Modeling High Resolution 3-D Cloud Fields for Earth-space Communication Systems

Lorenzo Luini and Carlo Capsoni

***Abstract*—A methodology to synthesize three-dimensional spatially correlated cloud fields from Numerical Weather Prediction (NWP) products is presented. The target area is 200 km×200 km and the horizontal spatial resolution is 1 km×1 km. The field synthesis relies on the stochastic approach proposed by Bell and the main model's parameters are extracted from high-resolution cloud fields observed by the MODIS sensor. The model's inputs are the fractional cloud cover and the average integrated cloud liquid water content provided by an NWP dataset (the ERA40 reanalysis in this study). Also the vertical profile of clouds is modelled, based on the analysis of data collected by the Cloud Profiling Radar on-board the CloudSat satellite. Tests on the model performance indicate that both first-order (Complementary Cumulative Distribution Function - CCDF) and second-order (spatial distribution) statistics of the integrated cloud liquid water content are reproduced with good accuracy in several sites in Europe. The proposed model is one of the main blocks of a simulator of weather disturbances affecting radio wave propagation, primarily intended to support the design and performance assessment of Earth-space Communication Systems (EHF range or optical wavelengths) but also of possible interest for all the applications involving radiative transfer in the atmosphere.**

***Index Terms*— Electromagnetic propagation, cloud effects, Fade Mitigation Techniques, radiative transfer.**

I. Introduction

EHF carriers are nowadays becoming very attractive to satellite communication (SatCom) system operators because they offer wide bandwidth for the provision of advanced multimedia and interactive services. Above 10 GHz, the atmosphere has a definite impact on Earth-space links and, while rainfall always represents the prevalent impairment affecting radio waves [1], the contribution of suspended liquid water becomes significant at frequencies above 20 GHz and in low elevation links, not only in terms of specific attenuation, but also for the high occurrence probability of clouds (40–80% of the yearly time in Europe). Even more, at optical wavelengths, which, in principle, would enable Earth-space communication systems with extremely high data rates, the presence of clouds along the path is the limiting factor because of the large density and marked optical extinction properties of micrometric droplets [2].

In the field of wave propagation, which this contribution addresses, the prediction of cloud effects on satellite links is tackled by few models of different complexity and applicability. A class of semiempirical models addressing the EHF range, such as those proposed by Altshuler and Marr [3] and by Dintelmann and Ortgies [4], relate cloud attenuation A_C to different meteorological quantities (e.g. the surface absolute humidity) by defining expressions whose coefficients have been regressed on existing measurements. Other ones are more physically sound since they preliminarily introduce a cloud model to evaluate cloud attenuation. Among them, it is worth mentioning the methodology proposed by Dissanayake *et al.* in [5], which predicts cloud attenuation statistics based on the classification of clouds into four classes with associated key average properties (vertical and horizontal extent, water content) and probability of occurrence. The model developed by Salonen and Uppala, henceforth referred to as TKK (Teknillinen KorkeaKoulu) model [6], has received great attention (it is currently adopted in recommendation ITU-R P.840-6 [7]) because of its physical basis (it relies on the identification of cloud presence from vertical profiles of pressure, relative humidity and temperature (*PHT*), in turn derivable from radiosonde observations (RAOBS), and then it calculates specific attenuation in each layer according to the Rayleigh approximation [8]).

Cloud models for attenuation prediction currently available in the literature are limited in their applicability because intrinsically mono dimensional. In fact, they typically provide the vertical profile of the cloud, while its horizontal distribution is assumed to be uniform: as a result, models of this kind are restricted in estimating the impact of clouds on complex SatCom systems implementing site diversity schemes for the mitigation of high fades [9] or on Low Earth Orbit (LEO) satellite applications where the ground antenna changes elevation (from very low to high) and azimuth angles in some few minutes. Indeed, in these scenarios, the spatial correlation of clouds plays a relevant role. So far, this aspect has been addressed only in some works, such as [10], where the spatial distribution of clouds has been studied on continental scale (Europe, North and South America), although using meteorological data with coarse spatial resolution (2.5°×2.5° latitude/longitude grid). The correlation of clouds in space has been also duly investigated in [11] using 5 years of cloud cover data collected every 6 hours in 33 sites across Spain.

Manuscript received XXXX.
Lorenzo Luini and Carlo Capsoni are with the Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci, 35, 20133, Milano, Italy (e-mail: lorenzo.luini@polimi.it).

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/TAP.2014.2341297

This contribution presents SMOC (Stochastic Model Of Clouds), a methodology to synthesize high-resolution three-dimensional (3-D) cloud fields. The philosophy underlying the model is first explained in Section II; the discussion continues in Section III with the description of how the main parameters for the development of SMOC have been extracted from a set of cloud fields observed by the MODerate-resolution Imaging Spectroradiometer (MODIS) on-board the LEO Aqua satellite. Section IV details the full procedure to synthesize spatially correlated fields of integrated cloud liquid water content L from Numerical Weather Prediction (NWP) data. Section V focuses on how the vertical development of clouds is modeled starting from the observation of several vertical profiles of cloud liquid water content collected by the Cloud Profiling Radar (CPR) (CloudSat satellite). In Section VI the ability of SMOC in synthesizing realistic cloud fields is assessed, whilst Section VII finally draws some conclusions.

II. Rationale of the Cloud Model

SMOC is a methodology to synthesize a statistically meaningful dataset of high-resolution 3-D cloud fields. The model takes advantage of the stochastic approach proposed by Bell in [12], originally devised for rain field synthesis, which, as an intermediate step, generates random spatially correlated Gaussian fields. In this application, we synthesize spatially correlated cloud fields using the a priori knowledge of the fractional cloud cover and the average integrated cloud liquid water over the "target area", typically derivable from global meteorological products (e.g. NWP from the European Centre for Medium-range Weather Forecast, ECMWF [13]).

The main idea underpinning SMOC is that, by introducing suitable statistical properties of L (first-order statistics and spatial distribution) derived from high-resolution real cloud fields, it is possible to de-integrate average cloud quantities regularly provided worldwide over a coarse latitude/longitude grid with long sampling time (NWP), and, in practice, to synthesize realistic maps of integrated liquid water content, L, with fine spatial resolution (1 km×1 km) over areas in the order of 200 km×200 km (the "target area"). SMOC external input data and internal parameters are:

- E_L, the mean value of the cloud liquid water content over the target area, including $L = 0$ mm;
- S_L, the standard deviation of the cloud liquid water content over the target area, including $L = 0$ mm;
- f, the fractional cloud cover over the target area;
- $\rho(x,y)$, the spatial correlation of L.

E_L and f are extracted from NWP datasets, which typically provide such meteorological information worldwide (uniform latitude/longitude grid); on the other hand, S_L and $\rho(x,y)$ come from the processing of MODIS cloud maps as these quantities are not provided by NWP. The third dimension is introduced through an analytical profile for the vertical distribution of the liquid water content, with parameters directly dependent on L, as inferred from CloudSat observations.

