CLOUDS SPATIAL DISTRIBUTION AND INFRARED SOUNDING CONTAMINATION: AN IMPACT STUDY ON INFRARED SATELLITE SOUNDING OF THE ATMOSPHERE

V. Tramutoli
Università della Basilicata, Potenza, Italy

C. Pietrapertosa, V. Lanorte
Istituto di Metodologie Avanzate di Analisi Ambientale. IMAAA/CNR
Tito Scalo (PZ), Italy

1. INTRODUCTION

Cloud handling is a crucial task in the infrared sounding of the atmosphere from satellite. The presence of clouds can heavily reduce quality and coverage of the main satellite based IR products for meteorological and climatological applications. Most of these require completely clouds-free soundings so that cloud-contaminated infrared radiances are treated as missing data in the context of filtering schemes devoted to the production of clear-radiance 2-D fields. Several techniques, able to deal with partly-cloudy soundings, have been proposed, even if their reliability seems to depend on some specific conditions/assumptions (e.g. Susskind et al., 1984, Huang et al., 1986, Eyre, 1989). At the present improvements in the quality of satellite based products, seems to mainly depend on future instrument capability to achieve more and better distributed cloud-free soundings (see for example Amato et al., 1991, Cuomo et al., 1993) and/or on the reliability of new techniques able to obtain similar results by handling also partly cloud-contaminated radiances. In this paper cloud spatial distribution, at different scales, has been analyzed in order to assess its impact on new instrument design and on the most common TOVS-data applications. By using AVHRR data coming from several NOAA satellite passes over Europe, spatial resolutions from 3 to 100 km have been simulated and spatial distribution of differently clouds-contaminated soundings (up to 25% of the FOV area) investigated.

Section 2 describes the used dataset. Section 3 and 4 are devoted to obtain a description of cloud spatial distribution and a quantitative assessment of its impact on cloud-clearing techniques (based on the interpolation between nearest neighbor clear FOVs) and future IR sounders design. In Section 5 we will draw some preliminary conclusions.

2. DATA SOURCE

Table 1 shows the characteristics of AVHRR passes selected for our analysis. Figure 1 shows fields of brightness temperature in channel 4 for all passes. A clearness-index ($C_c$), taking values from 0 (for completely overcast) to 100 (for completely clear) FOVs, has been introduced in order to measure the percentage of FOV’s area which is cloud-free. Full resolution AVHRR cloud masks, products of CMS operational cloud detection
Fig. 1: Used AVHRR passes. From left to right A, B, C, D, E and F (GAC reductions on AVHRR BT in channel 4).

<table>
<thead>
<tr>
<th>PASS</th>
<th>DATE</th>
<th>GMT</th>
<th>ORBIT</th>
<th>SCAN LINES</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>17.12.91</td>
<td>13:29</td>
<td>16635</td>
<td>4782</td>
</tr>
<tr>
<td>B</td>
<td>17.12.91</td>
<td>15:10</td>
<td>16636</td>
<td>4566</td>
</tr>
<tr>
<td>C</td>
<td>11.02.89</td>
<td>03:17</td>
<td>01970</td>
<td>4798</td>
</tr>
<tr>
<td>D</td>
<td>12.02.89</td>
<td>03:07</td>
<td>01984</td>
<td>4881</td>
</tr>
<tr>
<td>E</td>
<td>11.02.89</td>
<td>01:36</td>
<td>01969</td>
<td>4767</td>
</tr>
<tr>
<td>F</td>
<td>12.02.89</td>
<td>13:19</td>
<td>01962</td>
<td>4711</td>
</tr>
</tbody>
</table>

Table 1. Passes identification.

Figure 2. Percentage of the total clear detectable area at different spatial resolutions: averages over all passes
algorithm (Derrien et al., 1993), have been processed in order to obtain cloud masks at different spatial resolution and for \( C_c \geq 75, 95 \) and 100. From now on we will use the term clear in order to indicate FOVs which have a clearness-index not lower than the specified threshold \( C_c \).

