Statistical downscaling of water vapour satellite measurements from profiles of tropical ice clouds

Giulia Carella[1,2], Mathieu Vrac[1], Hélène Brogniez[2], Pascal Yiou[1], and Hélène Chepfer[3]

[1]Laboratoire des Sciences du Climat et de l'Environnement (LSCE/IPSL, CNRS - CEA - UVSQ - Université Paris-Saclay), Orme des Merisiers, Gif-sur-Yvette, France

[2]Laboratoire Atmosphères, Milieux, Observations Spatiales (LATMOS/IPSL, UVSQ Université Paris-Saclay, Sorbonne Université, CNRS), Guyancourt, France

[3]Laboratoire de Météorologie Dynamique (LMD/IPSL, Sorbonne Université, Ecole Polytechnique, CNRS), Paris, France

Correspondence: Hélène Brogniez (helene.brogniez@latmos.ipsl.fr)

Abstract. Multi-scale interactions between the main players of the atmospheric water cycle are poorly understood, even in present-day climate, and represent one of the main sources of uncertainty among future climate projections. Here, we present a method to downscale observations of relative humidity available from the passive microwave sounder SAPHIR at a nominal horizontal resolution of 10 km to the finer resolution of 90 m using scattering ratio profiles from the lidar CALIPSO. With the scattering ratio profiles as covariates, an iterative approach applied to a non-parametric regression model based on Quantile Random Forest is used to effectively incorporate into the predicted relative humidity structure the high-resolution variability from cloud profiles. Results are presented for tropical ice clouds over the ocean: based on the coefficient of determination (with respect to the observed relative humidity) and the Continuous Rank Probability Skill Score (with respect to the climatology), we conclude that we are able to successfully predict, at the resolution of cloud measurements, the relative humidity along the whole troposphere, yet ensuring the best possible coherence with the values observed by SAPHIR. By providing a method to generate pseudo-observations of relative humidity (at high spatial resolution) from simultaneous co-located cloud profiles, this work will help revisiting some of the current key barriers in atmospheric science. A sample of co-located SAPHIR and CALIPSO observations, together with the downscaled relative humidity predictions, is made available at http://dx.doi.org/10.14768/20181022001.1.

1 Introduction

The atmospheric water cycle consists of complex processes covering a wide range of scales. At small scales, the components of the atmospheric water cycle - water vapour, clouds, precipitation (rain and snow), aerosols - interact amongst each other and with their surrounding environment through micro-physical, radiative and thermo-dynamical processes. At global scales, the atmospheric water cycle interplays with the global atmospheric circulation and the Earth radiative balance. These complex multi-scale interactions are not well understood and how the global atmospheric water cycle works in present-day climate is the subject of intense research, e.g. within the World Climate Research Program (WCRP) core project "Global Earth Water cycle Exchanges" (GEWEX, http://www.gewex.org/) and within the WCRP grand challenge on "cloud, circulation and climate

sensitivity" (https://www.wcrp-climate.org/grand-challenges). Given this poor understanding, it is challenging to anticipate how the atmospheric water cycle will evolve in the future as climate warms (Boucher et al., 2013).

A symptomatic example of this lack of knowledge is the difficulty of state-of the art climate models to reproduce the observed clouds and precipitation in present-day climate (Nam et al., 2012), Cesana and Chepfer (2012), Zhang et al. (2005), Kay et al.
5 (2016), Klein et al. (2017), Lacour et al. (2017)). One of the reasons is that small-scale processes act at space and time scales smaller than the model grid-box and smaller than the model time step, therefore those processes are not represented explicitly in climate models. As a consequence, on a longer term (hundred years), the projections on how clouds and precipitation will evolve in the future differ amongst models (Vial et al., 2016). Observations collected by field experiments and ground-based sites have provided essential knowledge on how the atmospheric water cycle works at small scale (<100 m) (Campbell et
10 al., 2002; Intrieri et al., 2002; Shupe et al., 2006; Long et al., 2009; Wild, 2009; Manara et al., 2016), but these observations are sparse and limited in space. Thanks to their global cover and their long life-time, satellites have observed the water cycle components on a global scale for over 25 years (Gruber and Levizzani, 2008; Raschke et al., 2012; Stubenrauch et al., 2013). However, these satellites lack some essential capabilities, such as documenting the detailed vertical structure of the water cycle components. Since 2006, the space lidar CALIPSO (Winker et al., 2017) and the space radar CloudSat (Stephens et al., 2008)
15 provide a more detailed view of aerosols, clouds, and precipitation (light rain and snow), on a global scale. These active sensors provide new surface-blind detailed vertical profiles of aerosols (Liu et al., 2009; Sekiyama et al., 2010), clouds (Mace et al., 2009; Vaughan et al., 2009; Chepfer et al., 2010), snow precipitation (Palerme et al., 2014), Arctic atmosphere (Kay et al., 2008; Cesana and Chepfer, 2012), light rain precipitation (Lebsock and L'Ecuyer, 2011), atmospheric heating rate profiles and surface radiation (Kato et al., 2011; Stephens et al., 2012).

20 Similarly, atmospheric reanalyses, although suited for the study of integrated contents of water vapour (Obligis et al., 2009; Schröder et al., 2017), exhibit noticeable biases in the tropical water and energy budget on the vertical. As suggested by comparisons between satellite observations of single-layer upper tropospheric humidity and atmospheric reanalyses (Chuang et al., 2010; Chiodo and Haimberger, 2010), reanalyses fail to reproduce the observed vertical correlation structure between the various layers of relative humidity in the upper troposphere, where moisture is mainly influenced by the shape of the
25 convective detrainment profile in deep convective clouds (Folkins et al., 2002), together with drying effects induced by mixing or air intrusion from the subtropics (Pierrehumbert, 1998; Brogniez et al., 2009). On the other hand, since 2011, the passive microwave sensor 'Sondeur Atmosphérique du Profil d'Humidité Intertropical par Radiométrie' (SAPHIR), provides over the entire tropical belt (30°S - 30°N) observations of water vapour even in the presence of (non-precipitating) clouds, which are largely transparent at frequencies above 100 GHz (Brogniez et al., 2015). These detailed profiles are observed all over the
30 globe, and thus are good candidates to help improving our current understanding on how the atmospheric water cycle works.

However, if the new generation of space clouds observations has the relevant spatial resolution (60 m on the vertically, 333 m horizontally, (Chepfer et al., 2010)) and the global cover to document processes over the entire Earth, the water vapour observations do not. The water vapour measured by SAPHIR is observed at larger spatial resolutions (with a footprint size at nadir of 10 km) which implies that small scale heterogeneities will be missed, critical for understanding the full water
35 cycle processes. To better understand the atmospheric water cycle, and the multi-scales interplays, it is thus of strong interest

to build a pseudo-observations dataset that contains, over the entire tropical belt and during several years, simultaneous co-located profiles of water vapour and clouds at high spatial resolution relevant to process studies (480 m vertically and 330 m horizontally, Chepfer et al. (2010)). It is the purpose of this paper to build such a pseudo-observation dataset.

Among the clouds forming in the troposphere, tropical ice clouds are of particular interest, because of their extensive

5 horizontal and vertical coverage and their long lifetime (Sassen et al., 2008), and above all because they are intimately related to water vapour (Udelhofen and Hartmann, 1995). Although the approach that we present in this study could in principle be extended to other cloud types, here we decided to focus on ice clouds, for which the connection to water vapour is expected to be stronger (Luo and Rossow, 2004; Tian et al., 2004).

When combining measurements from different platforms care must be taken to account for the different spatial resolutions

10 of the instruments (Atkinson , 2013). For spaceborne instruments, the horizontal spatial resolution or support is determined by the sensor's instantaneous field of view and is approximately equal to the size of a pixel in an image provided by that sensor. Although ideally we would like all spaceborne measurements to have the finest possible horizontal spatial resolution, in practice there is a limit imposed by the trade-off between spatial resolution, revisit time and spatial coverage: on the one hand, CALIPSO and CloudSat provide images with a fine horizontal spatial resolution (see section 2.2) but have a sparse coverage

15 and a long revisit time due to their polar orbiting; on the other hand, SAPHIR, owing to the low inclination of its orbit, is characterized by a much higher revisit frequency and a more complete coverage, but has a lower horizontal spatial resolution (see section 2.1). The support therefore provides a limit on what a spaceborn sensor can retrieve and effectively acts as 'filter on reality' (Atkinson , 2013): different instruments with different supports will indeed view the Earth differently.

Statistical downscaling methods (Bierkens et al., 2000; Vaittinada et al., 2015) involve reconstructing a coarse-scale measured

20 variable at a finer resolution based on statistical relationships between large- and local-scale variables. Although the typical application for these methods is to derive sub-grid scale climate estimates from GCMs outputs or reanalysis data to drive impact studies (Gutierrez et al., 2018), recent studies have started adopting the standard downscaling techniques to enhance the resolution of satellite images using available covariate data at a finer resolution (Liu and Pu, 2008; Malone et al., 2012). Following the approach taken in these studies, here we are interested in modelling, at the finer scale of the clouds measurements,

25 the statistical relationship between the water vapour layered-vertical structure associated to ice clouds in the tropical belt and the vertical profiles of clouds provided by CALIPSO. The method employed in this study provides a general framework to effectively perform a downscaling of SAPHIR observations of relative humidity and, for unsampled locations and times, to predict the (downscaled) water-vapour vertical structure using cloud profiles only.

The paper is organized as follows. In section 2 we present the satellite data sources used in this study; section 3 describes

30 the general approach and the methods used to downscale water vapour observations from vertical cloud profiles; results are discussed in section 4 and finally, conclusions and future perspectives are drawn.

2 Data

2.1 SAPHIR

SAPHIR is a cross-track passive microwave sounder on board the Megha-Tropiques mission. It observes the Earth's atmosphere with an inclination of 20 degrees to the equator, a footprint size at nadir of $10x10 \, \mathrm{km}^2$, with a 1700-km swath made of scan lines
5   containing 130 non-overlapping footprints (for more details see e.g. Brogniez et al. (2016) and references therein). Operating since October 2011, SAPHIR provides indirect observations of the relative humidity (RH) in the tropics (28°S - 28°N) by measuring the upwelling radiation with six double-sideband channels close to the 183.3-GHz water vapour absorption. In this work, we used the layer-averaged RH (six layers distributed between 100 and 950 hPa) derived by Brogniez et al. (2016). In this study, the authors adopted a purely statistical technique to retrieve for each atmospheric layer the full distribution of
10   RH from the space-borne observations of the upwelling radiation and training RH data derived from radiosondes profiles. This retrieval scheme was found to have similar performances compared to other methods that also rely on some other physical constraints (e.g. the surface emissivity, temperature profile, and a prior for RH profiles for brightness temperature simulations). Figure 1a, shows an example, for each atmospheric layer, of the mean of the retrieved RH distribution, derived as detailed in Brogniez et al. (2016).

15   2.2 CALIPSO

The lidar profiles in the GCM-Oriented Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) Cloud Product (CALIPSO-GOCCP, Chepfer et al. (2010)), are designed to compare in a consistent way the cloudiness derived from satellite observations to that simulated by General Circulation Models (GCMs, (Chepfer et al., 2008)). CALIPSO is a nearly sun-synchronous platform that crosses the equator at about 01:30 LST (Winker et al., 2009) and carries aboard the
20   Cloud-Aerosol LIdar with Orthogonal Polarization (CALIOP). CALIOP measures, every 330 m along track with a foot size of 90 m, the Attenuated Backscattered (ATB) profile at 532 nm. The lidar scattering ratio (SR) is measured relative to the backscatter signal that a molecular atmosphere (without clouds or aerosols) would have produced. Within a cloud the SR value represents a signature of the amount of condensed water within each layer convoluted with the optical properties of the cloud particles that depend on their size and shape. Values of SR greater than five are taken as indications of layers containing clouds
25   (Fig. 1b, see Chepfer et al. (2010) for more details). On the other hand, values of SR lower than 0.01 correspond to layers that are not documented by CALIPSO. Indeed, layers located below clouds opaque to radiations are not sounded by the laser (Guzman et al., 2017; Vaillant de Guélis et al., 2017).