III. Horizontal distribution of clouds: main features

A. The high-resolution cloud field database

The database used for the investigation of the horizontal distribution of clouds originates from the remote sensing observations collected by MODIS, which travels a 705-km high, sun-synchronous, near-polar orbit, thus achieving the full coverage of the Globe in less than two days. The MODIS sensor is a scientific payload mainly designed to provide measurements on large-scale global dynamics including changes in Earth's cloud cover, radiation budget and processes occurring over the oceans, land, and in the lower atmosphere. Radiance data are acquired by 36 optical channels (wavelength in the 0.4-14.4 μm range) with high spatial resolution (from 250 m to 1 km footprint, linear size) implementing automatic in-flight calibration procedures [14]. Raw data are first processed by the MODIS Characterization Support Team (MCST) to provide high quality calibrated products to the MODIS Science Team (MST) for diversified Earth science applications [15].

Among the various atmospheric high-resolution products made freely available by the National Aeronautics and Space Administration (NASA) there are maps of integrated liquid water content L whose dimensions are 200 km×2000 km and whose spatial resolution is 1 km×1 km, definitely suitable to adequately catch the spatial variability of L within clouds. In particular, we have downloaded 3090 swaths collected over Europe (20° N ≤ latitude ≤ 62° N and −10° E ≤ longitude ≤ 37° E) in 2010 using the Mirador web interface [16]. As an example, Fig. 1 shows the spatial distribution of the integrated liquid water content as observed by MODIS along one swath.

Fig. 1. Example of the spatial distribution of integrated cloud liquid water, L, as observed by MODIS along a swath over Europe.

B. Data processing and characterization of clouds spatial distribution

As a preliminary step, original swaths were subdivided into 10 maps as wide as 200 km×200 km in order to deal with dimensions typical of NWP products (e.g. 2°×2° latitude/longitude grid).

The resulting 13183 maps containing clouds (the full dataset with 30900 maps includes also cloud-free images) were processed to identify possible relationships between S_L and f or E_L because NWP products do not provide S_L, whose value, as anticipated in Section II, is a necessary element of SMOC for cloud field synthesis. The second couple of variables turned out to be the most appropriate one and the conditional probability density function $p(S_L|E_L)$ was found to be well approximated by the lognormal function, whose expression is:

$$p(S_L \mid E_L) = \frac{1}{S_L\,\sigma_p\sqrt{2\pi}}\exp\left[-\frac{\left(\ln S_L - \mu_p\right)^2}{2\sigma_p^2}\right] \quad (1)$$

where μ_p and σ_p, which are both function of E_L, are the mean and standard deviation values of the natural logarithm of S_L, respectively.

As an example, Fig. 2 depicts $p(S_L|E_L)$ for two classes of E_L (low values on the top and high values on the bottom), including μ_p and σ_p of the fitting MLE (Maximum Likely Estimation) lognormal distribution.

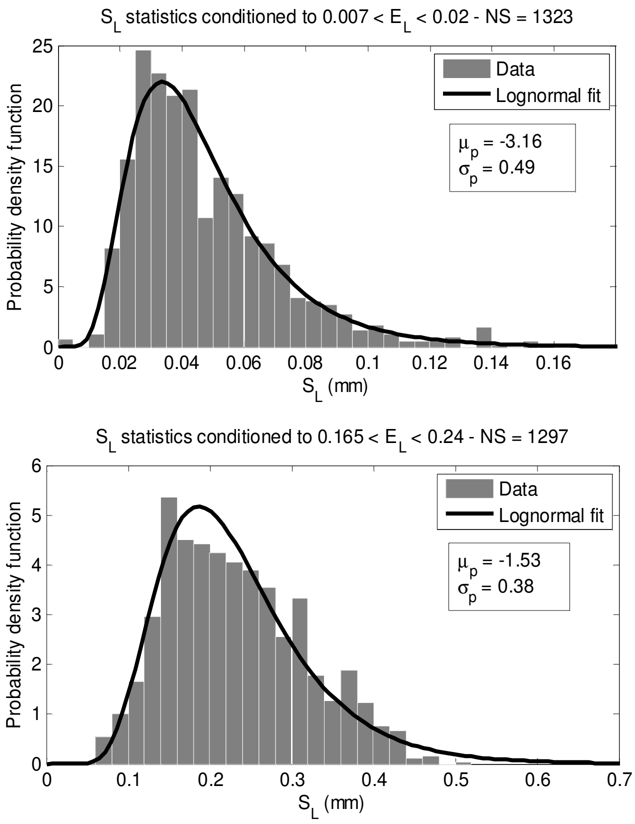

Fig. 2. Examples of $p(S_L|E_L)$, the statistical distribution of S_L conditioned to E_L; low and high values of E_L on the top and bottom side, respectively. Empirical data and MLE lognormal distributions.

In order to fully characterize $p(S_L|E_L)$, being E_L exponentially distributed as shown in Fig. 3, eleven E_L bins of different width have been defined in such a way to include in each of them approximately the same number of samples ($NS \approx 1300$). For each class, the MLE lognormal distributions fitted to the empirical $p(S_L|E_L)$ show root mean square (RMS) values of the percentage relative difference error that never exceed 10%. Moreover, as it is clear from Fig. 4 and Fig. 5, both μ_p and σ_p reveal quite a regular trend with E_L (squared dots represent the center values of each class), and, thus, can be properly fitted by the following simple expressions:

$$\begin{aligned}\mu_p(E_L) &= -5.61\,E_L^{-0.076} + 4.69\\ \sigma_p(E_L) &= 1.03\,E_L^{-0.029} - 0.81\end{aligned} \quad (2)$$

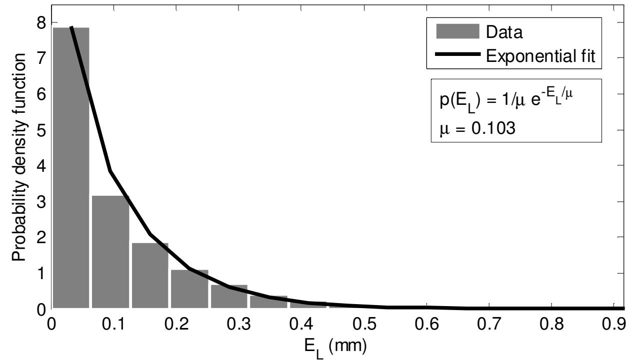

Fig. 3. Distribution of E_L derived from MODIS data; empirical data and MLE exponential distribution.

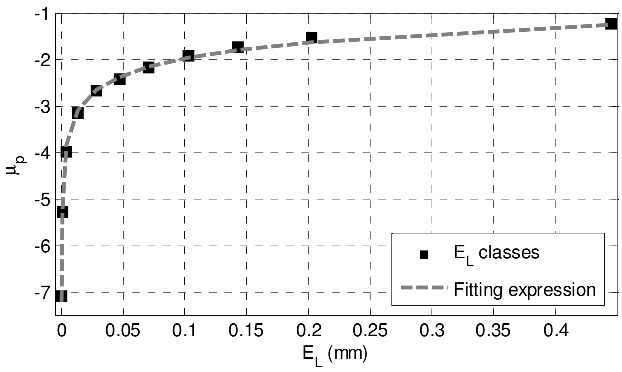

Fig. 4. Trend of μ_p with E_L.