3. CLEAR DETECTABLE AREA: DIFFERENTIAL ANALYSIS

To understand clear areas distribution across the scene we used a continuous scanning design and, to better control the evolution of the measured quantities as simulated spatial resolution is varied, square FOVs of linear sizes, \( L = 96, 48, 24, 12, 6, 3 \), AVHRR pixels (i.e. between 3.3 Km and 105.6 Km at the nadir) have been simulated. So, each simulated FOV of linear size \( L_i \) is completely inside the one corresponding to the immediately lower linear resolution \( L_{i-1} \) and the FOV's number increases, exactly inside the same area, by four times. Using the AVHRR data set described in section 2, cloud masks have been computed for each of the above linear spatial resolutions using values 75, 95 and 100 for \( C_c \) thresholds. Let is \( N_{C_c}(L) \) the number of clear \( L \times L \) boxes when the cloud contamination threshold is \( C_c \). For this way, as an example, \( N_{75}(3) \), is the number of clear \( 3 \times 3 \) boxes when \( C_c = 75 \). At every fixed \( C_c \) the maximum "clear" detectable area \( A_{C_c}(3) \) is the one corresponding to the higher spatial resolution \( L_6 = 3km \), i.e. the one detected as clear using the cloud contamination threshold \( C_c \) and FOV's sizes of \( 3 \times 3 \) AVHRR pixels. Measuring this quantity in \( 3 \times 3 \) AVHRR boxes (B3) units we can write:

\[
A_{C_c}(3) = N_{C_c}(3)
\]

As consequence of the chosen, sampling design and \( \{L_i\} \) values, we can define in the same way the quantity

\[
A_{C_c}(L_i) = N_{C_c}(L_i) \times 4^{6-i} \quad \{i = 1...6\}
\]

which simply represents, in B3 units, the clear detectable area, at the resolution and cloud contamination threshold, corresponding to \( L_i \) and \( C_c \). Note that \( L_i \) values decrease from 96 to 3 as \( i \) increases from 1 to 6. To evaluate the net fractional contribution, coming from each increasing of spatial resolution, to the total clear detectable area \( A_{C_c}(3) \) we measured the quantity:

\[
F_{C_c}(L_i) = 100 \cdot \frac{A_{C_c}(L_i) - A_{C_c}(L_{i-1})}{A_{C_c}(3)}
\]

which represents, at each spatial resolution, associated with the linear FOV size \( L_i \), the percentage of new clear FOVs (over the total clear detectable area \( A_{C_c}(3) \)), detected as a consequence of a 4 times increasing of spatial resolution from \( L_{i-1} \) to \( L_i \). It is easy to demonstrate that: \( \sum_{i=1}^{6} F_{C_c}(L_i) = 100 \) At each fixed \( C_c \) this index gives a quantitative indication on how much convenient, in term of new clear detectable area, is to increase
spatial resolution, compared with the maximum clear detectable area at the highest spatial resolution considered (L=3 km). Results, averaged over all passes, are showed in Figure 2. We can see as increasing spatial resolution become more and more important as far less clouds-contaminated radiances (i.e. greater values of \( C_c \)) are required. Using a \( C_c = 75 \) threshold, quite half of the maximum detectable clear area (\( A_{75}(3) \)) is already detected at \( L = 96 km \) with improvements only around a 10% at each doubling of linear spatial resolution. To obtain a comparable \( 0.5 \cdot A_{C_c}(3) \) result we need to use at least \( L = 24 km \) for \( C_c = 95 \) and \( 6 km < L < 12 km \) when a \( C_c = 100 \) threshold is required. Figure 3 graphically explains, for a selected box of 300 x 216 AVHRR pixels coming from pass B, the joint effects of different choices of \( C_c \) and \( L_t \). Figure 3a shows the full resolution AVHRR cloud mask (35% of AVHRR FOVs are clouds-free and black colored). Figures from 3b to 3c show results obtained respectively for \( C_c = 75, 95 \) and 100. Different colored areas correspond to different spatial resolution and show on the pictures new clear detected areas passing from lower to higher spatial resolutions.