Following Chepfer et al. (2010), layers corresponding to values located below the surface ($SR = -888$), rejected values ($SR = -777$), missing values ($SR = -9999$) and noisy observations ($-776 < SR < 0$) were all set to missing. Moreover, in
30   order to reduce the noise and the number of missing data, each SR profile (40 equidistant layers with height interval of 480 m) was averaged as the following: in the boundary layer (below 2 km), the original vertical spacing was used (four layers in total), while, above, the layers were averaged every 1 km, giving in total $p = 21$ vertical layers. Only the averaged SR profiles without any missing layer were retained: the choice of setting to missing all noisy layers, implies retaining mostly night-time

data only (after excluding the averaged profiles with missing layers, the percentage of day-time profiles dropped from about 50% to less then 15%). Finally, in order to further reduce the noise in the profiles, these were transformed using a Principal Component Analysis (PCA) analysis, where 90% of the variance was retained.

3 Methods

A three-step method was applied to downscale water vapour observations from vertical cloud profiles. First, we co-located SAPHIR and CALIPSO observations (section 3.1); then, using a statistical clustering technique, we selected only CALIPSO profiles corresponding to ice clouds 3.2), and finally we applied the downscaling method (section 3.3).

3.1 SAPHIR-CALIPSO co-location

To identify the times and locations where the orbits of SAPHIR and CALIPSO overlap, we first extracted all the observations at Nadir falling within a distance of 50 km and within 30 min (for details of the software used for the co-location of the orbits see http://climserv.ipsl.polytechnique.fr/ixion). SAPHIR measurements (both at and off-Nadir) corresponding to the selected orbits were then matched to CALIPSO observations falling within each SAPHIR pixel, defined as the 10 km circle around its geographical coordinates (see Fig. 1c). In the following analysis, each SAPHIR measurement at coarse resolution ($M = 1, ..., N$) encapsulates $n(M)$ CALIPSO observations at fine scale ($m = 1, ..., n(M)$), where $n(M)$ changes depending on the spatial alignment of the two satellites. Fig. 2 shows a sample of the co-located CALIPSO and SAPHIR profiles.

The following analysis is focused on tropical ice clouds over the ocean, where the correlation between cloud profiles and tropospheric humidity is expected to be larger (Schröder et al., 2017).

3.2 Selection of tropical ice cloud profiles

In order to select only profiles characterized by tropical ice clouds, the co-located samples were separated into clusters based on indicators of the type of clouds present at the moment of the observation.

The clusters were obtained by a k-means unsupervised classification of the reconstructed SR profiles (e.g. Lloyd (1982)). Since layers with SR values in the same range are associated to the same micro-physical properties, for clustering only, the recontructed SR profiles were binned according to the interval boundaries suggested in Chepfer et al. (2010), as detailed in Fig. 5 in their study. Given an optimal number of clusters (k), this method partitions the observations into k clusters with each observation belonging to the cluster with the nearest mean by minimizing the within-cluster-sum of squares (wss). Since the initial assignment of the observations to a cluster is random, the algorithm is run several times (here 100) and the partition with the smallest wss is chosen amongst the different ensemble members. However, when k is not known a priori, it must be selected from a range of plausible values (here: $k \in \{2, \ldots, 15\}$), and chosen so that adding another cluster does not produce a drastic decrease in wss, and therefore does not improve significantly the quality of the clustering. For example for reconstructed SR profiles in July 2013 over the Indian Ocean, this criterion yields between 8 and 13 clusters (not shown).

As Fig. 3 shows, both clusters 1 derived by k-means with k=8 and k=13 show a similar mean SR profile, with layers classified as cloudy mostly in the upper troposphere. As a further check that these profiles correspond indeed to ice clouds, we compared the k-means result with the clusters derived by combining the cloud phase flags associated with each vertical level as defined in Cesana and Chepfer (2013) (e.g. a profile corresponding only to clear-sky and liquid observations is classified as

5   LIQUID, see caption in Fig. 3 for more details). As Fig. 3 shows, again a similar characteristic SR profile is observed for the flag-based profiles corresponding to ICE and ICE-MIX observations. Therefore, in the following, the k-means classification is used to select all SAPHIR-CALIPSO co-located observations belonging to SR clusters characterized by this typical mean SR profile (in Fig. 3, clusters outlined by a red square).

3.3   Downscaling of water vapour measurements from cloud profiles

10 Given the SAPHIR-CALIPSO co-located samples belonging to ice cloud-type clusters as derived in the previous section, SAPHIR relative humidity at the l-th pressure level (RH_l, here corresponding to the mean of the distribution in Brogniez et al. (2016)) can be estimated in terms of an unknown function Φ of the SR profile

$$RH_l \sim \Phi(SR_1, SR_2, ..., SR_p) \tag{1}$$

where $SR_1, SR_2, ..., SR_p$ designate SR at each altitude level ($p = 21$, following the vertical averaging implemented as

15 described in section 3.2) and here represent the covariate data sources, also known as predictors. The method to downscale SAPHIR observations of relative humidity from CALIPSO SR profiles consists in a two-stage regression model implemented directly on the observed spatial resolution (Liu and Pu, 2008; Malone et al., 2012). First, RH_l is estimated based on the chosen statistical regression model (section 3.3.1). Secondly, the same regression model is applied iteratively to the predictions $\widehat{RH_l}$ and at each iteration step the multi-site results are corrected to harmonize the average of the estimates at fine resolution with

20 its value at coarser scale (section 3.3.2).

This downscaling scheme differs from the classical downscaling approach, where local variables, generally point-scale observations, are generated from large-scale variables, available at the much coarser grid-scale resolution typical of climate models and reanalyses outputs, based on a model trained on the available local variables. This approach cannot be applied in the case under study, since there are no RH observations at the resolution of cloud measurements available. On the other

25 hand, by including covariates at a finer resolution in the regression model of Eq. (1) we can incorporate their higher-resolution variability in the estimates of the response variable (here RH), while maintaining, through the iterative procedure, the so-called 'mass balance' with the original measured values.

3.3.1   Choice of the regression model

The aim of this section is to compare different regression models for RH_l given the set of predictors $SR_1, SR_2, ..., SR_p$ and to

30 select the model with the 'best' predictions in a sense that will be clarified later. The models tested in this study are summarized in Table 1.

Random Forests (RF, Breiman (2001)), similarly to other machine learning techniques, does not require to specify the functional form of the relationship between the response variable and the predictors and, provided a large learning sample, has been shown to perform well (Hastie and Tibshirani, 2009) in the context of prediction of a response variable even with a non-linear relationship with a set of predictors. RF belongs to the family of classification and regression decision trees (Breiman

5   et al., 1984). Decision trees split the predictor space into boxes (or leaves) such that the homogeneity of the corresponding values of the response variable in each box is maximized. For regression trees, the homogeneity is defined as the sum of the residual-sum of squares (rss) with respect to the mean of the response variable within each box. As described in detail for example in Hastie and Tibshirani (2009), this method is implemented by sequentially splitting the predictor space into the regions $x_i < c$ and $x_i \geq c$ where the predictor x_i and the cutting-point c give the greatest possible reduction in rss. This binary

10 split is repeated until a minimum number of observations in each leaf is reached or because of an insufficient decrease in rss. Another possibility, which prevents overfitting, is to grow a tree with a large number of leaves but prune it at each split by controlling the trade-off of between the tree complexity (i.e. the number of leaves) and the fit to the data. Finally, the model estimate of the response variable is given by the mean of all the observations in each terminal leaf and for predictions for a new set of values of the predictors, one has then simply to follow the path in the tree until the final leaf is found. In order to reduce

15 the variance in the predictions, Breiman (1996) proposed to grow a tree on several bootstrapped samples of the original data and then take the average result from the different trees (*bagging*). This approach is justified by the property that by taking the average of N independent observations with variance σ^2 we reduce the variance by σ^2/N. To avoid overfitting, the number of bootstrapped samples and that of the corresponding trees can be adjusted, while the trees are not pruned. With RF, the variance in the predictions can be even further reduced by retaining at each split a random selection from the full set of predictors,

20 therefore reducing the correlation between the trees generated by bootstrapping only.

Bagging and RF only estimate the conditional mean of the response variable but not its distribution, which can give information on the uncertainty in the predictions. On the other hand, Quantile Regression Forests (QRF, Meinshausen (2006)), by computing the Cumulative Distribution Function (CDF) of the response variable in each terminal leaf instead of its mean, represent a straightforward extension of the RF method, allowing to estimate any quantile of the response variable.

25 Non-parametric methods, like RF and QRF, do not allow to specify the functional form of the relationship between the response variable and the predictors. For this reason, we also tested the results obtained with a Generalized Additive Model (GAM, Hastie and Tibshirani (1986)), which is a statistical semi-parametric regression technique. A GAM is a Generalized Linear Model (GLM) with predictors involving a sum of non-linear smooth functions:

$$g\left(E\left[y|\mathbf{x}\right]\right) = \sum_{i=1}^{p} f_i\left(x_i\right) + \varepsilon \qquad (2)$$

30 where $g(\cdot)$ is a link function between the expectation of the response variable y (here the RH of an atmospheric layer l) conditionally on a set of p predictors $x_1, ..., x_p$ (here $SR_1, ..., SR_p$) and a sum of unknown univariate smooth functions of each predictor, $f_i(\cdot)$. ε represents a zero-mean Gaussian noise. Here, RH_l is assumed to follow a beta distribution, which is the usual choice for continuous proportion data, and its canonical link function, the logit $g(x) = \log\left(\frac{x}{1-x}\right)$, is used (Wood, 2011),

which assures that all values are in the (0,1) interval. To estimate each f, we can represent it as a weighted sum of known basis functions $z_k(\cdot)$

$$f(x) = \sum_k \beta_k z_k(x) \tag{3}$$

in such a way that Eq.(2) becomes a linear model, and only the β_k are unknown. Here, we chose to represent the basis functions as piecewise cubic polynomials joined together so that the whole spline is continuous up to second derivative. The borders at which the pieces join up are called knots, and their number and location control the model smoothness. To fit the model in Eq. (2), we used the approach of Wood (2011): the appropriate degree of smoothness of each spline is determined by setting a maximal set of evenly spaced knots (i.e. $bias(f) \ll var(f)$) and then controlling the fit by regularization, by adding a 'wiggliness' penalty $\int f''(x)dx = \boldsymbol{\beta}^T S \boldsymbol{\beta}$ to the likelihood estimation:

$$\mathcal{L}(\boldsymbol{\beta}) - \boldsymbol{\beta}^T \mathbf{S} \boldsymbol{\beta} \tag{4}$$

where $\mathcal{L}$ is the likelihood function of the $\boldsymbol{\beta}$ parameters and $\mathbf{S}$ the penalty matrix, with elements for the kth - $\widetilde{k}$th terms $S_{k\widetilde{k}} = \int z''_k(x) z''_{\widetilde{k}}(x) \, dx$.

Ideally, we would like to account for a neighbouring structure, i. e. neighbouring SR profiles should be characterized by similar model parameters. This effect can be accounted for by assuming, under the Markovian property, that the model parameters for the mth profile are independent of all the other parameters given the set of its neighbours $\mathcal{N}(m)$. This neighbouring structure can then be modelled by adding to Eq. (2) a smooth term with penalty

$$\Gamma(\boldsymbol{\gamma}) = \sum_{m=1}^{n} \sum_{\widetilde{m} \in \overline{\mathcal{N}(m)}} (\gamma_m - \gamma_{\widetilde{m}})^2 \tag{5}$$

where γ_m is the smooth coefficient for region m and $\overline{\mathcal{N}(m)}$ denotes the elements of $\mathcal{N}(m)$ for which $\widetilde{m} > m$. The penalty in Eq. (5) can be then rewritten as $\Gamma(\boldsymbol{\gamma}) = \boldsymbol{\gamma}^T \mathbf{S} \boldsymbol{\gamma}$ with $S_{m\widetilde{m}} = -1$ if $\widetilde{m} \in \mathcal{N}(m)$ and $S_{m\widetilde{m}} = n(m)$ where $n(m)$ is the number of profiles neighbouring profile m (not including m itself). This specification is very computationally efficient, given the sparsity of the parameters precision matrix, and is known as Gaussian Markov random field (GMRF, Rue and Held (2005)). Here, we implemented this augmented model by defining two CALIPSO SR profiles as neighbours if they belong to the same SAPHIR pixel.

