Improved GOES Sounder Coverage of Cloud-Broken Data Fields

Alan E. Lipton

Optical Effects Division, Air Force Research Laboratory, Hanscom Air Force Base, Massachusetts

11 March 1997 and 18 June 1997

ABSTRACT

Geostationary satellite sounder radiances have typically been averaged over several individual fields of view before the radiances are used for retrieving thermodynamic profiles or for assimilation in weather prediction models. The purpose of the averaging is to compensate for data noise. Cloudy fields of view are excluded from averaging. In areas without a sufficient number of clear fields of view, complete profiles are not retrieved. Clouds thus cause gaps in sounder coverage. This note describes an automated method to select a set of averaging areas for a given field of sounder data, such that the gaps in coverage caused by clouds are as small and as few as possible. Test results are shown, indicating that the method can provide substantially better coverage than is obtained with a commonly used method.

1. Introduction

The sounder instrument employed on the Geostationary Operational Environmental Satellite (GOES) system has 1 visible-band channel and 18 infrared channels that are used to retrieve profiles of atmospheric temperature and water vapor, as well as other parameters. Each channel has a nominal field of view of 10 km at nadir (Menzel and Purdom 1994); however, data noise is great enough that it is preferable not to produce retrievals at every individual field of view (IFOV) but to average data from several neighboring IFOVs before undertaking the retrieval process.

The retrieval procedure that is operational at the National Environmental Satellite, Data, and Information Service (NESDIS) involves dividing a field of GOES sounder data into squares of 5 × 5 IFOVs and averaging data within each square to create sounding fields of view (SFOVs) at a spacing of about 50 km (Hayden and Schmit 1994). Each averaging square is centered a distance of five IFOVs from the centers of each neighboring square, so the SFOVs are all independent in the sense that there are no instances where an IFOV is included in more than one SFOV. Any IFOVs that have been determined to be cloudy are omitted from the averaging process. The operational cloud avoidance procedure is detailed in Hayden (1988). If an insufficient number of the IFOVs within the square are clear, no sounding is produced for the square (although a partial sounding, covering altitudes above cloud top, may be produced if the cloud cover is uniform). At the time the research reported in this note was performed, the required number of clear IFOVs was nine. In February 1997, the required number was changed to one, and the other numbers have been considered for use in the future (T. J. Schmidt 1997, personal communication). An alternative method to create SFOVs, which relies on cluster analysis of the satellite radiance data, has been proposed by Hillger and Purdom (1990) and demonstrated by Fuelberg et al. (1995).

A problem with procedures that divide sounder data fields into a grid of uniform rectangles (e.g., squares) before averaging is that they do not, in general, provide the most thorough possible SFOV coverage of each data field. This problem stems from the fact that the edges of the cloud-free areas of a sounder data field are not always positioned optimally with respect to the boundaries of the averaging rectangles. Consider the example where at least nine clear IFOVs must be averaged for the SFOV to be accepted. In this case, a 3 × 3 clear area within a predominantly cloudy region is big enough to allow a retrieval to be performed, provided that the clear area is entirely within the bounds of an averaging rectangle. If the area occurs along a boundary between rectangles, no rectangle would have the nine clear IFOVs required to make a retrieval.

This note describes a method to significantly improve SFOV coverage of cloud-broken data fields when data averaging is required, in relation to the coverage that can be obtained by dividing the data in some gridded fashion before performing averaging. The method was developed in the course of an investigation of convective development in a largely cloudy region (Lipton et al. 1995) using data from the predecessor of the current

Corresponding author address: Dr. Alan E. Lipton, Atmospheric and Environmental Research, Inc., 840 Memorial Drive, Cambridge, MA 02139.
E-mail: alipton@aer.com

GOES sounder, the VISSR (Visible and Infrared Spin Scan Radiometer) Atmospheric Sounder (Menzel et al. 1983). In that investigation, the SFOVs generated by dividing the data field into rectangles did not provide sufficient coverage of the case study domain to allow for meteorological analysis. The method described below yielded SFOV coverage that was substantially broader.

The proposed method relies on an editing procedure designed so that the spacing between SFOVs will be about the same as if the data field were divided into a regular grid of rectangles, but the SFOVs are located such that the sounder data field is covered as thoroughly as possible with SFOVs. The method is illustrated using data from *GOES-8.*

2. Method

Say we wish to create SFOVs by averaging data from $n \times m$ rectangles of IFOVs, where m is the number of lines of data and n is the number of elements along a line. Cloudy IFOVs are excluded from the averages, and SFOVs with fewer than some specified number (J) of cloud-free IFOVs are omitted. The usual way to proceed would be to divide the data field into a grid of $n \times m$ rectangles and average data within each rectangle. This procedure will be referred to as the grid method. It results in SFOVs uniformly spaced $n \times m$ apart, except where SFOVs are omitted due to clouds or other data problems. The NESDIS operational method is a particular example of this grid method.

The alternative method proposed here will be referred to as the edit method. It begins by forming SFOVs spaced 1×1 apart, using an $n \times m$ averaging rectangle. As with the grid method, cloudy IFOVs are excluded from averages, and SFOVs with fewer than J cloud-free IFOVs are omitted. Because the SFOV spacing is smaller than the averaging dimensions, there can be a large amount of overlap between the data averaged in one SFOV and the data in its neighbors. It would be possible to compute retrievals for every one of these SFOVs, but it would be a waste of computer resources because the SFOVs do not have independent information; they are largely redundant to each other. The edit method uses an editing process to selectively omit some of these SFOVs so that the ones retained have a minimum of redundancy.