Figure 3. Selected box (300x200AVHRR pixels) from pass B: a) cloud mask: clear FOVs are black colored; clear detected area using different spatial resolutions (different colors) for b) \( C_c = 75 \); c) \( C_c = 95 \); d) \( C_c = 100 \)
Two obvious conclusions can be drawn:

1) In terms of fractional dimension of the clear detected area: degrading cloud masks at low spatial resolution give results which are roughly comparable with the ones achievable, for clouds-free or less clouds-contaminated radiances, by increasing spatial resolution. Looking, in fact at Figure 2, a 65% of the maximum clear detectable area $A_{C_e}(3)$ is obtained with $L \approx 24 \text{km}$ at $C_e = 75$, $L \approx 12 \text{km}$ at $C_e = 95$ and $L \approx 6 \text{km}$ at $C_e = 100$.

2) In terms of uniformity of clear FOVs distribution it is evident from pictures in Figure 3 that improvements are coming quite only from increased spatial resolutions at every $C_e$.

The impact of clear FOVs spatial distribution on cloud clearing (one of the most common steps in TOVS-data processing), can be better understood from the following tests.

3.1 Assessing the impact on cloud-clearing

We computed the quantity:

$$FC_{C_e}(R, L) = 100 \cdot \frac{A(R)}{A - A_{C_e}(L)}$$

where:

$A(R)$ is the area where clear radiances can be predicted by using whatever interpolation technique able to produce clear radiances in cloudy locations by using only available clear soundings within a circle of radius $R \text{ km}$; $A$ is the total area of the scene; $A_{C_e}(L)$ is the "clear" detected area at the spatial resolution corresponding to a square box of linear size $L$ and by using a $C_e$ threshold. For this way $FC_{C_e}(R, L)$ gives the percentage of cloudy area that can be "cloud-cleared" using a linear FOV size $L$, a cloud contamination threshold corresponding to $C_e$ and an useful interpolation radius $R$. Values of $FC_{C_e}(R, L)$ against $L$ averaged over all passes, are plotted for $C_e = 75, 95$ and $100$, in Figures 4a, 4b and 4c, respectively at $R = 50, 100$ and $200 \text{ km}$. Starting from the lowest spatial resolution (corresponding to $L = 96$ AVHRR pixels) each step, right to left on the abscissa, corresponds to an increasing of spatial resolution by a factor 4. Obviously $FC$ is an increasing function of $L$ and, at each fixed $L$, takes values which are greater as far as lower values of $C_e$ are considered. If we consider, as a starting point, a spatial resolution corresponding to a fixed $L$ and the most conservative thresholds $C_e = 100$, in order to increase $FC$ we have two way: a) to reduce linear FOV size (which require improvements in instrument design), b) to reduce $C_e$ thresholds (which requires progress in techniques dealing with cloud-contaminated radiances). More technology, more science, is an alternative which we, simplistically, could apply to these two possibilities (which obviously could be pursued together) in order to remember that each of them requires technical or scientifical
improvements. Looking to plots 5a, 5b and 5c, we can see as improving spatial resolution gives always better results than radiances quality degradation as soon values of $L \leq 24$ are reached. These effects become more and more noticeable as far higher spatial resolutions are considered and lower values of $R$ are used. At higher values of $R$ (see Figure 4c) they become less marked or appreciable only at the immediately higher spatial resolutions. By using $R = 200 \ km$ no substantial improvements ($FC$ is quite constant) we find by increasing spatial resolution from $L = 48$ up to $L = 24$, using cloud contaminated radiances and, also for $C_c = 100$, the improvement is less marked. So, compared with the one based on cloud mask degradation, to increase spatial resolution seems to be the most appropriate strategy in order to improve cloud clearing products. Nevertheless results obtained at higher interpolation radii show that values of $R$, large enough, reduce the dependence on clear FOVs' distribution. This is a not unexpected result if we consider that greater values of $FC$ could be reached in the similar manner whenever as consequence of a more uniform distribution of clear FOVs (think, to the emerging of new clear isolated FOVs in large cloudy areas obtained, for example, by increasing spatial resolution), whether using, in the filtering step, interpolation radius $R$ large enough to include clear FOVs beyond the cloudy front. Obviously different will be, into the two cases, the quality of clear radiances retrieval and more and more different depending on channels energy level (Tramutoli et al., 1994, Cuomo et al., 1995).