Another possibility, although more computationally expensive, is to explicitly include in our model the spatial correlation structure of the predictors by a fusion of geostatistical and additive models, known as geoadditive models (Kammann and Wand, 2003). These models allow accounting not only for the non-linear effects of the predictors (under the assumption of additivity) but also for their spatial distribution: two SR profiles, and therefore the corresponding water vapour structures, are more likely to be dependent if they are close, by some metric. Given a set of geographical locations $\mathbf{s}$, a (bivariate) smooth

term $f(\mathbf{s})$ can be represented as the random effect $f(\mathbf{s}) = (1, \mathbf{s}^T)\boldsymbol{\gamma} + \sum_j w_j C(\mathbf{s}, \mathbf{s}_j)$ with $w \sim N(0, (\lambda C)^{-1})$, $\boldsymbol{\gamma}$ a vector of parameters and $C(\mathbf{s}, \mathbf{s}_j) = c(||x - x_j||)$ a non-negative function such that $c(0) = 1$ and $\lim_{d \to \infty} c(d) = 0$, which is interpretable as the correlation function of the smooth f (Wood, 2011). By adding this term to the model in Eq. (2), we explicitly include the spatial autocorrelation in the SR data without changing the mathematical structure of the minimization problem, and we can still use the GAM basis-penalty representation (Wood, 2011). Here, we assumed an isotropic exponential correlation function $C(\mathbf{s}, \mathbf{s}_j) = \exp(- \parallel \mathbf{s} - \mathbf{s}_j \parallel /r)$ with the range r chosen equal to the size of SAPHIR pixels (10 km).

Following Ferro (2008), Ferro et al. (2014), and Taillardat et al. (2016), to assess the prediction skills of such models, scoring rules can be used to assign numerical scores to probabilistic forecasts and measure their predictive performance. Given an observation y, for a model ensemble forecast with members $x_1, .., x_K$ a fair estimator (Ferro et al., 2014) of the continuous ranked probability score (CRPS) is

$$CRPS(y) = \frac{1}{K} \sum_{i=1}^{K} | x_i - y | - \frac{1}{2K(K-1)} \sum_{i=1}^{K} \sum_{j=1}^{K} | x_i - x_j | \qquad (6)$$

where lower values of the CRPS indicate better predictive skills. For regression techniques that estimate the conditional mean only (RF, GAM, GAM with GRMF, and the geoadditive method), the CRPS score accounts only for the accuracy of the forecast (the second term in Eq. (6) is zero), while for probabilistic methods, like the QRF method, it also accounts for the forecast precision. Typically, in order to directly compare a prediction system to a reference forecast (e.g. a climatology), the continuous ranked probability skill score (CRPSS) is needed

$$CRPSS = 1 - \frac{CRPS_{mod}}{CRPS_{ref}} \qquad (7)$$

The CRPSS is positive if and only if the model forecast is better than the reference forecast for the CRPS scoring rule.

3.3.2 Iterative downscaling

Following the approach of Liu and Pu (2008) and Malone et al. (2012), the predictions were further optimized by ensuring that, for all layers, the observed relative humidity is as close as possible to the average of the predicted RH distributions within the corresponding encapsulating SAPHIR pixel. This approach is meant to preserve the so-called 'mass balance' with the coarse scale SAPHIR information, and can be easily implemented with the following iterative approach:

1 within each SAPHIR pixel (M), update the predictions $\widehat{RH_l}$: $\widetilde{RH_l}(m) = \widehat{RH_l}(m) + RH_l(M) - \frac{1}{n(M)} \sum_{j \in n(M)} \widehat{RH_l}(j)$

2 with the chosen regression model, regress the updated predictions $\widetilde{RH_l}$ with respect to the set of predictors $SR_1, SR_2, ..., SR_p$

3 if the coefficient of determination (R^2) with respect to the observed relative humidity $RH_l(M)$ of the updated predictions is larger than that of the previous iteration than repeat steps [1]-[2], otherwise stop at previous iteration.

For ensemble models, like QRF, the update predictions and R^2 are computed on the median of the distribution only.

4   Results and discussion

Figure 4 shows, for ice cloud profiles in the Indian Ocean in July 2013 (k=8), the comparison of the CRPSS computed for the forecast derived for the different regression methods (QRF, RF, GAM, GAM with GRMF, and the geoaddive method) with respect to the reference CRPS computed from the empirical distribution of the observations. In order to validate the regression results with independent test data, the predictions were performed using a 5-fold cross validation scheme and, to reduce the computation time, were limited to the first iteration step (here we are interested in comparing the performance of the different models rather than performing the full downscaling). For the RF and QRF method, the sensitivity of the results to the model parameters (number of trees and number of predictors selected at random at each split) was also investigated using a grid search; however, for both models, variations in the prediction skills (both in terms of R^2 and the CRPSS score) were found negligible with respect to the choice of these parameters, that were therefore set to their default values (c.f. the `randomForest` R package, R Core Team (2017)). The largest CRPSS is obtained using the QRF method, with a median value larger than 0.5 for all layers. The RH predicted with the RF method are also significantly better that what we would obtain from the empirical distribution of the observations, although the probabilistic approach taken in QRF is more skill-full. On the other hand, all GAM-derived methods have a lower score, with CRPSS median values overall below 0.5, although, apart from the highest and lowest layers, all medians are above zero. As the CRPSS reveals, full non-parametric methods that do not rely on any assumption on the probability distribution of the response and that are free to learn any functional form from the training data, perform significantly better.

A positive value of the CRPSS for all RH layers indicates a high level of correlation along the full vertical profile, which is expected for ice clouds: within and in the neighbourhood of regions of deep convection, which is their primary source (Hartmann et al., 2001), air masses are rapidly transported from the boundary layer through the free troposphere into the tropopause region (Corti et al., 2006). This is also shown in the scatter plot of the median of the QRF-predicted distribution vs. the observed relative humidity (Fig. 5): for layers L1-L5, the data are distributed close to the identity line, with the model explaining a large proportion of the variance of the observed RH ($R^2 \geq 0.7$). On the other hand, as expected for ice clouds which populate the upper troposphere, lower correlation values are found for the lowest layer (L6, $R^2 \sim 0.3$). To assess the importance of the cloud structure on the predicted relative humidity at different layers, we can compute, for each predictor, the decrease in accuracy obtained by randomly permuting its values (Fig. 6): the larger this value, the more important a predictor is. For the higher layers, as expected, this metric highlights the larger contribution of SR layers corresponding to layers classified as cloudy, which are observed above $\sim 10\,\mathrm{km}$ (c.f. Fig. 3). On the other hand, for layers closer to the surface, the contribution of lower, (on average) non-cloudy SR layers is found to be equally important because of the moisture that originates over warm waters. Similar results can be found for different choices of the number of clusters (k=13), season (January), and region (Pacific Ocean), as shown in Fig. 7. These results are also independent (not shown) on the temporal difference and the spatial alignment of the co-located samples, on the distance from the cost, or on the uncertainty (standard deviation) in the observed relative humidity by SAPHIR.

Overall, these results suggest that, at the instantaneous scale of cloud measurements, the water vapour response along the whole troposphere in correspondence with ice cloud profiles is well predicted only accounting for their capability to backscatter radiation (given by the observed SR profile). While the large-scale link between relative humidity and the cloud properties (vertical distribution, phase and opacity) has been well documented in previous studies (Martins et al., 2011; Reverdy et al.,

5   2012), this work represents the evidence that this relationship can also be detected at much smaller spatio-temporal scales. The emergence of a clear signal at these fine scales, also highlights the limitations of SAPHIR measurements: although SAPHIR observes the water vapour field at a much finer horizontal resolution than what is currently available in reanalysis products, in order to explain physical processes, downscaled observations are needed. Figure 8 compares, for a selection of ice cloud profiles ($n(M) > 25$), the corresponding layers of relative humidity observed by SAPHIR with the median of the downscaled results

10   derived by implementing the iterative QRF scheme. For all layers, the iteration typically stops after 2-3 steps and, although increases the R^2 between SAPHIR observations and the predicted relative humidity by only few percent, ensures consistency with the observed data, as described in section 3.3.2. The goal of the downscaling scheme implemented in this work is to reconstruct the variation of the relative humidity field at the fine resolution of cloud measurements within each SAPHIR coarsely resolved pixel: as Fig. 8 shows, the downscaled values exhibit variations within the same SAPHIR pixel depending

15   on the corresponding SR profile (Fig. 8c) that cannot be observed by SAPHIR (Fig. 8b). As discussed at the beginning of this section, a measure of the reliability of these variations can be derived from the spread of the predicted distribution, given here as the interquartile range (Fig. 8d).

The intra-pixel RH variations are further analysed in Fig. 9, which shows for a single SAPHIR pixel overlaid on the observed values, the downscaled predictions from the QRF and the geoadditive model. For the latter, the predictions were

20   extended outside the observed CALIPSO locations on the direction orthogonal to CALIPSO track line up to 1 km on each side. The relative humidity field at these new locations was predicted using the model fitted through the iterative scheme for the available CALIPSO observations and assuming that each SR profile was also representative of the cloud distribution for locations shifted along the direction orthogonal to CALIPSO track within a distance of 1 km. As expected and shown by Fig. 9b, the largest part of the variance is explained by the SR predictors, while variations related to the spatial smooth are almost

25   not noticeable with the scale used in the plot, compared to the variations in the predictions for a given SR profile.

Although the CRPSS quantifies the quality of the predictions (w.r.t. the climatology) conditionally on the regression model and the predictors, for direct validation, observations of relative humidity at the scale of the clouds measurements would be required. In principle, the network of radiosonde measurements, which provides RH quality-checked data (Durre et al., 2006) and has been used in previous studies for validation of satellite measurements, including SAPHIR (Sivira et al., 2015; Brogniez

30   et al., 2016), could be used for validation purposes. However, in practice, its limited spatial coverage, with also most of the observations falling over land, hampers the feasibility of this approach. On the other hand, probabilistic approaches, like the QRF method, by assessing the uncertainty in the predictions through the spread of the distribution, allow the quantification of the confidence in those predictions and therefore, in a way, provide an indirect estimate of their quality.

5 Conclusions

We have presented a method to downscale observations of relative humidity (RH) available from the passive microwave sounder SAPHIR at a nominal horizontal resolution of 10 km to the finer resolution of 90 m using scattering ratio (SR) profiles from the lidar CALIPSO. The method was applied to ice clouds profiles over the tropical oceans, where the connection
5 to water vapour is expected to be stronger.

By using an iterative regression model of the satellite-derived RH with the SR profiles as covariates, we were able to successfully predict the relative humidity along the whole troposphere at the resolution of cloud measurements. The method also ensures that the average of the predicted RH distributions within the corresponding encapsulating SAPHIR pixel is as close as possible to the observed value. Amongst the different regression models tested, the best results were obtained using a
10 Quantile Random Forest (QRF) method, with a coefficient of determination (R^2) with respect to the observed relative humidity larger than 0.7 and a CRPSS with respect to the climatology with a median value larger than 0.5 for all layers down to 800 hPa. High explanatory power along the full vertical profile is expected for ice clouds, for which deep convection, by transporting air masses from the boundary layer up to the tropopause region, is their primary source.