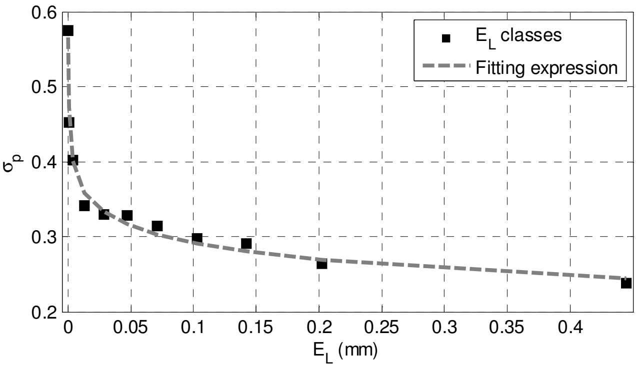

Fig. 5. Trend σ_p with E_L.

In turn, the distribution of L within each 200 km×200 km map (conditioned to $L > 0$ mm) was found to be lognormal as shown in Fig. 6, where a sample cloud field observed by MODIS (top) and the statistical characterization of L (bottom) is provided in terms of Cumulative Distribution Function (CDF). In addition to E_L, S_L, and f, the figure title also includes μ_{LN} and σ_{LN} of the MLE lognormal distribution fitting L data with evident satisfactory accuracy.

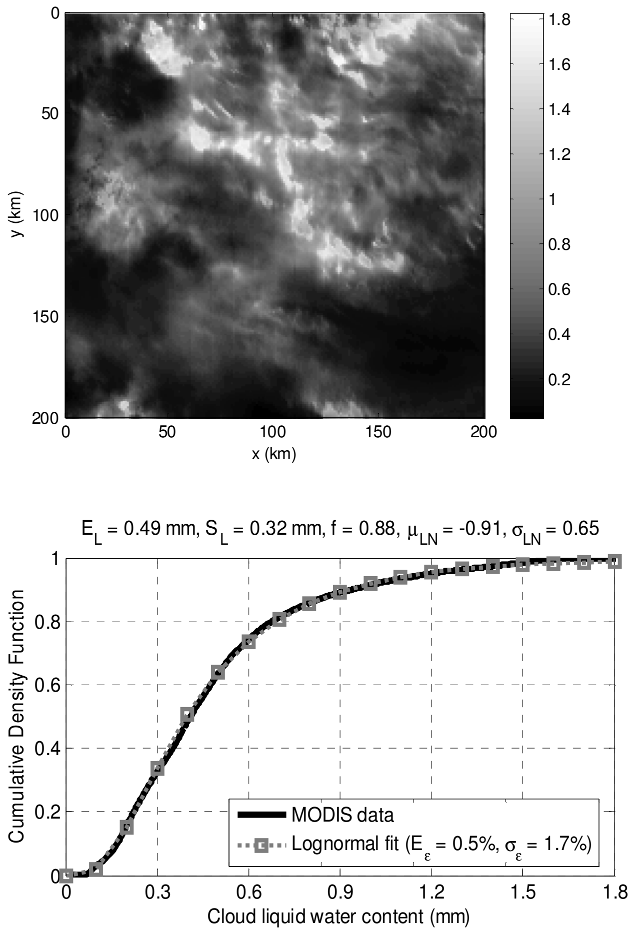

Fig. 6. MODIS cloud field example (top) and statistical characterization of L (bottom).

The appropriateness of the lognormal approximation for L is quantified in the figure legend which also reports the average (E_ε) and standard deviation (σ_ε) values of the relative error figure ε defined as:

$$\varepsilon(P) = 100\frac{L_f(P) - L_m(P)}{L_m(P)} \tag{3}$$

In (3), $L_m(P)$ and $L_f(P)$ are the L values (mm) extracted from the reference (MODIS) and fitted CDFs, respectively, at given probability levels P covering the full 0-1 range, with step of 0.001.

Fig. 7 shows the trend of the average E_ε (solid line) and σ_ε (dashed line) as a function of the fractional cloud cover f, together with the percentage number of MODIS cloud fields considered in each class (gray bars). Besides showing that f tends to be rather uniformly distributed between 0 and 1, except for the prevalence of fields with full cloud coverage, results confirm that L in each map tends to be lognormally distributed, being the approximation slightly more accurate for larger coverage values.

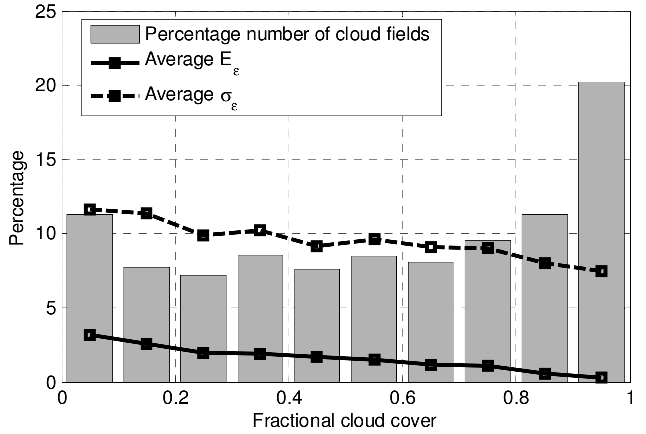

Fig. 7. Trend of the average E_ε (solid line) and σ_ε (dashed line) as a function of the fractional cloud cover f, together with the percentage number of cloud fields considered in each class (gray bars).

The final information required to synthesize cloud fields is the spatial distribution of L (including pixels with $L = 0$) that we have investigated by means of the spatial correlation index defined as [17]:

$$\rho(\mathbf{x},\mathbf{y}) = \frac{\mathrm{E}[L(\mathbf{x}) \cdot L(\mathbf{y})] - \mathrm{E}[L(\mathbf{x})]E[L(\mathbf{y})]}{\sigma[L(\mathbf{x})]\,\sigma[L(\mathbf{y})]} \tag{4}$$

E[•] and σ[•] are the mean and standard deviation operators, whilst $L(\mathbf{x})$ and $L(\mathbf{y})$ are the cloud liquid water content time series, respectively relative to pixels $\mathbf{x}$ and $\mathbf{y}$ in each 200 km×200 km cloud map. In calculating ρ, we have assumed that the cloud field is stationary in space (like for rain fields [17]) and independent of the site in Europe the target area refers to; this implies that the spatial correlation between two points depends (mostly) on their distance and only marginally on their position, i.e.:

$$\rho(\mathbf{x},\mathbf{y}) = \rho(d = |\mathbf{x} - \mathbf{y}|) \tag{5}$$

Fig. 8 depicts the spatial correlation of L obtained by averaging ρ values relative to all the couples of pixels at the same distance d (light dashed gray line). The figure also reports the spread of ρ around its average value (gray scale density scatter plot, higher concentration in darker areas), which allows to visually infer the degree of the cloud field spatial stationarity, which has been assumed in this work for modeling purpose. For convenience, the average spatial correlation of rainfall as obtained from a set of rain fields derived by the NIMROD weather radar network [18] has been added to Fig. 8 in order to show the much higher spatial variability of precipitation with respect to clouds.