3.2 Isolated clear FOVs

All filtering techniques, based on the nearest neighbor clear defined FOVs receive important benefits when clear isolated FOVs are detected into large cloudy areas as well in their broken edges. Cluster indexes (see Amato et al., 1991) usually have a quick decreasing giving evidence of a more uniform distribution of clear FOVs across the scene. We decided to investigate how, clear FOV characteristic (that we will call class $I$) to have around it, 0,1,...,8 cloudy FOVs, depends from the used spatial resolution and chosen cloud contamination thresholds. Results are resumed in Figures from 5a to 5c where the percentages (averaged over all passes) of new clear spots having $I = 0$, (which means growing inside an already clear area), 1, 2, ... 8 (which means isolated), cloudy spots around, are plotted varying linear spatial resolution from $L = 96$ to $L = 6$ and for $C_c = 75, 95, 100$. Must be noted that results obtained for isolated FOVs (class $I = 8$) at $L = 96$ are not significant being only 3 the total number of such a FOV over all passes and not more than 1 for a single pass. It is immediately evident from all plots that more than 70% of clear FOVs have at least 4 clear nearest FOVs, as soon as values of $L \leq 48$ are considered. Hence clear FOVs tend to clusterize more and more strongly as higher values of $C_c$ and higher spatial resolution are considered. In Figure 6 the total number $N8$ of isolated clear
FOVs (class I = 8) over all passes is shown as spatial resolution and cloud contamination thresholds are varied. Decimal logarithms of absolute numbers are reported to show as the increasing of clear isolated FOVs, as the spatial resolution increases, is relevant at every $C_c$ and specially fast at $L \leq 24\text{km}$.

4. CLEAR FOVS DISTRIBUTION

It is not useless to remember that in the Numerical Weather Forecast (NWF) context, cloud-cleared radiances are usually processed at about 120-km and a grid scale of about 100-km is the one used in a General Circulation Model (GCM). If we imagine our scene divided following a requested G-km resolution grid, the choice of new instrument designs, capable to improve clear soundings number and distribution, could be addressed by a different way. The penetration probability $P_G(\geq 1, L)$ to achieve at least one clear $L\times L$ sounding inside a $G\times G$ area could give us an independent indication on clear soundings distribution across the scene directly comparable with NWP requirements. The trade-off between this quantity and the chosen instrument design (in terms of FOV size and sampling density) strongly depends on the spatial distribution of clear areas, at different sizes, across the scene.

4.1 Random distribution test

If we assume, for example, that clear areas, at every size, are independently distributed across the scene (which means probabilities of individual FOVs covering clouds-free areas are independent), a Poisson statistics could be used to describe the probability $P_G(I, L)$ to find $I$ clear soundings of size $L\times L$ inside a $G\times G$ area. In this case

$$P_G(0, L) = e^{-m}$$

$$P_G(I, L) = \frac{m^I}{I!} e^{-m}$$

$$P_G(\geq 1, L) = 1 - P_G(0, L) = 1 - e^{-m}$$

where $m$ (which is function of the chosen values for $G$ and $L$) is the mean rate of all individual FOVs covering clouds-free areas. Under the above-mentioned assumptions, by increasing sampling density by a factor $k^2$ also the mean rate of individual success will increase by the same quantity. Under fixed $NE\Delta T$ constraint to increase sampling density we must increase the detector (or IFOV) area by the same factor. In these conditions will be:

$$P_G(\geq 1, kL) = 1 - P_G(0, kL) = 1 - e^{-mk^2}$$

which is an increasing function of $k$. So, under the hypothesis that the probabilities of individual FOVs covering clouds-free areas are independent and by using the practical
constraint of fixed $NE \Delta T$ for the detector, we obtain strong indications toward an increasing of sampling density also at lower spatial resolutions. Such a result is of evident contrast with suggestions, exactly opposite, we received from the analysis performed in Section 3. In order to verify if the assumption that probabilities of individual FOVs covering clouds-free areas are independent, we analyzed clouds-free area distribution over all AVHRR scenes described in section 2. All AVHRR scenes have been divided into rectangular boxes $34 \times 39$ AVHRR pixels large (which is the spacing between HIRS/2 FOVs) and soundings at different spatial resolution, simulated simply by dividing these boxes into 2, 3, 4, 6 and 9 equal parts. Cloud contamination masks have been computed, in the usual way, at the different spatial resolutions, expressed in fractions of the HIRS/2 box, $P = 1/2, 1/3, 1/4, 1/6, 1/9$, and for $C_c = 75, 95, 100$ values. Inside HIRS/2 boxes we computed the number $I$ of clear P-FOVs and built the quantity:

$$M_{C_c}(I, P) = 100 \cdot \frac{\text{number of HIRS/2 boxes with } I \text{ clear } P - \text{FOVs inside}}{\text{total number of HIRS/2 boxes}}$$

which, at high values of HIRS/2 boxes, represents the probability (expressed in percentages) to find one HIRS/2 box with $I$ clear FOVs inside. Also in this case we use the word "clear" for FOVs having clearness-indexes not lower than $C_c$. Corresponding values predicted by Poisson statistics:

$$P_{C_c}(I, P) = 100 \cdot \frac{m^I}{I!} e^{-m}$$

have been computed being $m$:

$$m = m_{C_c}(P) = \frac{\text{Num. of clear } P - \text{FOVs}}{\text{Num. of HIRS/2 boxes}}$$

known for each chosen threshold $C_c$ and simulated spatial resolution $P$. Values of $M_{C_c}(I, P)$ and $P_{C_c}(I, P)$ have been computed for all passes at each selected spatial resolution and cloud contamination threshold, corresponding to the previously listed $P$ and $C_c$ values. Figure 7 shows, for passes B and E, at $P = 1/9$ ($L = 1/3$) differences between measured $M_{C_c}(I, P)$ and predicted $P_{C_c}(I, P)$ for different $I$ and $C_c$ values. Looking at these results, (but similar results come from the analysis of the other passes) it is easy to see that differences are marked specially at the lower values of $I$ and are important (from our special point of view) also at the higher ones. Generally speaking observed clear soundings distribution exhibits less "isolated" clear soundings and more in cluster than foreseen from the Poisson statistics. The effect is more marked as far higher spatial resolution and higher (lower) values of $C_c$, at low (high) values of $I$, are involved. Our
**Fig. 4.** Percentage of remaining cloudy-area cloud cleared by interpolating with:

- a) $R=50\, \text{km}$;
- b) $R=100\, \text{km}$; $R=200\, \text{km}$

**Fig. 5.** Percentage of clear FOVs having I first neighbors cloudy when:

- a) $C_c=100$;
- b) $C_c=95$;
- c) $C_c=75$

**Fig. 6.** Number of clear holes at different spatial resolution $L$ and $C_c=75, 95$ and $100$.

**Fig. 7.** Observed and random distributed (Poisson) clear soundings. Differences in number of successes per HIRS box are given in percentage over all HIRS boxes at spatial resolution $L=1/3$ ($P=1/9$) and at various $C_c$. Passes B and E.
test (at more than 99.9% of confidence level after $\chi^2$ tests) clearly confirms as clear area distribution is not Poissonian at least for linear scales ranging from 10km to 40km.

4.2 Clustering, self-similarity, fractality

Results obtained by other studies (e.g. Yano et al., 1987, Duroure et al., 1990, Weger et al., 1993, Serio et al., 1995), performed on real cases too, suggest that clear areas tend to clusterize quite at every spatial scale. Scaling properties in cloudy (and raining) areas have been well studied in the past (see for example Lovejoy, 1987) and well defined fractal dimensions measured. The number $N(>L)$ of clear areas with linear dimension greater than $L$ is well then expected to follow a power law with a cumulative distribution function like $F(L) = 1 - C \cdot L^{-\alpha}$ (with $C$ constant), being:

$$1 - F(L) = C \cdot L^{-\alpha}$$

the probability to find one clear area with linear dimension greater than $L$. The number $N(>L)$ of clear areas with linear dimension greater than $L$ will follows a similar power law with the same exponent.