By providing a method to generate profiles of water vapour (at high spatial resolution) from simultaneous co-located cloud
15 profiles, this work will be of great help to revisit some of the current key barriers in atmospheric science. While SAPHIR record only stretches back to 2011, CALISPO cloud measurements are available since 2006, a period that includes three El Niño/Southern Oscillation (ENSO) cycles. A 10-year long high resolution water vapour-clouds combined dataset might allow:

- to study how small scale water cycle processes behave when exposed to strong variations in large scale circulation regimes such as those associated to El Niño cycles

20 - to 'evaluate' how small scale water vapour inhomogeneities affect the water vapour simulated by standard reanalyses (e.g. ERA-Interim Dee et al. (2011), NCEP Kalnay et al. (1996), etc.), which are known to badly parameterize clouds and to have biases in water vapour in the upper troposphere (Schröder et al., 2017)

- to put the results of past and current field experiments into a larger scale context, e.g. identifying if results of specific campaigns are representative of large portions of the tropical belt

25 - to guide the parametrization of unresolved subgrid-scale water vapour/clouds processes to reduce cloud feedback uncertainties (Randall et al., 2003) in climate models which ultimately will contribute to improve model-based estimates of climate sensitivity

- to evaluate the description of water vapour/cloud interactions in regional models - e. g. WRF, Meso-NH (Chaboureau et al., 2002; Fan et al., 2007), which although having a fine-enough grid-spacing to allow explicit simulations of the
30 mesoscale dynamics associated with convective clouds (Guichard and Couvreux, 2017) still integrate parametrizations to represent sub-grid-scale motions, micro-physics, and radiative processes

- to test the validity of the fixed anvil temperature hypothesis (Hartmann and Larson, 2002) and estimate the changes to long-wave fluxes with warming, for example using simulated CALIPSO profiles from model variables (Chepfer et al., 2008)

- to quantify the limits of current and future space missions by characterizing the spatial inhomogeneities in water vapour fields that cannot be observed by present satellites and will likely not be observed within the next tow decades (e.g. 2017-2027 Decadal Survey for Earth Science and Applications from Space) due to technological limits.

We also note that the method developed in this study could in principle be extended to other types of clouds, although additional covariates might be required (e.g., for liquid clouds, including the radar reflectivity as measured by the radar CloudSat, might increase the model explanatory power in the presence of light precipitation). Finally, the downscaling method
10 presented here could be also applied to other satellite products, with the underlying assumption of using covariate data that are strongly related to the target variable.

Sample availability. A sample of co-located SAPHIR and CALIPSO observations, together with the downscaled relative humidity predictions, is available using the following digital object identifier (doi): http://dx.doi.org/10.14768/20181022001.1. This dataset corresponds to ice cloud profiles in July 2013 over the whole tropical oceans, derived assuming $k=13$ number of clusters.

15 *Author contributions.* GC developed the methodology, and drafted the manuscript. MV, HB, PY, and HC supervised and supported the development of the methodology and provided detailed comments on the manuscript.

Competing interests. The authors declare no competing interest.

Acknowledgements. The authors are thankful to Patrick Raberanto (Laboratoire de Météorologie Dynamique) for his help with the co-location of SAPHIR and CALIPSO orbits. The authors would like also to thank the IPSL mesocenter and ESPRI teams from IPSL for providing
20 computing and storage resources, and CNES and NASA for providing SAPHIR and CALIPSO Level 1 data. GC was supported by the Paris-Saclay Initiative de Recherche Strategique SPACEOBS (ANR-11-IDEX-0003-02). The authors also acknowledge the support of CNES, program EECLAT.

References

Atkinson, P. M.: Downscaling in remote sensing, International Journal of Applied Earth Observation and Geoinformation, 22, 106-114, doi: 10.1016/j.jag.2012.04.012, 2013.

Bierkens, M. F. P., Finke, P. A., and, De Willigen, P.: Upscaling and Downscaling Methods for Environmental Research, Kluwer Academic, Dordrecht, The Netherlands, 2000.

Bodas-Salcedo, A., Webb, M.J., Bony, S., Chepfer, H., Dufresne, J., Klein, S.A., Zhang, Y., Marchand, R., Haynes, J.M., Pincus, R., and John, V.O.: COSP: Satellite simulation software for model assessment, Bull. Amer. Meteor. Soc., 92, 1023–1043, doi: 10.1175/2011BAMS2856.1, 2011.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P.,Satheesh, S.K., Sherwood, S., Stevens, B., and Zhang, X.Y.: Clouds and aerosols. In Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. T.F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Doschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, Eds. Cambridge University Press, 571-657, doi: 10.1017/CBO9781107415324.016, 2013.

Brasseur, G., and Jacob, D.: Parameterization of Subgrid-Scale Processes. In Modeling of Atmospheric Chemistry, 342-398, Cambridge: Cambridge University Press, doi: 10.1017/9781316544754.009, 2017.

Breiman, J. F., Stone C. J., and Olshen R. A.: Classification and Regression Trees, CRC Press, 368 pp., 1984.

Breiman, L.: Bagging predictors, Mach. Learn., 24 (2), 123-140, 1996.

Breiman, L.: Random forests, Mach. Learn., 45 (1), 5-32, doi: 10.1023/A:1010933404324, 2001.

Brogniez, H., Roca, R., and Picon, L.: A Study of the Free Tropospheric Humidity Interannual Variability Using Meteosat Data and an Advection–Condensation Transport Model, J. Climate, 22, 6773–6787, doi: 10.1175/2009JCLI2963.1, 2009.

Brogniez H., Clain, G., and Roca, R.: Validation of Upper Tropospheric Humidity from SAPHIR onboard Megha-Tropiques using tropical soundings, J. Appl. Meteorol. Climat., 54, 896-908, doi: 10.1175/JAMC-D-14-0096.1, 2015.

Brogniez, H., Fallourd, R., Mallet, C., Sivira, R., and Dufour, C.: Estimating confidence intervals around relative humidity profiles from satellite observations: Application to the SAPHIR sounder, J. Atmospheric Ocean. Technol., 33(5), 1005-1022, doi:10.1175/ JTECH-D-15-0237.1, 2016.

Campbell, J.R., Hlavka, D.L., Welton, E.J., Flynn, C.J., Turner, D.D., Spinhirne, J.D., Scott, V.S., and Hwang, I.H.: Full-Time, Eye-Safe Cloud and Aerosol Lidar Observation at Atmospheric Radiation Measurement Program Sites: Instruments and Data Processing, J. Atmos. Oceanic Technol., 19, 431–442, doi: 10.1175/1520-0426(2002)019<0431:FTESCA>2.0.CO;2, 2002.

Cesana G., and Chepfer, H.: Evaluation of the cloud water phase in a climate model using CALIPSO-GOCCP, J. Geophys. Res., doi: 10.1002/jgrd.50376, 2013.

Cesana, G., Chepfer, H., Winker, D.M., Getzewich, B., Cai, X., Okamoto, H., Hagihara, Y., Jourdan, O., Mioche, G., Noel, V., M. Reverdy: Using *in situ* airborne measurements to evaluate three cloud phase products derived from CALIPSO, J. Geophys. Res. Atmos., 121, 5788–5808, doi: 10.1002/2015JD024334, 2016.

Chepfer, H., Bony, S., Winker, D., Chiriaco, M., Dufresne, J.-L., and Sèze, G.: Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model, Geophys. Res. Lett., 35, L15704, doi: 10.1029/2008GL034207, 2008.

Chepfer, H., Bony, S., Winker, D., Cesana, G., Dufresne, J. L., Minnis, P., Stubenrauch, C. J., and Zeng, S.: The GCM-Oriented CALIPSO Cloud Product (CALIPSO-GOCCP), J. Geophys. Res., 115, D00H16, doi: 10.1029/2009JD012251, 2010.

Cesana, G., and Chepfer, H.: How well do climate models simulate cloud vertical structure? A comparison between CALIPSO-GOCCP satellite observations and CMIP5 models, Geophys. Res. Lett., 39, L20803, doi: 10.1029/2012GL053153, 2012.

Chaboureau, J.-P., Cammas, J.-P., Mascart, P. J., Pinty, J.-P., and Lafore, J.-P.: Mesoscale model cloud scheme assessment using satellite observations, J. Geophys. Res., 107(D16), doi: 10.1029/2001JD000714, 2002.

5   Chiodo, G., and Haimberger, L.: Interannual changes in mass consistent energy budgets from ERA-Interim and satellite data, J. Geophys. Res., 115, D02112, doi: 10.1029/2009JD012049, 2010.

Chuang, H., Huang, X., and Minschwaner, K.: Interannual variations of tropical upper tropospheric humidity and tropical rainy-region SST: Comparisons between models, reanalyses, and observations, J. Geophys. Res., 115, D21125, doi: 10.1029/2010JD014205, 2010.

Collis, R. T. H., Russell, P. B.: Lidar measurement of particles and gases by elastic backscattering and differential absorption, Hinkley E. D.

10   (eds) Laser Monitoring of the Atmosphere, Topics in Applied Physics, 14, Springer, Berlin, Heidelberg, doi: 10.1007/3-540-07743-X_18, 1976.

Corti, T., Luo, B. P., Fu, Q., Vömel, H., and Peter, T.: The impact of cirrus clouds on tropical troposphere-to-stratosphere transport, Atmos. Chem. Phys., 6, 2539-2547, doi: 10.5194/acp-6-2539-2006, 2006.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P.,

15   Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q.J.R. Meteorol. Soc., 137: 553–597, doi: 10.1002/qj.828, 2011.

Dufresne, J., and Bony, S.: An Assessment of the Primary Sources of Spread of Global Warming Estimates from Coupled Atmosphere–Ocean

20   Models, J. Climate, 21, 5135–5144, doi: 10.1175/2008JCLI2239.1, 2008.

Durre, I., Vose, R. S., and Wuertz, D. B.: Overview of the Integrated Global Radiosonde Archive, J. Climate, 19, 53–68, doi: 10.1175/JCLI3594.1, 2006.

Fan, J., Zhang, R., Li, G., and Tao, W.-K.: Effects of aerosols and relative humidity on cumulus clouds, J. Geophys. Res., 112, D14204, doi: 10.1029/2006JD008136, 2007.

25   Ferro, C., Richardson, D. S., and Weigel, A. P.: On the effect of ensemble size on the discrete and continuous ranked probability scores, Meteor. Appl., 15, 19-24, doi: 10.1002/met.45, 2008.

Ferro, C.: Fair scores for ensemble forecasts, Quart. J. Roy. Meteor. Soc., 140, 1917-1923, doi: 10.1002/qj.2270, 2014.

Folkins, I., Braun, C., Thompson, A. M., and Witte, J.: Tropical ozone as an indicator of deep convection, J. Geophys. Res., 107(D13), doi: 10.1029/2001JD001178, 2002.

30   Fu, Q., Krueger, S. K., and Liou, K. N.: Interactions of radiation and convection in simulated tropical cloud clusters, J. Atmos. Sci., 52, 1310-1328, doi: 10.1175/1520-0469(1995)052<1310:IORACI>2.0.CO;2 , 1995.

Gruber, A., and Levizzani, V.: Assessment of global precipitation products, WCRP Series Report 128 and WMO TD-No. 1430, WMO: Geneva, Switzerland, 2008.

Guichard, F., and Couvreux, F.: A short review of numerical cloud-resolving models, Tellus A: Dynamic Meteorology and Oceanography,

35   69:1, 1373578, doi: 10.1080/16000870.2017.1373578, 2017.

Gutierrez, J.M., Maraun, D., Widman, M., Huth, R., Hertig, E., Benestad, R., Roessler, O., Wibig, J., Wilcke, R., Kotlarski, S., San Martin, D., Herrera, S., Bedia, J., Casanueva, A., Manzanas, R., Iturbide, M., Vrac, M., Dubrovsky, Ribalaygua, J., Portoles, J., Raty, O., Raisanen, J., Hingray, B., Raynaud, D., Casado, M.J., Ramos, P., Zerenner, T., Turco, M., Bosshard, T., Stepanek, P., Bartholy, J., Pongracz, R.,

Keller, D.E., Fischer, A.M., Cardoso, R.M., Soares, P.M.M., Czernecki, B., Page, C.: An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment, Int. J. Climatol.,1-36, doi: 10.1002/joc.5462, 2018.