As clarified in [12] by Bell, the stochastic approach proposed here to synthesize realistic cloud fields starts from the generation of random Gaussian fields, whose spatial correlation $\rho_G(d)$ needs to be provided as input to the generation process. To this aim, we have estimated the average $\rho_G(d)$ by first turning each MODIS cloud field into a truncated Gaussian field, which, under the assumption of lognormal distribution for L, corresponds to inverting (9) reported in Section IV below. Afterwards the spatial correlation of the random Gaussian process has been calculated from converted maps using the same definition of ρ in (4) and assuming again spatial stationarity. The resulting average $\rho_G(d)$ is well represented by the following analytical expression (the distance d is expressed in km):

$$\rho_G(d) = 0.35\, e^{-\frac{d}{7.8}} + 0.65\, e^{-\frac{d}{225.3}} \tag{6}$$

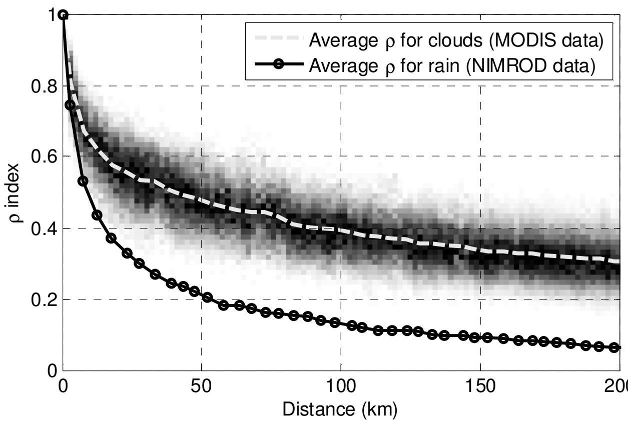

Fig. 8. Average decorrelation with distance of the integrated cloud liquid water calculated from MODIS data (light gray dashed line) and of rainfall calculated from NIMROD data (black solid line with circles, results extracted from [18]). Also depicted is the spread of ρ around its average value (gray scale density scatter plot, higher concentration in darker areas).

IV. HORIZONTAL CLOUD FIELDS SYNTHESIS

Based on the expressions in (2), the horizontal cloud field synthesis in a target area can be achieved from the knowledge of E_L and f. In turn, this information can be extracted from NWP products; in this work, we made reference to the ECMWF ERA40 dataset [13]. In particular, we extracted E_L and f with temporal sampling of 6 hours (i.e. nearly instantaneous values every 6 hours) and spatial resolutions of 2°×2° (latitude×longitude) respectively, the latter approximately corresponding to 200 km×200 km in Europe.

The procedure consists in the following steps:

1. From the ERA40 database, extract the fractional cloud cover (f_{ERA}) and the spatial average of the integrated cloud liquid water (L_{ERA}). For a given site with coordinates (lat, lon), f_{ERA} and L_{ERA} will result from the bilinear interpolation of the values relative to the four surrounding grid pixels, as suggested by the ECMWF. Thus $E_L = L_{ERA}$ and $f = f_{ERA}$.
2. From E_L, derive μ_p and σ_p as from the expressions in (2), which completely define $p(S_L|E_L)$.
3. Randomly extract S_L from the lognormal distribution $p(S_L|E_L)$ derived at step 2. As a result, S_L values associated to consecutive E_L (= L_{ERA}) and f (= f_{ERA}) samples will be uncorrelated.
4. Generate a random Gaussian field $g(x,y)$ with the spatial correlation ρ_G in (6) according to the procedure outlined in [12].
5. Calculate μ_{LN} and σ_{LN} of the lognormal distribution characterizing the cloud map to be generated (obviously for $L > 0$ mm), by inverting the following equation system:

$$\begin{cases} E_L = f \exp\left(\mu_{LN} + \sigma_{LN}^2/2\right) \\ S_L = \sqrt{f \exp\left(2\mu_{LN} + 2\sigma_{LN}^2\right) - f^2 \exp\left(2\mu_{LN} + \sigma_{LN}^2\right)} \end{cases} \tag{7}$$

Whilst E_L and S_L come from the NWP database, the right-hand sides of the equations in (7) express the mean and standard deviation values of a mixed random variable whose value is 0 with probability 1-f (cloud free fraction of the map) and is extracted from a lognormal distribution (with parameters μ_{LN} and σ_{LN}) with probability f.
Thus, the explicit expressions for μ_{LN} and σ_{LN} are:

$$\begin{cases} \mu_{LN} = \ln\left[\dfrac{1}{f^{1.5}} \dfrac{E_L^2}{\sqrt{E_L^2 + S_L^2}}\right] \\ \sigma_{LN} = \sqrt{\ln\left[f\left(\dfrac{S_L^2}{E_L^2} + 1\right)\right]} \end{cases} \tag{8}$$

6. Turn the Gaussian field $g(x,y)$ into a lognormal (cloud) field $C(x,y)$ according to:

$$\begin{aligned} &C(x,y) = 0, \quad g < g_{th} \\ &C(x,y) = \exp\left\{\mu_{LN} + \sqrt{2}\sigma_{LN}\operatorname{erfc}^{-1}\left[\frac{1}{f}\operatorname{erfc}\left(\frac{g(x,y)}{\sqrt{2}}\right)\right]\right\}, \quad g \geq g_{th} \end{aligned} \tag{9}$$

where $g_{th} = \sqrt{2}\,\operatorname{erfc}^{-1}(2f)$ and erfc is the complementary error function.

As a result, SMOC allows to generate horizontal cloud fields (random lognormal fields) maintaining the basic integral information (f_{ERA} and L_{ERA}) and reproducing the spatial correlation observed in real cloud fields. Fig. 9 shows a sample synthetic cloud field (L in mm) reflecting the input values extracted from the ERA40 database with high accuracy (f_{ERA} = 0.76 and L_{ERA} = 0.2 mm): considering the generation of approximately 7300 synthetic fields, the root mean square value of the relative percentage error in reproducing f_{ERA} and L_{ERA} (definition as in (3)) is 0.1% and 3.2%, respectively.

