is not completely clear ~~,~~ if and how the area-perimeter relation might change due to this error. Exemplarily, we investigate this for a sphere~~; we~~ . We calculate the area and perimeter according to the methods explained above and compare these values with the analytically calculated quantities. The area-perimeter relations are represented in Figure **??**

~~Although there is~~ Despite a systematic shift in perimeters, the resulting fractal dimension $D \sim 0.999564$ (as a linear fit to the double logarithmic scatter plot) is very close to the analytically determined "fractal" dimension $D = 1$. Thus, we are confident that the systematic overestimation of perimeters does not disturb the quantitative investigations of area-perimeter relations.

Appendix D: Weibull plots

Here we give a short explanation for the use of the Weibull plot. The general form of a Weibull probability distribution of a non-negative random variable x is given by

$$f(x) = \gamma p x^{p-1} \exp(-\gamma x^p) \tag{D1}$$

with the shape parameter $p \in \mathbb{R}_+$, responsible for the slope of the distribution, and the scale parameter $\gamma \in \mathbb{R}$. The cumulative distribution (i.e. via integration) can be written as

$$F(x) = \int_0^\infty f(t)\, dt = 1 - \exp(-\gamma x^p), \text{ i.e. } 0 \leq F(x) \leq 1. \tag{D2}$$

This equation can be reformulated as follows.

$$1 - F(x) = \exp(-\gamma x^p) \quad \Leftrightarrow \quad \frac{1}{1-F(x)} = \exp(\gamma x^p) \tag{D3}$$

$$\Leftrightarrow \quad \log\left(\frac{1}{1-F(x)}\right) = \gamma x^p \tag{D4}$$

$$\Leftrightarrow \quad \log\left(\log\left(\frac{1}{1-F(x)}\right)\right) = \log(\gamma x^p) \tag{D5}$$

Finally, we end with this equation

$$\log\left(\log\left(\frac{1}{1-F(x)}\right)\right) = \log(\gamma) + p \cdot \log(x) \tag{D6}$$

thus constituting a linear relationship between the quantities $\eta := \log(-\log(1-F(x)))$ and $\xi = \log(x)$ with a slope of p. Thus, using a scatter plot of variables (ξ, η) we can derive the parameter p via linear regression.

Appendix E: ~~Fractal dimension - additional plots~~

~~Here we present the determination of the fractal dimension for three pressure levels in case of July 2020.~~

Appendix E: ~~Results using filled holes~~

~~In this appendix, we present the corresponding figures based on the data set with filled holes. These results complement the analysis shown in the main text, where primarily the unfilled-hole data were discussed.~~

~~The purpose of including the filled-hole data is to demonstrate the robustness of the identified statistical characteristics of ISSR horizontal extensions with respect to the treatment of internal gaps. While minor quantitative differences may occur, the overall behavior of the distributions, including their seasonal dependence and characteristic regimes, remains consistent.~~

~~The following figures therefore provide a direct comparison to the results presented in Sections **??** and **??**.~~

. The code for evaluation is available at Zenodo. DOI: 10.5281/zenodo.15531354

. The ERA5 data are available at the MARS archive of ECMWF.

. HS, PS and PR designed the study. HS and SN developed the data evaluation methods, HS carried out the data analysis. HS, PS and PR discussed the results, and wrote the manuscript.

. The authors state that there are no competing interests.

. We thank two anonymous reviewers for important comments leading to an overall improvement of the manuscript. Philipp Reutter acknowledges support by the DFG within the Transregional Collaborative Research Centre TRR301 TPChange, Project-ID 428312742 , project C01. Peter Spichtinger acknowledges support by the DFG within the Transregional Collaborative Research Centre TRR301 TPChange, Project-ID 428312742, project B07. The study contributes to the project "Big Data in Atmospheric Physics (BINARY)", funded by the Carl Zeiss Foundation (grant P2018-02-003).