Guzman, R., Chepfer, H., Noel, V., Vaillant de Guelis, T., Kay, J.E., Raberanto, P., Cesana, G., Vaughan, M. A., and D. M. Winker: Direct atmosphere opacity observations from CALIPSO provide new constraints on cloud-radiation interactions, J. Geophys. Res. Atmos., 122(2), 1066-1085, doi: 10.1002/2016JD025946, 2017.

Hartmann, D. L., Moy, L. A., and Fu, Q.: Tropical convection and the energy balance at the top of the atmosphere, J. Climate, 14, 4495-4511, doi: 10.1175/1520-0442(2001)014<4495:TCATEB>2.0.CO;2 2001.

Hartmann, D. L., and Larson, K.: An important constraint on tropical cloud - climate feedback, Geophys. Res. Lett., 29(20), 1951, doi: 10.1029/2002GL015835, 2002.

Hastie, T., and Tibshirani, R.: Generalized additive models (with discussion), Statistical Science 1, 297-318, 1986.

Hastie, T., Tibshirani, R., and Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. Springer, 745 pp, 2009.

Haynes, J. M., L'Ecuyer, T. S., Stephens, G. L., Miller, S. D., Mitrescu, C., Wood, N. B., and Tanelli, S.: Rainfall retrieval over the ocean with spaceborne W-band radar, J. Geophys. Res., 114, D00A22, doi:10.1029/2008JD009973, 2009

Houze, R. A., and Betts, A. K.: Convection in GATE, Rev. Geophys. Space Phys., 19, 541-576, doi: 10.1029/RG019i004p00541, 1981.

Intrieri, J. M., Fairall, C. W., Shupe, M. D., Persson, P. O. G., Andreas, E. L., Guest, P. S., and Moritz, R. E.: An annual cycle of Arctic surface cloud forcing at SHEBA, J. Geophys. Res., 107(C10), 8039, doi: 10.1029/2000JC000439, 2002.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bull. Amer. Meteor. Soc., 77, 437–472, doi: 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

Kammann, E. E., and Wand, M. P.: Geoadditive models, Journal of the Royal Statistical Society: Series C (Applied Statistics), 52, 1-18, doi: 10.1111/1467-9876.00385, 2003.

Kato, S., Rose, F. G., Sun-Mack, S., Miller, W. F., Chen, Y., Rutan, D.A., Stephens, G.L, Loeb, N. G., Minnis, P., Wielicki, B. A., Winker, D. M., Charlock, T. P., Stackhouse Jr. P. W., Xu, K.-M., Collins, W. D.: Improvements of top-of-atmosphere and surface irradiance computations with CALIPSO-, CloudSat-, and MODIS-derived cloud and aerosol properties, J. Geophys. Res., 116, D19209, doi: 10.1029/2011JD016050, 2011.

Kay, J. E., L'Ecuyer, T., Gettelman, A., Stephens, G., and O'Dell, C.: The contribution of cloud and radiation anomalies to the 2007 Arctic sea ice extent minimum, Geophys. Res. Lett., 35, L08503, doi: 10.1029/2008GL033451, 2008.

Kay, J.E., Bourdages, L., Miller, N.B., Morrison, A., Yettella, V., Chepfer H, and Eaton, B.: Evaluating and improving cloud phase in the Community Atmosphere Model version 5 using spaceborne lidar observations, J. Geophys. Res.-Atmos., 121(8), 4162-4176, doi: 10.1002/2015JD024699, 2016.

Klein, S.A., Hall, A., Norris, J.R. et al.: Low-Cloud Feedbacks from Cloud-Controlling Factors: A Review, Surv. Geophys., 38, 1307-1329, doi: 10.1007/s10712-017-9433-3, 2017.

Lacour, A., Chepfer, H., Shupe, M.D., Miller, N.B., Noel, V., Kay, J., Turner, D.D., and Guzman, R.: Greenland Clouds Observed in CALIPSO-GOCCP: Comparison with Ground-Based Summit Observations, J. Clim., 30(15), 6065-6083, doi: 10.1175/JCLI-D-16-0552.1, 2017.

Lebsock, M. D., and L'Ecuyer, T. S.: The retrieval of warm rain from CloudSat, J. Geophys. Res., 116, D20209, doi: 10.1029/2011JD016076, 2011.

Lloyd, S. P.: Least squares quantization in PCM, IEEE Transactions on Information Theory, 28(2), 129-137, doi: 10.1109/TIT.1982.1056489, 1982.

5   Liu, D.S., Pu, R.L.: Downscaling thermal infrared radiance for subpixel land surface temperature retrieval, Sensors 8, 2695-2706, doi: 10.3390/s8042695, 2008.

Liu, Z., Vaughan, M., Winker, D., Kittaka, C., Getzewich, B., Kuehn, R., Omar, A., Powell, K., Trepte, C., and Hostetler, C.: The CALIPSO Lidar Cloud and Aerosol Discrimination: Version 2 Algorithm and Initial Assessment of Performance, J. Atmos. Oceanic Technol., 26, 1198–1213, doi: 10.1175/2009JTECHA1229.1, 2009.

10   Long, C. N., Dutton, E. G., Augustine, J. A., Wiscombe, W., Wild, M., McFarlane, S. A., and Flynn, C. J: Significant decadal brightening of downwelling shortwave in the continental United States, J. Geophys. Res., 114, D00D06, doi: 10.1029/2008JD011263, 2009.

Luo, Z., and Rossow, W.B.: Characterizing Tropical Cirrus Life Cycle, Evolution, and Interaction with Upper-Tropospheric Water Vapor Using Lagrangian Trajectory Analysis of Satellite Observations, J. Climate, 17, 4541–4563, doi: 10.1175/3222.1, 2004

Mace, G. G., Zhang, Q., Vaughan, M., Marchand, R., Stephens, G., Trepte, C., and Winker, D.: A description of hydrometeor layer occurrence
15   statistics derived from the first year of merged Cloudsat and CALIPSO data, J. Geophys. Res., 114, D00A26, doi: 10.1029/2007JD009755, 2009.

Malone, B. P., McBratney, A. B., Minasny, B., Wheeler, I., A general method for downscaling earth resource information, Computers & Geosciences, 41, 119-125, doi: 10.1016/j.cageo.2011.08.021, 2012.

Manara, V., Brunetti, M., Celozzi, A., Maugeri, M., Sanchez-Lorenzo, A., and Wild, M.: Detection of dimming/brightening in Italy
20   from homogenized all-sky and clear-sky surface solar radiation records and underlying causes (1959–2013), Atmos. Chem. Phys., 16, 11145-11161, doi: 10.5194/acp-16-11145-2016, 2016.

Martins E., Noel, V., Chepfer, H.: Properties of cirrus and subvisible cirrus from nighttime CALIOP, related to atmospheric dynamics and water vapour, J. Geophys. Res., J. Geophys. Res., 116, D02208, doi:10.1029/2010JD014519, 2011.

Meinshausen, N.: Quantile regression forests, J. Mach. Learn. Res., 7, 983-999, 2006.

25   Nam, C., Bony, S., Dufresne, J.-L., and Chepfer, H.: The 'too few, too bright' tropical low-cloud problem in CMIP5 models, Geophys. Res. Lett., 39, L21801, doi: 10.1029/2012GL053421, 2012.

Obligis, E., Rahmani, A., Eymard, L., Labroue, S., and Bronner, E.: An Improved Retrieval Algorithm for Water Vapor Retrieval: Application to the Envisat Microwave Radiometer, IEEE Transactions on Geoscience and Remote Sensing, 47(9), 3057-3064, 2009.

Palerme, C., Kay, J. E., Genthon, C., L'Ecuyer, T., Wood, N. B., and Claud, C.: How much snow falls on the Antarctic ice sheet?, The
30   Cryosphere, 8, 1577-1587, doi: 10.5194/tc-8-1577-2014, 2014.

Pierrehumbert, R. H.: Lateral mixing as a source of subtropical water vapour, Geophys. Res. Lett., 25(2), 0094-8276, doi: 10.1029/97GL03563, 2002.

Pierrehumbert, R. H.: The hydrologic cycle in deep time climate problems, Nature, 419, 191-198, doi: 10.1038/nature01088, 2002.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2017.

35   Randall, D., Khairoutdinov, M., Arakawa, A., and Grabowski, W.: Breaking the Cloud Parameterization Deadlock, Bull. Amer. Meteor. Soc., 84, 1547-1564, doi: 10.1175/BAMS-84-11-1547, 2003.

Raschke, E., Kinne, S., Stackhouse, P.W.: GEWEX Radiative Flux Assessment (RFA) Volume 1: Assessment. A Project of the World Climate Research Programme Global Energy and Water Cycle Experiment (GEWEX) Radiation Panel, WCRP Report 19/2012, World Meteorological Organization (WMO), Geneva, Switzerland, 2012.

Reverdy M., Noel, V., Chepfer, H., Legras, B.: On the origins of subvisible cirrus clouds in the tropical upper troposphere, Atm. Chem. Phys., 12, 12081-12101, doi:10.5194/acp-12-12081-2012, 2012.

Rue, H., and Held, L.: Gaussian Markov random fields, Theory and applications, Boca Raton: CRC=Chapman & Hall, 2005.

Sassen, K., Wang, Z., and Liu, D.: Global distribution of cirrus clouds from CloudSat/Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) measurements, J. Geophys. Res., 113, D00A12, doi: 10.1029/2008JD009972, 2008.

Sekiyama, T. T., Tanaka, T. Y., Shimizu, A., and Miyoshi, T.: Data assimilation of CALIPSO aerosol observations, Atmos. Chem. Phys., 10, 39-49, doi: 10.5194/acp-10-39-2010, 2010.

Sivira, R., Brogniez, H., Mallet, C., and Oussar, Y.: A layer-averaged relative humidity profile retrieval for microwave observations: design and results for the Megha-Tropiques payload, Atmospheric Measurement Techniques, 8, 1055-107, doi: 10.5194/amt-8-1055-2015, 2015.

Schröder, M., Lockhoff, M., Shi, L., August, T., Bennartz, R., Borbas, E., Brogniez, H., Calbet, X., Crewell, S., Eikenberg, S., Fell, F., Forsythe, J., Gambacorta, A., Graw, K., Ho, S.P., Höschen, H., Kinzel, J., Kursinski, E.R., Reale, A., Roman, J., Scott, N., Steinke, S., Sun, B., Trent, T., Walther, A., Willen, U., and Yang, Q.: GEWEX water vapor assessment (G-VAP). WCRP Report 16/2017 World Climate Research Programme (WCRP): Geneva, Switzerland 2017, 216, pp, available at urlhttps://www.wcrp-climate.org/resources/wcrp-publications, 2017.

Shupe, M.D., Matrosov, S.Y., and Uttal, T.: Arctic Mixed-Phase Cloud Properties Derived from Surface-Based Sensors at SHEBA, J. Atmos. Sci., 63, 697–711, doi: 10.1175/JAS3659.1, 2006.

Stephens, G. L., Tsay, S. C., Stackhouse, P. W., and Flatau, P. J.: The relevance of the microphysical and radiative properties of cirrus clouds to climate and climatic feedback, J. Atmos. Sci., 47, 1742-1754, 1990.

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O'connor, E. J., Rossow, W. B., Durden, S. L., Miller, S. D., Austin, R. T., Benedetti, A., and Mitrescu, C.: The CloudSat mission and the A-TRAIN, Bull. Amer. Meteor. Soc., 83, 1771-1790, doi: 10.1175/BAMS-83-12-1771, 2002.

Stephens, G. L., and Coauthors: CloudSat mission: Performance and early science after the first year of operation, J. Geophys. Res., 113, D00A18, doi: 10.1029/2008JD009982, 2008.