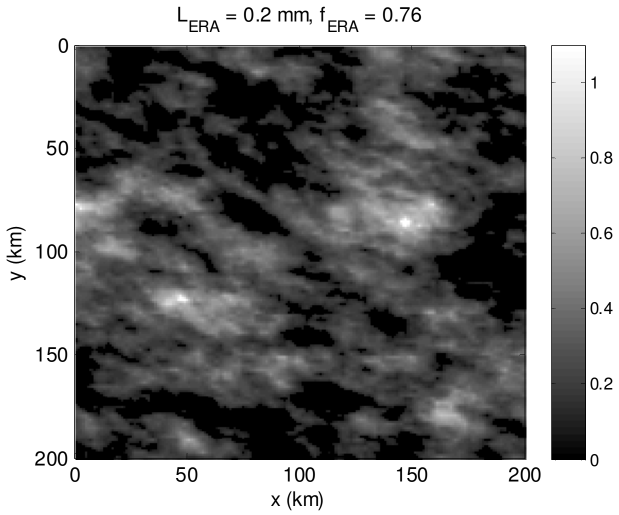

Fig. 9. Example of a cloud field generated by SMOC starting from ERA40 data with 2°×2° spatial resolution and 6-hour temporal resolution.

V. Vertical development of clouds

To investigate the vertical profile of clouds, we have taken advantage of the data collected by the NASA Earth Observation Satellite CloudSat. Launched in 2006, the LEO satellite orbits in formation as part of the A-Train constellation (Aqua, CloudSat, CALIPSO, PARASOL and Aura satellites) and features a 94-GHz nadir-looking radar (Cloud Profiling Radar, CPR) designed to observe clouds and precipitation from space. As an advantage over passive sensors on-board EO satellites for cloud monitoring, CPR allows to measure, with high spatial detail (the footprint is 1.4 km×1.7 km and the vertical profile is sampled every 240 m), the full distribution of the liquid water content $w(h)$ in clouds between the surface and 25 km of altitude [19].

We focused on the 2B-CWC-RVOD product developed and distributed by the Cooperative Institute for Research in the Atmosphere (CIRA) at the Colorado State University [20]. Indeed, these data are expected to maximize the accuracy in estimating $w(h)$ because, besides standard calibration and quality checks common to all 2B Level Products, 2B-CWC-RVOD data originate from the combination of CPR-derived profiles and from the concurrent Visible Optical Depth measurements collected by the MODIS sensor that is part of the A-Train constellation as well [20].

The processing of a full year (2009) of CloudSat data collected over Europe allowed to identify and isolate more than 50000 single cloud vertical profiles (we assume hereinafter that only one cloud is prevailing over the target area, as clearly shown by CloudSat-derived profiles of $w(h)$), two examples of which are shown in Fig. 10: data in the left graph, associated to a low cloud of rather limited vertical extent, were collected over Romania in January, whilst the right graph shows a much thicker cloud lying over Southern Italy in June (black solid lines with circles).

The preliminary visual inspection of CloudSat-derived profiles pointed out that, for most clouds, the trend of $w(h)$ with height is asymmetric (the peak value of the liquid water content being typically closer to the cloud base) and that, as in the sample Fig. 10, profiles slowly decay to zero with increasing height. According to these features, we have selected the following analytical expression to model $w(h)$ (h km a.m.s.l.):

$$\tilde{w}(h) = \begin{cases} \dfrac{L}{b^a\,\Gamma(a)}\left(h-h_0\right)^{a-1} e^{-(h-h_0)/b} & \text{for } h \geq h_0 \\ 0 & \text{for } h < h_0 \end{cases} \tag{10}$$

In (10), a and b are parameters regulating the shape of $\tilde{w}(h)$, h_0 is the cloud base height (addressed later on in this section), Γ is the Gamma function and L is the integrated liquid water content of the cloud. As a matter of fact, the analytical expression in (10) resembles the Gamma probability density function (PDF), the only difference being that its integral from 0 to infinity is L instead of 1. The choice of (10) allows to easily constrain the analytical profile to a given integrated liquid water content L but the expression needs to be truncated to model real clouds (in fact $\tilde{w}(h) \rightarrow 0$ g/m^3 only for $h \rightarrow +\infty$): we set $\tilde{w}(h) = 0$ for $\tilde{w}(h) < w_{th}$. In order to identify the optimum value for w_{th} and, more in general, to investigate in detail the vertical distribution of $w(h)$, each cloud in the database has been characterized in terms of L and in terms of the best set of a and b parameters maximizing the agreement between the measured profile and (10) (see Fig. 10). The overall assessment of (10) to model clouds indicates that the average (over the whole database) root mean square value of the error $\varepsilon_w(h) = \tilde{w}(h) - w(h)$ is 0.038 g/m^3, a good score considering, as a reference, that the average value of $w(h)$ is approximately 0.16 g/m^3. These results are achieved by setting w_{th} = 0.06L, a good trade-off between maximizing the accuracy in the cloud thickness D estimate (considering the whole cloud database, the mean and root mean square values of the relative error $\varepsilon_{\Delta h} = 100\,(\tilde{D} - D)/D$ are 3.1% and 15.2%, respectively) and minimizing the underestimation of L caused by the truncation of the profile (considering L values ranging from 10^{-3} to 2.5 mm, the mean and root mean square values of the relative error $\varepsilon_L = 100\,(\tilde{L} - L)/L$ are -4.6% and 4.8%, respectively). The latter constrain has been privileged in determining w_{th} because of its direct impact on the future use of SMOC to predict cloud attenuation induced on electromagnetic waves.

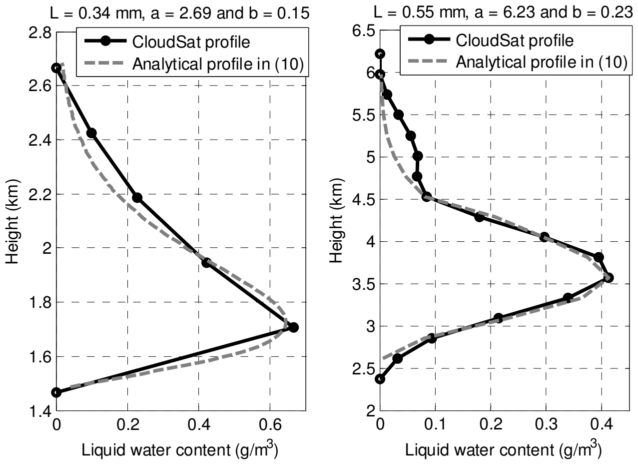

Fig. 10. Sample vertical profiles of the liquid water content $w(h)$ as measured by the CPR on-board the CloudSat satellite and as estimated using the expression in (10); left side: profile collected over Romania in January 2009, right side: profile collected over Southern Italy in June 2009.

Fig. 11 depicts the density scatter plot (gray scale, higher concentration in darker areas) between the cloud liquid water content L and the parameter a; data indicate that the probability of occurrence is higher for "light" clouds (L roughly lower than 0.2 mm), and that the regression curve between L and a is (white line with circles):

$$a = 4.27\, e^{-4.93(L+0.06)} + 54.12\, e^{-61.25\,(L+0.06)} + 1.71 \tag{11}$$

Similarly, the relationship between a and b, reported in Fig. 12 again in terms of density scatter plot (gray scale), is (white line with circles):

$$b = 3.17\, a^{-3.04} + 0.074 \tag{12}$$

As a result, from the knowledge of L and exploiting (11) and (12) to estimate a and b in (10), a realistic cloud vertical profile can be derived for each of the pixels in the synthetic maps of L generated by SMOC.