$$N(>L) \propto L^{-\alpha} \quad (1)$$

A self-similarity property like:

$$N(>rL) \propto r^{-\alpha}N(>L) \quad (2)$$

is well then expected to hold. Figure 8 presents, for pass B, log-log plots of the number of $L \times L$ boxes not completely cloudy against $L$ as obtained for different values of $C_c$. It follows quite perfectly a power law like:

![Fig. 8. Log-log plot of observed number of not completely cloudy areas of size LxL, against L. Results for pass B and Cc=75, 95 and 100.](image1)

![Fig. 9. Self similarity test with scaling factor r=2. Predicted against measured number of clear FOVs of linear size not less than L. Results for Pass B and Cc=75, 95 and 100.](image2)
\[ N(L) \propto L^{-D} \]

whose exponents, for different choices of \( C_e \), are indicated on the plots in Figure 8 as slopes. Fractal dimensions \( D \) of clear-sky class, have been computed for all passes and values comprise between 1.75 and 1.96 have been found as far as threshold \( C_e \) have been varied between 75 and 100. Both power law and self similarity properties have been verified by comparison with measured quantities for all the scenes and for different thresholds \( C_e \). In Figure 9 the quantity

\[ S(> L) \propto N(> rL) \cdot (L)^\alpha \]

predicted by using (2) at lower resolution \((r=2)\) is plotted against measured values \( N(> L) \) in the case of pass B. The analysis performed over all passes gave \( \alpha \) exponents between 1.4 and 1.8 for \( C_e = 75 \); between 2.2 and 2.5 for \( C_e = 95 \); between 2.7 and 3.3 for \( C_e = 100 \). The agreement between predicted and measured values is quite excellent and confirms that clear area distribution follows the power law (1) and exhibits self-similarity properties at least for the investigated scales between 3km and 100km.

5. CONCLUSIONS

Spatial distribution of clouds has been analyzed belong this paper in the case of six satellite passes over Europe. Its impact on processing techniques, based on clouds-free or weakly cloud-contaminated radiances, has been evaluated in terms of number and distribution of "clear-defined" soundings at spatial resolutions ranging between 3 and 100 km (linear). Tests on the fractional increase of clear detected area showed as results achievable by including weakly cloud-contaminated radiances at low spatial resolution, are roughly comparable with the ones achievable, for clouds-free or less clouds-contaminated radiances, by increasing spatial resolution. On the contrary a more uniform, clear soundings distribution across the scene, was found quite only by increasing spatial resolutions. These effects is more and more noticeable as far higher spatial resolutions are considered and lower values of the interpolation radius \( R \) are used in cloud-filtering processes. At \( R = 200 \) km no substantial improvements were found by increasing spatial resolution up to \( L = 24 \). This is a not unexpected result if we consider that the same area of the scene could be rebuilt as consequence of a more uniform distribution of clear soundings as well (but not with the same reliability) by using, in the filtering step, larger interpolation radii. These results have been supported on a more general basis by the analysis of spatial distribution of clouds-free areas. Our results clearly refuse random distribution of clouds-free areas. Clustering, self-similarity and fractality of clouds-free areas, already suggested by several independent studies, have been, on the contrary, confirmed at every scales, between 3km and 100km, belong this work. For this way superior performances of higher spatial
resolution instrument designs, compared with over-sampling strategies (at lower spatial resolution) can be easily predicted.

6. REFERENCES


Tramutoli,V., Serio,C., 1994. A radial predicting filter to recover clear-column infrared radiance field from satellite. In SPIE, 2315, 789-798


498
TECHNICAL PROCEEDINGS OF
THE NINTH INTERNATIONAL TOVS STUDY CONFERENCE

Igls, Austria

20-26 February 1997

Edited by
J R Eyre
Meteorological Office, Bracknell, U.K.

Published by
European Centre for Medium-range Weather Forecasts
Shinfield Park, Reading, RG2 9AX, U.K.

May 1997