Stephens, G.L., Wild, M., Stackhouse, P.W., L'Ecuyer, T., Kato, S., and Henderson, D.S.: The Global Character of the Flux of Downward Longwave Radiation, J. Climate, 25, 2329-2340, doi: 10.1175/JCLI-D-11-00262.1, 2012.

Stephens, G., Winker, D., Pelon, J., Trepte, C., Vane, D., Yuhas, C., L'Ecuyer, T., and Lebsock, M.: CloudSat and CALIPSO within the A-Train: Ten Years of Actively Observing the Earth System, Bull. Amer. Meteor. Soc., 99, 569–581, doi: 10.1175/BAMS-D-16-0324.1, 2018.

Stubenrauch, C.J., Rossow, W.B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H., Di Girolamo, L., Getzewich, B., Guignard, A., Heidinger, A., Maddux, B.C., Menzel, W.P., Minnis, P., Pearl, C., Platnick, S., Poulsen, C., Riedi, J., Sun-Mack, S., Walther, A., Winker, D., Zeng, S., and Zhao, G.: Assessment of Global Cloud Datasets from Satellites: Project and Database Initiated by the GEWEX Radiation Panel, Bull. Amer. Meteor. Soc., 94, 1031–1049, doi: 10.1175/BAMS-D-12-00117.1, 2013.

Taillardat, M., Mestre, O., Zamo, M., and Naveau, P.: Calibrated Ensemble Forecasts Using Quantile Regression Forests and Ensemble Model Output Statistics, Mon. Wea. Rev., 144, 2375-2393, doi: 10.1175/MWR-D-15-0260.1, 2016.

Tian, B., Soden, B. J., and Wu, X.: Diurnal cycle of convection, clouds, and water vapor in the tropical upper troposphere: Satellites versus a general circulation model, J. Geophys. Res., 109, D10101, doi: 10.1029/2003JD004117, 2004.

Udelhofen, P. M., and Hartmann, D. L.: Influence of tropical cloud systems on the relative humidity in the upper troposphere, J. Geophys. Res., 100(D4), 7423–7440, doi: 10.1029/94JD02826, 1995.

5   Vaillant de Guélis T., H. Chepfer, V. Noel, R. Guzman, D. Winker, R. Plougonven: Using space lidar observations to decompose Longwave Cloud Radiative Effect variations over the last decade, Geophys. Res. Let., 44, 11, 994-12, 003, doi: 10.1002/2017GL074628, 2017.

Vaittinada Ayar, P., Vrac, M., Bastin, S., Carreau, J., Déqué, M., Gallardo, C.: Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: Present climate evaluations, Clim. Dyn., 46, 1301-1329, doi: 10.1007/s00382-015-2647-5, 2015.

10   Vaughan, M.A., Powell, K.A., Winker, D.M., Hostetler, C.A., Kuehn, R.E., Hunt, W.H., Getzewich, B.J., Young, S.A., Liu, Z., and McGill, M.J.: Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements, J. Atmos. Oceanic Technol., 26, 2034–2050, doi: 10.1175/2009JTECHA1228.1, 2009.

Vial, J., Bony, S., Dufresne, J., Roehrig, R.: Coupling between lower-tropospheric convective mixing and low-level clouds: Physical mechanisms and dependence on convection scheme, J. Adv. Model Earth System, 8(4), 1892-1911, doi: 10.1002/2016MS000740, 2016.

15   Wielicki, B. A., Harrison, E. F., Cess, R. D., King, M. D., and Randall, D. A.: Mission to Planet Earth: Role of clouds and radiation in climate, Bull. Amer. Meteor. Soc., 76, 2125-2153, doi: 10.1175/1520-0477(1995)076<2125:MTPERO>2.0.CO;2, 1995.

Wild, M.: Global dimming and brightening: A review, J. Geophys. Res., 114, D00D16, doi: 10.1029/2008JD011470, 2009.

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K.A., Liu, Z., Hunt, W. H., and Young, S. A.: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms, J. Atmos. Oceanic Technol., 26, 2310-2323, doi: 10.1175/2009JTECHA1281.1, 2009.

20   Winker, D.M., Pelon, J., Coakley, J.A., Ackerman, S.A., Charlson, R.J., Colarco, P.R., Flamant, P., Fu, Q., Hoff, R.M., Kittaka, C., Kubar, T.L., Le Treut, H., Mccormick, M.P., Mégie, G., Poole, L., Powell, K., Trepte, C., Vaughan, M.A., and Wielicki, B.A.: The CALIPSO Mission, Bull. Amer. Meteor. Soc., 91, 1211–1230, doi: 10.1175/2010BAMS3009.1, 2010.

Winker, D., Chepfer, H., Noel, V. and Cai, X.: Observational Constraints on Cloud Feedbacks: The Role of Active Satellite Sensors, Surv Geophys., 38(6), 1483-1508, (2017) 38: 1483. https://doi.org/10.1007/s10712-017-9452-0, 2017.

25   Wood, S.N.: Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models, Journal of the Royal Statistical Society (B), 73(1), 3-36, doi: 10.1111/j.1467-9868.2010.00749.x, 2011.

Xu, K., Wong, T., Wielicki, B.A., Parker, L., Lin, B., Eitzen, Z.A., and Branson, M.: Statistical Analyses of Satellite Cloud Object Data from CERES. Part II: Tropical Convective Cloud Objects during 1998 El Niño and Evidence for Supporting the Fixed Anvil Temperature Hypothesis. J. Climate, 20, 819–842, doi: 10.1175/JCLI4069.1, 2007.

30   Zhang, M.H., Lin, W.Y., Klein, S.A., Bacmeister, J.T., Bony, S., Cederwall, R.T., Del Genio, A.D., Hack, J.J., Loeb, N.G., Lohmann, U., Minnis, P., Musat, I., Pincus, R., Stier, P., Suarez, M.J., Webb, M.J., Wu, J.B., Xie, S.C., Yao, M.-S., and Zhang, J.H.: Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements, J. Geophys. Res., 110, D15S02, doi: 10.1029/2004JD005021, 2005.

Zelinka, M. D., and Hartmann, D. L.: Why is longwave cloud feedback positive? J. Geophys. Res., 115, D16117, doi:
35   10.1029/2010JD013817, 2010.