The final information we need is the cloud base height h_0. We assume that the base height of all the clouds in a target area is fairly constant; the values of h_0 have been extracted from the CloudSat vertical profiles of $w(h)$ too: Fig. 13 shows the PDF of h_0 (km a.m.s.l.), as well as the fitting generalized extreme value distribution (shape parameter ξ = 0.484, scale parameter σ = 0.582 and location parameter μ = 0.987). Data indicate that most cloud bases lie around h_0 = 1 km, which is expected by the prevalence of stratiform clouds in Europe with rather limited integrated liquid water content (see the darker area in the density scatter plot of Fig. 11).

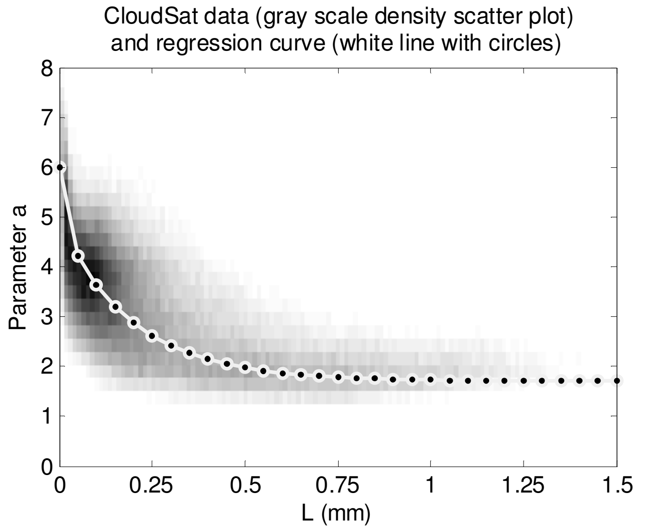

Fig. 11. Relationship between L and a in (10): density scatter plot (gray scale, higher concentration in darker areas) based on CloudSat data and regression curve (white line with circles).

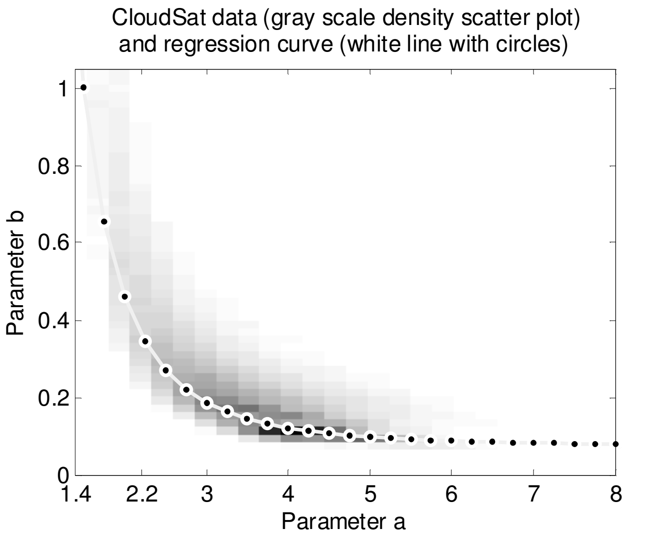

Fig. 12. Relationship between a and b in (10): density scatter plot (gray scale, higher concentration in darker areas) based on CloudSat data and regression curve (white line with circles).

As an example of the application of SMOC, Fig. 14 depicts the spatial distribution of $w(h)$ relative to the synthetic map of L shown in Fig. 9. Specifically, the data depicted in the bottom graph refer to the y/h plane for x = 70 km with h_0 = 1 km, whilst the top graph reports the associated integrated liquid water content L as a function of y. According to the model, the highest L values in the map are associated to thicker clouds characterized by large liquid water contents; moreover, the fractional cloud cover varies with the height, as it is typically the case of real cloud fields.

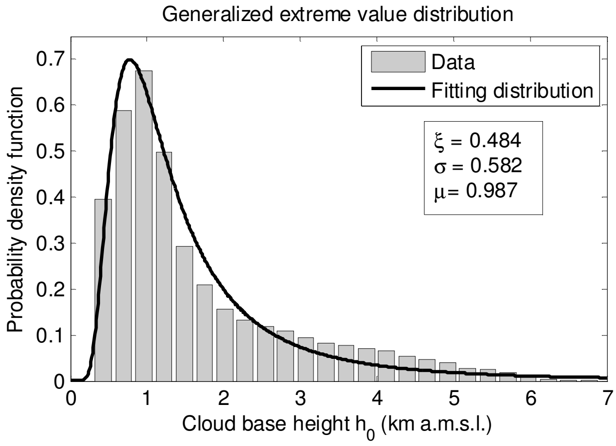

Fig. 13. Probability density function of h_0, the cloud base height, derived from the lowermost cloud of each CloudSat profile. Empirical data and MLE generalized extreme value distribution.

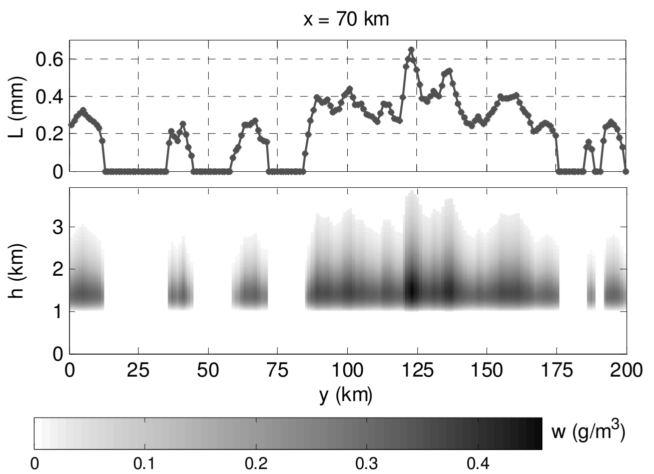

Fig. 14. Spatial distribution of w calculated by means of (10), (11) and (12) starting from the SMOC synthetic map of L shown in Fig. 9. Bottom graph: w on the y/h plane for x = 70 km; top graph: associated integrated liquid water content L as a function of y.

VI. Validation of SMOC

To validate SMOC, the full time series of f_{ERA} and L_{ERA} in the period 1996-2000 have been used. Approximately 7300 cloud fields have been generated and processed to calculate first- and second-order statistics of L. Concerning the former, the well documented quasi-ergodicity property of rain fields has been extended to clouds, which implies that all L values of each map have been included in the Complementary Cumulative Distribution Function (CCDF) of L [21],[22].

As reference statistics of L for the performance assessment of SMOC, we have exploited an extensive set of RAOBS data collected routinely twice a day for ten years (1980-1989) in 14 sites subject to very different climates (refer to Fig. 15 where the sites are indicated as circles). Specifically, temperature, pressure and relative humidity profiles have been used to derive the liquid water content (hence L) by means of the already mentioned TKK cloud detection algorithm [6].