Zorita, E. and von Storch, H.: The analog method as a simple statistical downscaling technique: Comparison with more complicated methods, J. Climate, 12, 2474-2489, 1999.
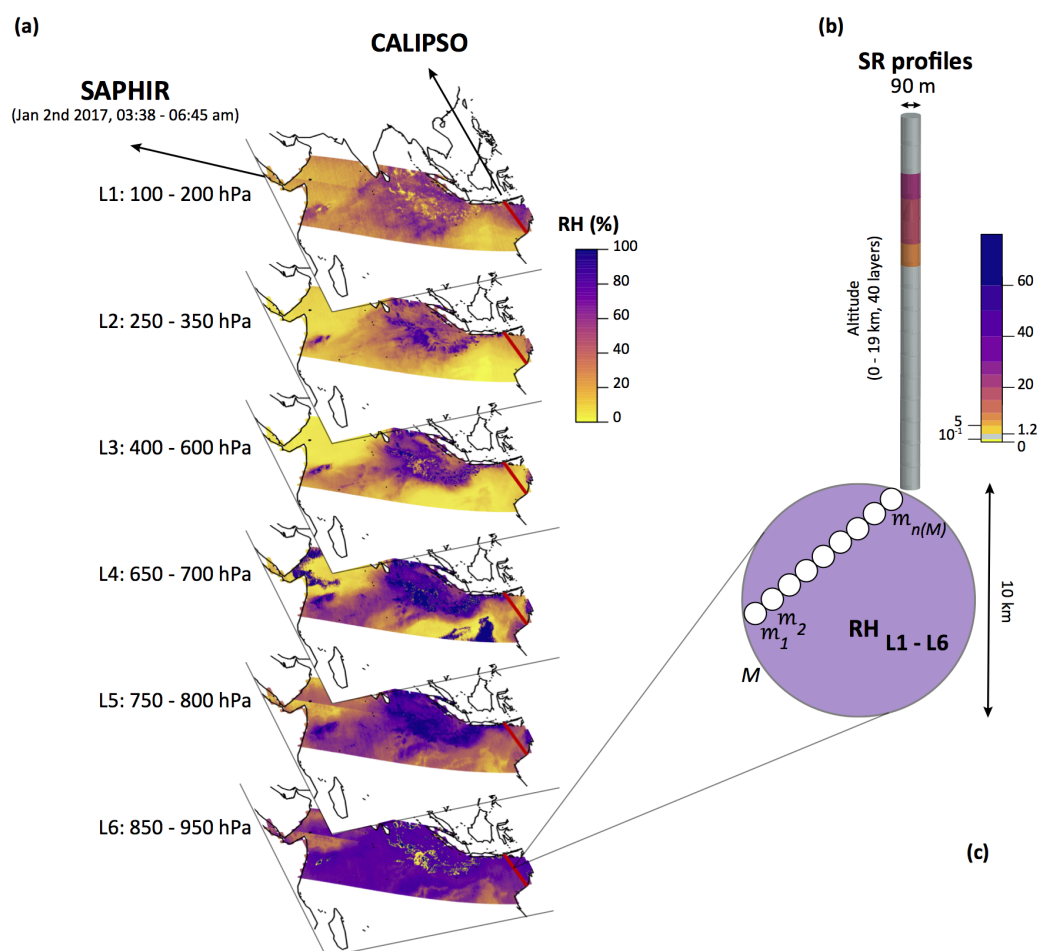
Figure 1. (a): RH (mean) observed by SAPHIR for all six pressure layers, in the Indian Ocean on January 2nd 2017 between 03:38 and 06:45 am. Overlaid is the CALIPSO track line (*red line*). (b): example of SR profile measured by CALIPSO. (c): schematic representation of SAPHIR-CALIPSO co-location: $M = 1, ..., N$ SAPHIR measurements at coarse resolution encapsulating $m = 1, ..., n(M)$ finely-resolved CALIPSO observations.
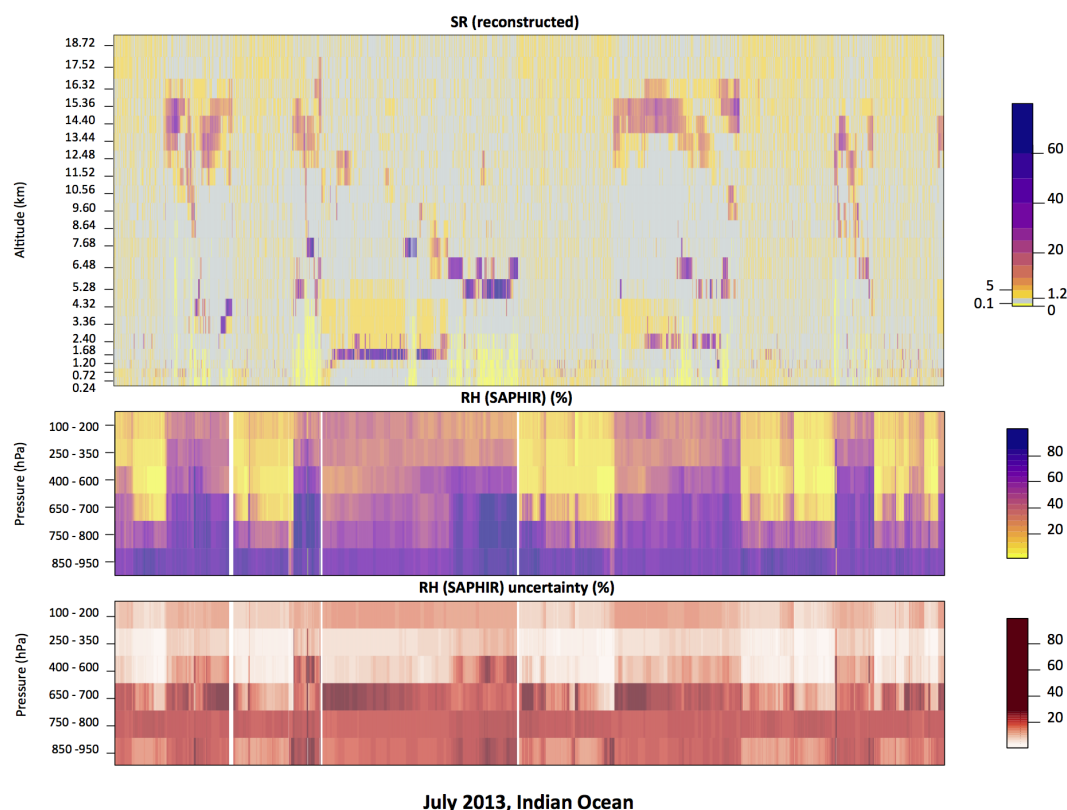
Figure 2. Reconstructed SR profiles for a selection of CALIPSO samples in the Indian Ocean, July 2013 (*top*) and co-located RH observations from SAPHIR (mean and uncertainty (standard deviation), *middle and bottom*). As in Chepfer et al. (2010), $SR > 5$ correspond to cloudy observations, $0 < SR < 0.01$ (*light yellow*) correspond to fully attenuated observations, and $1.2 < SR < 5$ (*grey*) correspond to unclassified observations. Note that the reconstructed SR were only used for layers indicating clouds to avoid mixing of cloud and clear sky values. The x-axis represents the co-location index.
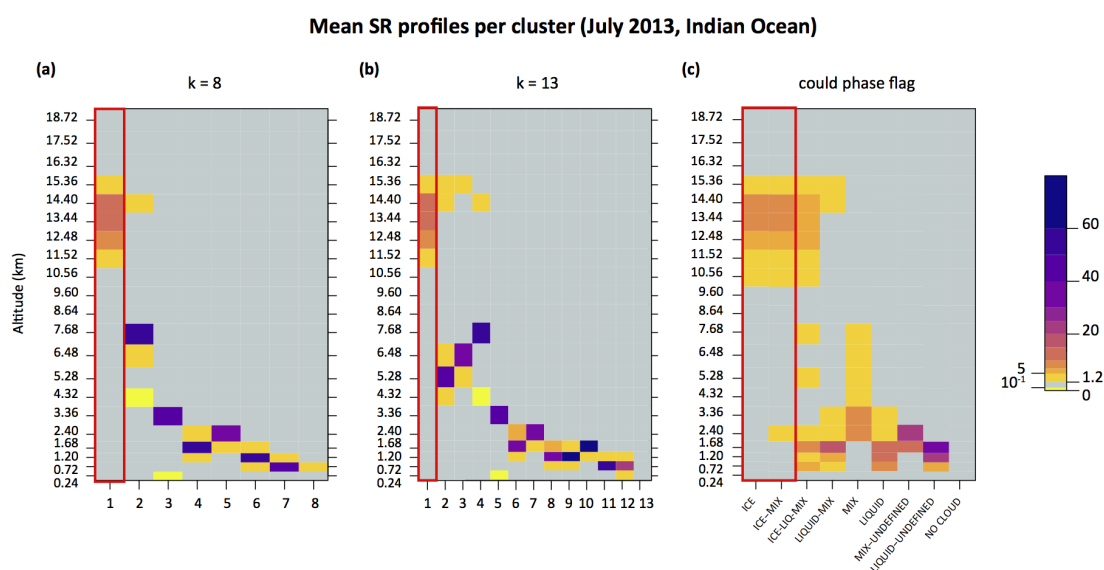
Figure 3. Mean SR profile per cluster for different choices of the clustering method (Indian Ocean, July 2013). (a): Mean SR profile per cluster obtained by a k-means classification setting k=8. (b): as (a) but setting k=13. (c): Mean SR profiles per cluster derived by combining the cloud phase flags in Cesana and Chepfer (2013). ICE: observations classified as ice only. LIQUID: observations classified as liquid only. MIX: profiles containing SR values derived by averaging observations classified as liquid and observations classified as ice. UNDEFINED: observations for which the cloud phase flag in Chepfer et al. (2010) is 'undefined', 'horizontally oriented' or 'unphysical'. The cluster type is then defined as the combination of these flags. Profiles characterized by other combinations of flags (e.g. FALSE LIQUID, FALSE ICE, etc.) correspond to less than 250 observations and have been omitted. Selected anvil-type clusters are outlined by a *red square*.
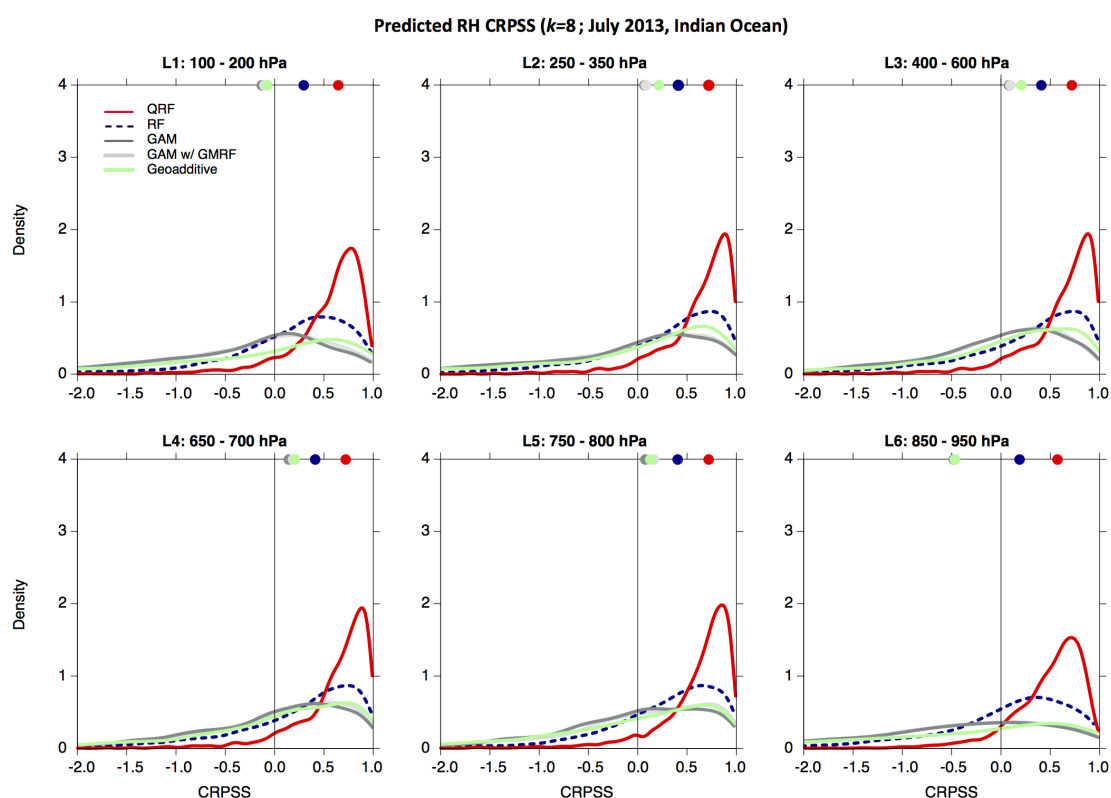
Figure 4. CRPSS score for ice cloud profiles (k=8) in the Indian Ocean, July 2013: QRF (*red solid line*), RF (*blue dashed line*), GAM (*dark grey solid line*), GAM with GMRF smoother (*light grey solid line*) and with the geoadditive method (*green solid line*). The dots at the top of each panel indicate the median of the distribution. Predictions are from the validation set within a 5-fold cross validation scheme.
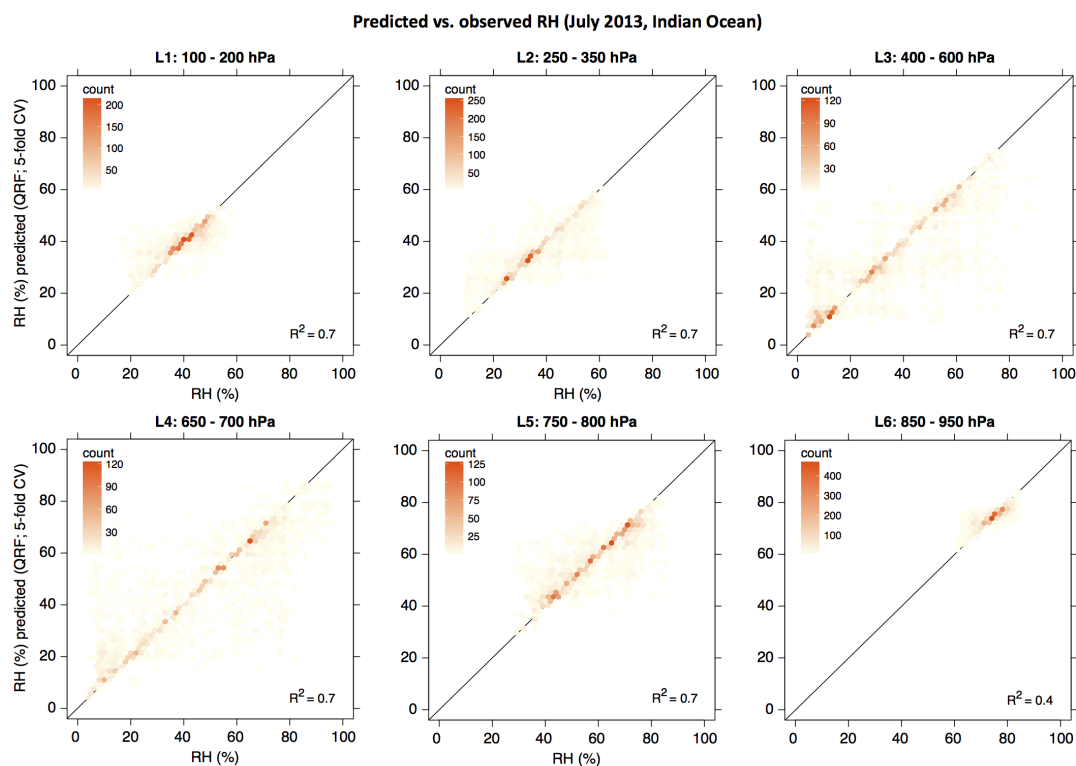
Figure 5. Scatter plot of the median of the predicted distribution vs. observed RH for ice cloud profiles (k=8) in the Indian Ocean, July 2013. Predictions are made using the QRF method and are from the validation set within a 5-fold cross validation scheme. R^2 is computed as $1 - \frac{\sum_i (y_i - \widehat{y_i})}{\sum_i (y_i - \overline{y})^2}$ where the y_i represent SAPHIR observations with mean $\overline{y}$ and $\widehat{y_i}$ are the cross-validation predictions.
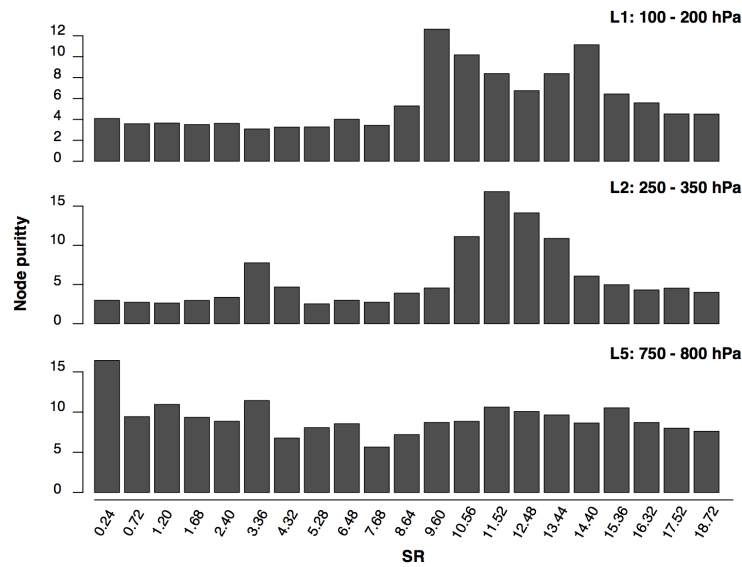
Figure 6. Variable importance (QRF method) for the predicted RH for ice cloud profiles (k=8) in the Indian Ocean, July 2013.
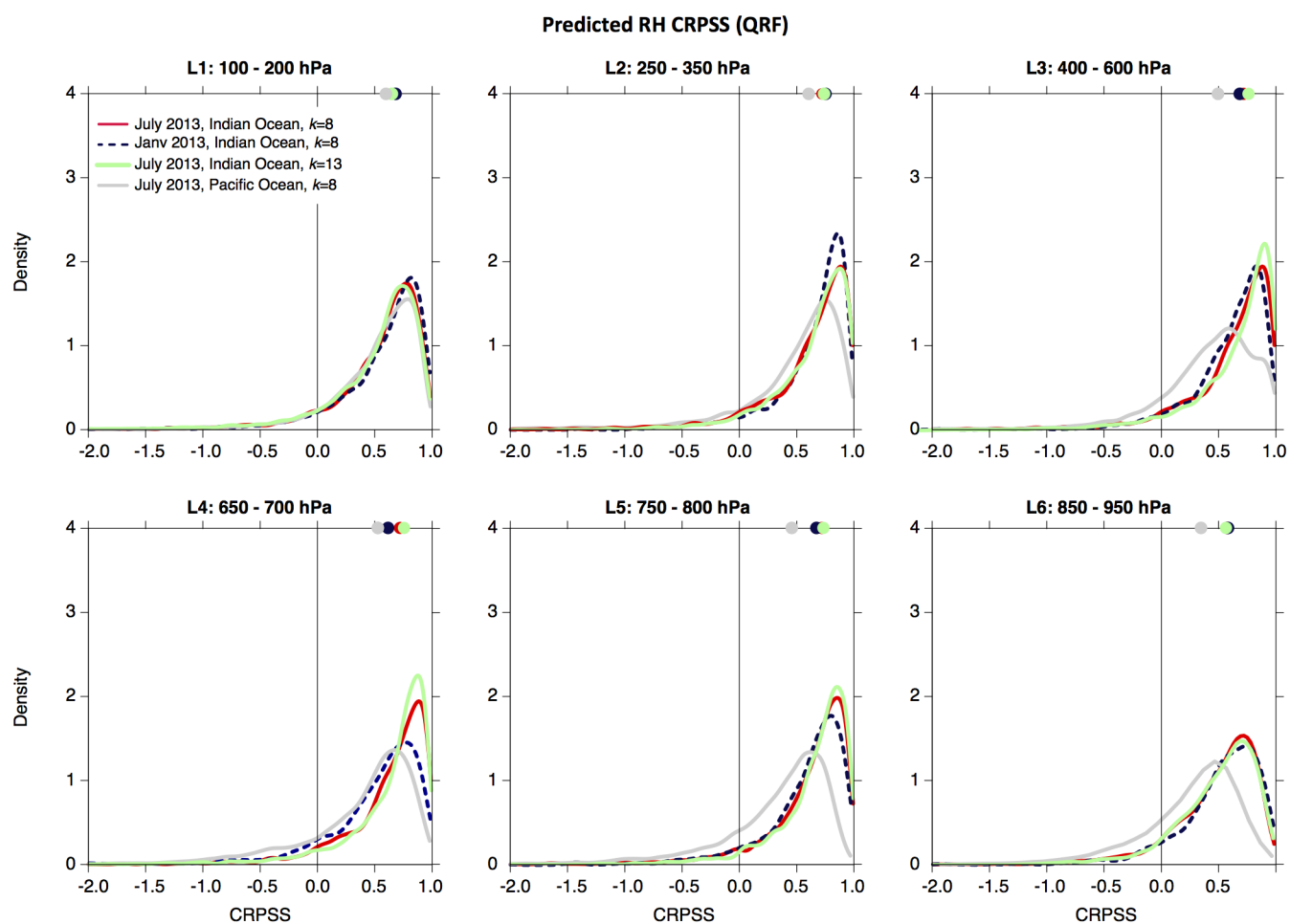
Figure 7. CRPSS score for ice cloud profiles (QRF method): Indian Ocean, July 2013 setting $k=8$ (*red solid line*) and setting $k=13$ (*blue dashed line*); Indian Ocean, January 2013 setting $k=8$ (*dark grey solid line*); Pacific Ocean, July 2013 setting $k=8$ (*light grey solid line*). The dots at the top of each panel indicate the median of the distribution.
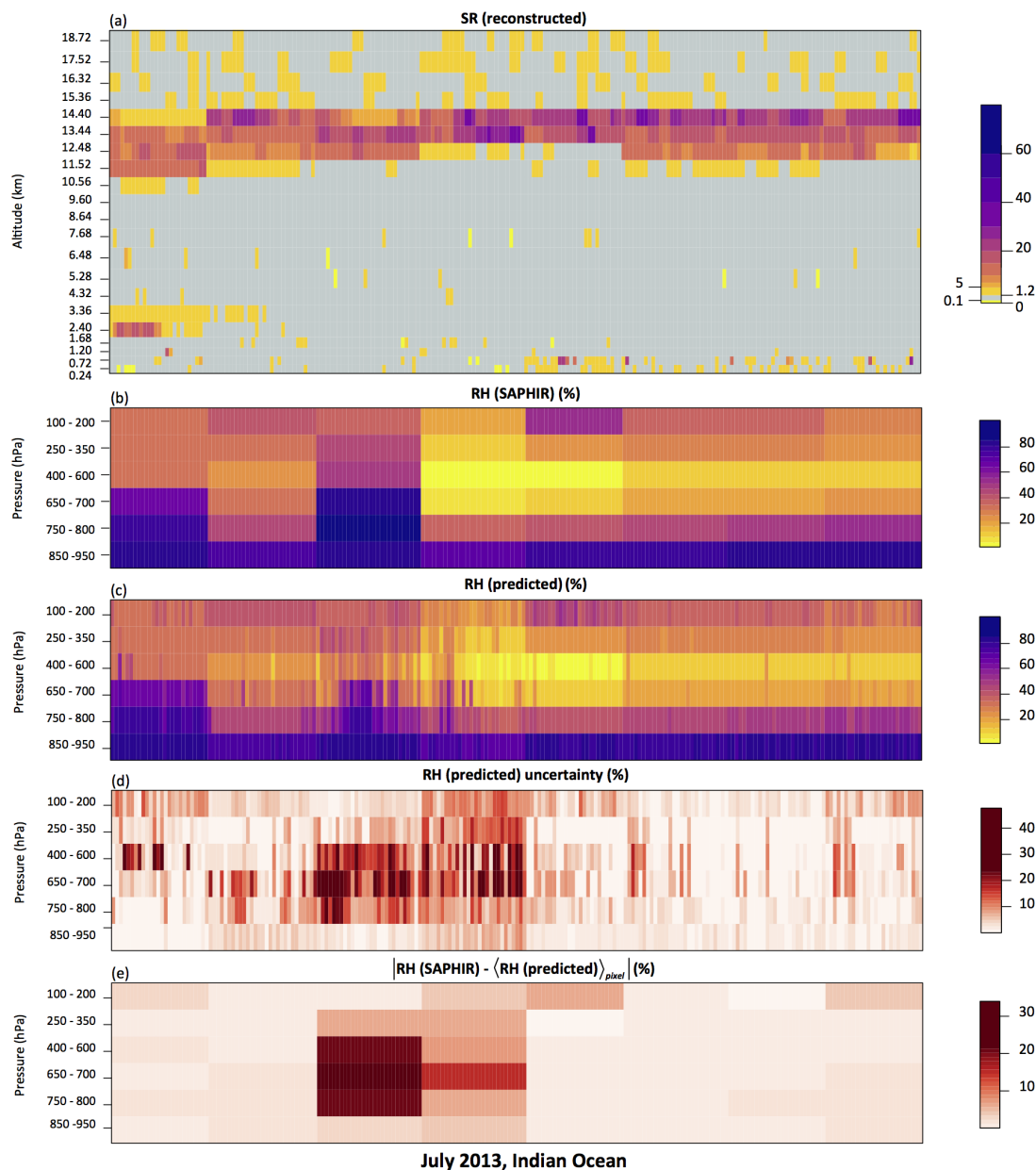
Figure 8. (*a*): SR profiles for a selection of ice cloud profiles from CALIPSO in the Indian Ocean, July 2013. The selected cloud profiles correspond to SAPHIR pixels with $n(M) > 25$. The scale is the same as in Fig. 2. (*b*): Co-located layered-RH observations from SAPHIR (mean). (*c*): Predicted layered-RH using the QRF method within the iterative scheme (median). (*d*): as (*c*) but for the interquartile range instead of the median. (*e*): for each layer, absolute difference between the observed RH from SAPHIR and the average over each SAPHIR pixel of the predicted RH. The x-axis represents the co-location index.
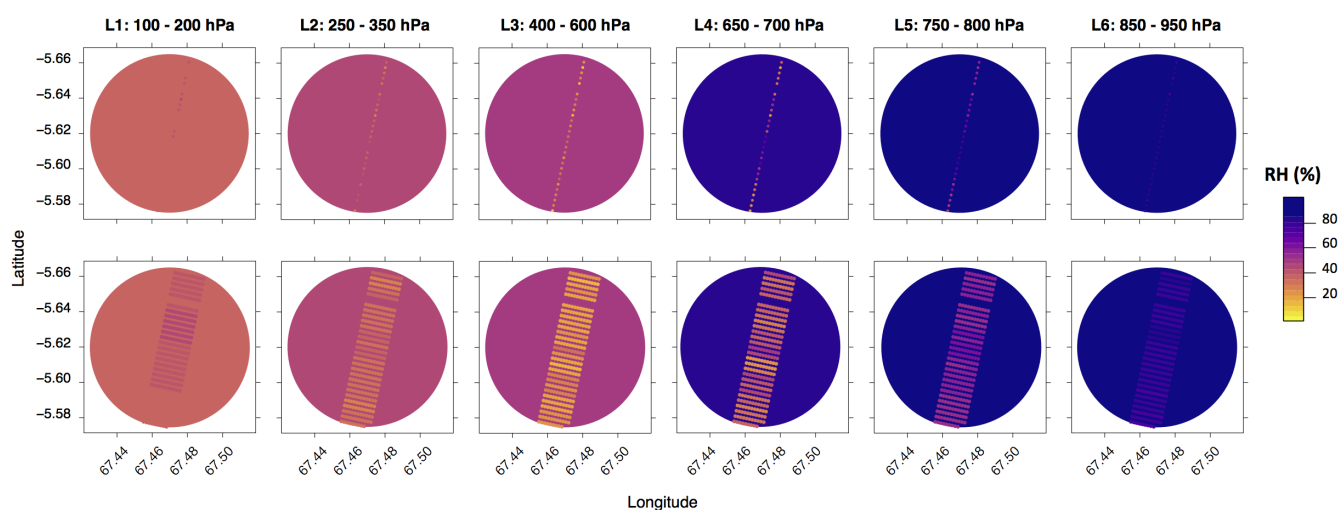
Figure 9. Example of predicted RH for a single SAPHIR pixel corresponding to ice cloud profiles using the QRF method (*top*, median) and the geoadditive model (*bottom*) within the iterative scheme.

Model	Model type	Spatial Correlation	Prediction type
RF	Non-parametric	-	Conditional mean
QRF	Non-parametric	-	Conditional quantiles
GAM	Semi-parametric	-	Conditional mean
GAM with GMRF smoother	Semi-parametric	Neighbour structure	Conditional mean
Geoadditive	Semi-parametric	Exponential correlation function	Conditional mean

Table 1. Summary of the regression models tested in this study.