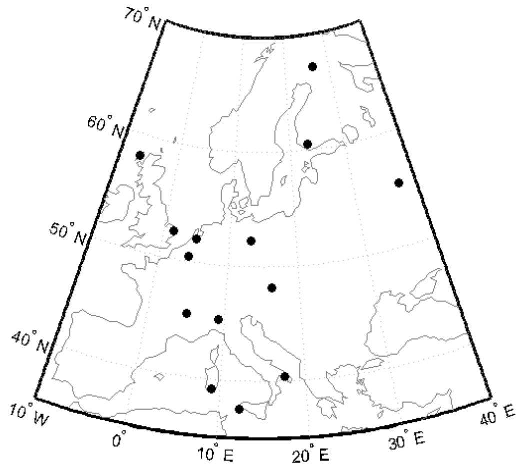

Fig. 15. Sites where RAOBS data have been collected.

As an example, Fig. 16 compares the CCDF of L estimated from SMOC with the one obtained from the RAOBS data collected at Milano Linate airport. Despite the two datasets are neither concurrent nor of the same duration, the comparison in Fig. 16 is meaningful because the analysis of RAOBS data revealed that 5 years are sufficient for the CCDF of L to practically be long-term representative.

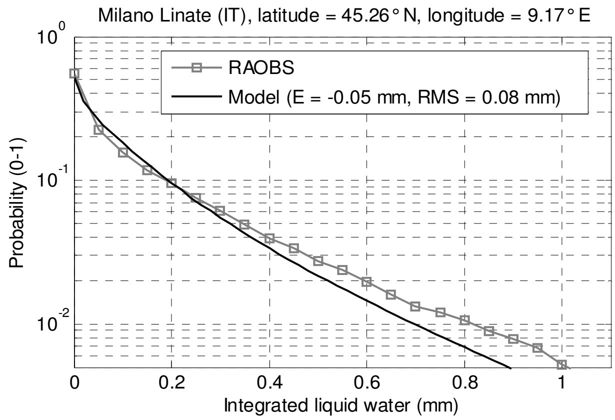

Fig. 16. Validation of SMOC against RAOBS data collected at Milano Linate airport (1980-1989) coupled with the TKK model. Input values to SMOC are f_{ERA} and L_{ERA} and values extracted from the ERA40 database (1996-2000).

The good agreement between the two curves in Fig. 16 is quantified in the figure legend, which reports the average (E_ψ) and root mean square (RMS_ψ) values of the error ψ ($P \geq 5\times10^{-3}$), defined as:

$$\psi(P) = L_E(P) - L_R(P) \tag{13}$$

In (13), $L_E(P)$ and $L_R(P)$ represent the estimated and reference integrated cloud liquid water contents, respectively, relative to the same probability level P.

Fig. 17 extends the prediction accuracy assessment by reporting E_ψ and RMS_ψ for all the 14 European sites where RAOBS data are available (as in Fig. 16, P ranges between

5×10^{-3} and 1). Results in Fig. 17 show that SMOC achieves an overall good accuracy in predicting first-order statistics of L, being the results dependent both on the mathematical formulation of the model and on the input f_{ERA} and L_{ERA}. Although not shown here for brevity's sake, the analysis on the model's input values indicates that the largest prediction errors reported in Fig. 17 (i.e. site 4, Hemsby – UK, and site 14, Moscow – Russia) are mainly associated to the underestimation of L_{ERA}. On the other hand, SMOC is not univocally tied to the ERA40 database; indeed, the model can receive the inputs from any (possibly more accurate) meteorological database including gridded values of fractional cloud cover and average integrated liquid water content.

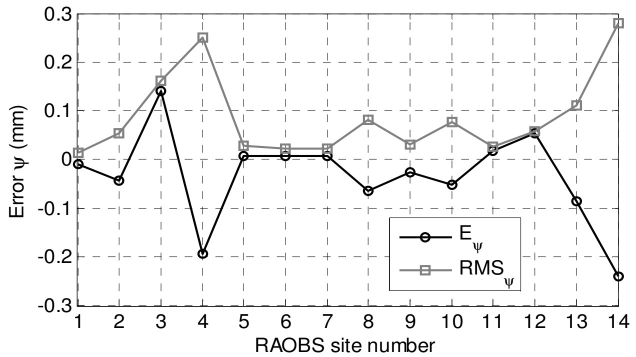

Fig. 17. Validation of SMOC against all RAOBS data available (1980-1989) coupled with the TKK model: first-order statistics. Input values to SMOC are f_{ERA} and L_{ERA} values extracted from the ERA40 database (1996-2000).

The ability of SMOC in reproducing the spatial distribution of L was tested with reference to the average decorrelation trend (ρ as defined in (4)) extracted from MODIS data as in Fig. 8. Fig. 18 compares such a trend (large black dashed line) with all the ones associated to the synthetic cloud fields generated by SMOC for the above 14 European sites (thin gray lines). Differences in ρ from site to site are plausible because of the different type of clouds occurring at different sites: the least and the most steep curves are associated to Stornway (UK, latitude = 58.13° N, longitude = -6.59° E) and Cagliari (IT, latitude = 39.15° N, longitude = 9.03° E), respectively experiencing mostly stratiform-like (large horizontal extent and low total liquid water content) and cumulus-like clouds (limited horizontal extent and high total liquid water content). Overall, the agreement between the MODIS curve and the average decorrelation trend obtained from SMOC (dashed gray line) is very good.

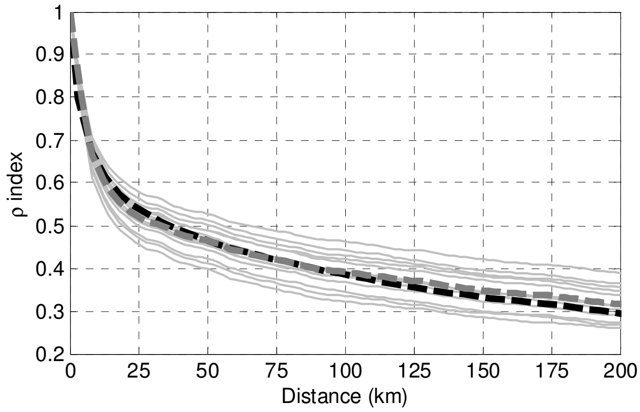

Fig. 18. Validation of SMOC against MODIS data: second-order statistics. Input values to SMOC are f_{ERA} and L_{ERA} values extracted from the ERA40 database (1996-2000).

VII. Conclusions

This contribution presents SMOC (Stochastic Model Of Clouds), a methodology to synthesize 3-D spatially correlated cloud fields in temperate regions (as wide as 200 km×200 km and with horizontal 1 km×1 km resolution) from Numerical Weather Prediction (NWP) products (reanalysis or forecasts), i.e. the fractional cloud coverage f and the average cloud liquid water content E_L, both relative to the target area of interest. SMOC relies on the stochastic approach originally proposed by Bell and its internal parameters have been determined from high-resolution cloud fields observed by the MODIS sensor over Europe. The analysis of such dataset revealed that L values tend to be lognormally distributed in each cloud map and that clouds are much more correlated in space than rainfall (at 100 km distance, the spatial correlation index ρ is roughly 0.4 and 0.15 for the former and the latter, respectively). Moreover, a large set of liquid water content profiles collected by the CPR on-board the CloudSat satellite was used to devise a simple yet effective model for the vertical development of clouds; profiles of liquid water content $w(h)$ follow an analytical expression that resembles the Gamma probability density function and whose parameters directly depend on L.

The model's accuracy has been evaluated against radiosonde data collected in 14 sites spanning from Northern (Sodankyla, Finland) to Southern (Trapani, Italy) Europe: overall, predicted CCDFs of L are in good agreement with the ones estimated from the RAOBS data used as input to the TKK cloud detection model (considering all sites, average root mean square of the error on the CCDF of L equal to 0.09 mm). Moreover, whilst the average spatial correlation characterizing the synthetic cloud fields generated by SMOC fairly well reproduces the one derived from the MODIS database, the trend of ρ with distance in SMOC maps varies from site to site because of the different type of clouds expected to occur. This points out the ability of SMOC to reflect the main local features of clouds inherently embedded in the ERA40 data, which, anyway, represent only one of the possible datasets that inputs to SMOC can be extracted from (e.g. also forecast data could be employed). Moreover, SMOC is flexible as its parameterization might be changed (if necessary) to extend its validity also to tropical/equatorial regions, which, in comparison to temperate regions, are more frequently subject to cumulus clouds with higher liquid water content.

SMOC represents a basic block of a simulator of weather disturbances affecting radio wave propagation, primarily intended to support the design and performance assessment of Earth-space Communication Systems (EHF range or at optical wavelengths) but also of possible interest for all the applications involving radiative transfer in the atmosphere.

ACKNOWLEDGMENT

This effort was partially sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA8655-13-1-3081. The U.S Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon. The authors would also like to acknowledge: the MODIS and CloudSat mission scientists and associated NASA personnel for the production of the data used in this work; the ECMWF for granting access to the cloud dataset included in the ERA40 database; Dr. Martellucci from the European Space Agency for the provision of radiosonde data.

REFERENCES

[1] R. K. Crane, "Electromagnetic Wave Propagation Through Rain," Wiley, 1996.

[2] L. Luini, R. Nebuloni, C. Capsoni, "Effectiveness of Multisite Diversity Schemes to Support Optical Systems in Scientific Missions," *Optical Engineering*, 53 (2), 026104 (Feb 18, 2014).

[3] E. Altshuler, R. Marr, "Cloud attenuation at millimetre wavelengths," *IEEE Transactions On Antennas and Propagation*, 37, 1473-1479, 1989.

[4] F. Dintelmann, G. Ortgies, "Semi-empirical model for cloud attenuation prediction," *Electron. Lett.*, 25, 1487-1479, 1989.

[5] A. Dissanayake, J. Allnutt, F. Haidara, "Cloud attenuation modelling for SHF and EHF applications," *Int. J. Satell. Commun.* 2001; 19:335-345.

[6] E. Salonen, S. Uppala, "New prediction method of cloud attenuation," *Electron. Lett.*, vol. 27, no. 12, pp. 1106–1108, Jun. 1991.

[7] Attenuation due to clouds and fog. Geneva, 2013, ITU-R recommendation P.840-6.

[8] H. J. Liebe, "MPM - An atmospheric millimeter-wave propagation model," *Int. J. Infrared and Millimeter Waves*, 1989, 10, pp. 631450.

[9] L. Luini, C. Capsoni, "A Rain Cell Model for the Simulation and Performance Evaluation of Site Diversity Schemes," *IEEE Antennas and Wireless Propagation Letters*, Vol. 12, No. 1, Page(s): 1327-1330, 2013.

[10] ESA/ESTEC/Contract 17760/03/NL/JA, "Characterisation and Modelling of Propagation Effects in 20-50 GHz Band: Final Report," Volume 4592 di ESA CR, 2006.

[11] P. Garcia, A. Benarroch, J. M. Riera, "Spatial distribution of cloud cover," *Int. J. Satell. Commun.* 2008; 26:141–155.

[12] T. L. Bell, "A Space-Time Stochastic Model of Rainfall for Satellite Remote-Sensing Studies," *Journal Of Geophysical Research*, Vol. 92, No. D8, pp. 9631-9643, 1987.

[13] S. M. Uppala, P. W. Kallberg, et al., "The ERA-40 Re-analysis," *Quart. J. Roy. Meteor. Soc.*, 131, pp. 2961-3012, 2005.

[14] X. Xiong, W. Barnes, "An Overview of MODIS Radiometric Calibration and Characterization," *Adv. Atmos. Sci.*, vol. 23, issue 1, pp. 69-79, 2006.

[15] Members of the MODIS Characterization Support Team, "MODIS Level 1B Product User's Guide," July 20, 2012, available online at http://mcst.gsfc.nasa.gov/sites/mcst.gsfc/files/file_attachments/M1054D_PUG_083112_final.pdf.

[16] C. Lynnes, R. Strub, E. Seiler, T. Joshi, P. MacHarrie, "Mirador: A Simple Fast Search Interface for Global Remote Sensing Data Sets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 92-96, Jan. 2009.

[17] L. Luini, C. Capsoni, "On the Relationship Between the Spatial Correlation of Point Rain Rate and of Rain Attenuation on Earth-Space Radio Links," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 10, Page(s): 5255 - 5263, October 2013.

[18] L. Luini, C. Capsoni, "The Impact of Space and Time Averaging on The Spatial Correlation of Rainfall," *Radio Sci.*, 47, RS3013, doi:10.1029/2011RS004915, 2012.

[19] G. L. Stephens, D. G. Vane, R. J. Boain, G. G. Mace, K. Sassen, Z. Wang, A. J. Illingworth, E. J. O'Connor, W. B. Rossow, S. L. Durden, S. D. Miller, R. T. Austin, A. Benedetti, C. Mitrescu, the CloudSat Science Team, "The CloudSat mission and the A-TRAIN: A new dimension to space-based observations of clouds and precipitation," 2002, *Bull. Am. Met. Soc.*, 83, 1771-1790.

[20] N. Wood, "Level 2B Radar-Visible Optical Depth Cloud Water Content (2B-CWC-RVOD) Process Description Document," CloudSat project, Version 5.1, 23 October 2008; available online at http://www.cloudsat.cira.colostate.edu/ICD/2B-CWC-RVOD/2B-CWC-RVOD_PDD_V5p1.pdf

[21] L. Luini, C. Capsoni, "MultiEXCELL: a new rain field model for propagation applications," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 11, Page(s): 4286 – 4300, November 2011.

[22] E. A. B. Eltahir, R. L. Bras, "Estimation of the fractional coverage of rainfall in climate models," *J. Climate*, vol. 6, pp. 639–644, 1993.