To describe the editing process, it is helpful to define what is meant by ‘‘coverage’’ of a data field by SFOVs. A particular IFOV location in a data field is considered to be covered if it is within the $n \times m$ averaging area of any SFOV. So, each SFOV can cover an $n \times m$ area of the data field. This definition stems from the idea that each SFOV is representative of the $n \times m$ area from which it drew its data, and an IFOV will be represented in the retrieval results as long as it was within the averaging area of at least one SFOV.

The editing procedure is done in two parts. In the first part, each SFOV is examined and is removed from the dataset unless its removal would change the status of K or more cloud-free IFOVs from covered to isolated (not covered). In other words, as an SFOV is examined, each clear IFOV within its averaging rectangle is considered. Those are the IFOVs at risk of being left isolated if the examined SFOV is removed. A count is made of the number of clear IFOVs that are not covered by any SFOV other than the one being examined. If that count is less than K, the SFOV being examined is removed. For example, if there are three SFOVs adjacent along a line, and $n \times m$ is 3×3 or larger, the middle SFOV would be removed because every IFOV within its averaging area would also be within the averaging areas of one or both of the other two SFOVs.

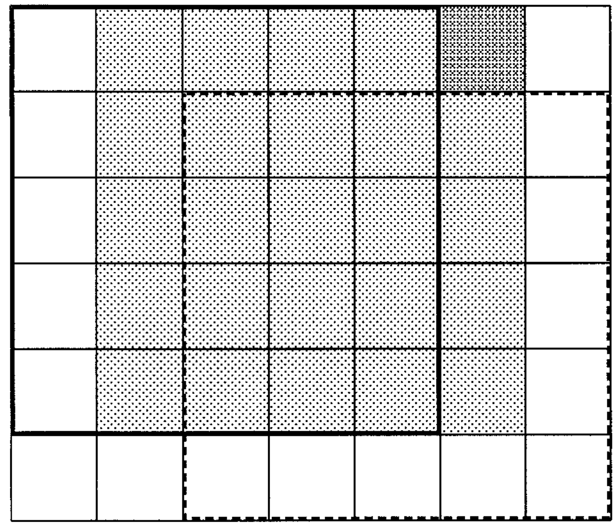

FIG. 1. An illustration of a possible configuration of SFOVs. Each box in the grid represents one IFOV. Three 5×5 SFOV averaging areas are shown: the middle one is shaded and the other two are outlined with heavy solid and dashed lines. The IFOV with the darker shading, which is assumed to be cloud free, is the only IFOV that lies inside the middle SFOV and outside the other two.

The parameter K will be referred to as the isolation limit and can be set to any value in the range $1 \leq K \leq nm$. If $K = 1$, a large proportion of the SFOVs will be removed, but there will be no reduction of the coverage of clear IFOVs. There may still be instances where two or more SFOVs are largely redundant to each other. An example of such a situation is illustrated in Fig. 1. The middle SFOV is almost entirely redundant to the the other two SFOVs, but there is one clear IFOV that is within the averaging area of the middle SFOV but within neither of the others.

A potential problem occurs when the selected value of K is larger than J since there will be some SFOVs that have no more than J clear IFOVs within their averaging areas. The editing process would remove all SFOVs with fewer than K clear IFOVs, overriding the chosen limit (J) on the number of clear IFOVs needed to create an SFOV. To avoid this problem, the editing

TABLE 1. The dates of the datasets processed and the coverage (as defined in the text) provided by the grid method and the edit method with $K = 4$.

	Proportional coverage (%)		
Date	Grid	$K = 4$	Increase
22 Jul 95	48	58	10
5 Aug 95	47	55	07
20 Aug 95	46	53	07
7 Sep 95	40	46	07
21 Sep 95	39	44	06
7 Oct 95	45	51	05
20 Oct 95	48	54	06
5 Nov 95	31	36	06
20 Nov 95	40	45	05
5 Dec 95	47	53	06
20 Dec 95	28	34	06
5 Jan 96	42	48	06

criterion is altered when $K > J$. In this case, an SFOV is removed unless its removal would convert J or more clear IFOVs and $K - J$ additional clear or cloudy IFOVs from covered to isolated. If $K = J$, for example, it would be possible for two SFOVs to have overlapping averaging areas and have some common clear IFOVs, as long as each SFOV had at least J clear IFOVs outside the area of overlap. When $K = nm$, SFOVs cannot overlap at all, meaning that there is no redundancy.

This first part of the editing procedure is done in a sequential, linear fashion: the editing starts in one corner of the data field and SFOVs are examined sequentially, moving through the data field to the opposite corner. The process is designed so that most of the retained SFOVs will occur at regular intervals of m lines from the top line of the dataset. This design feature results in contiguous, efficient coverage of any broad cloud-free portions of the data field.

Because of the sequential nature of the first part of the editing procedure, it can sometimes happen that an SFOV is removed unnecessarily due to overlap with another SFOV that will, itself, be removed later in the sequence of processing. To correct any unnecessary removals that occurred in the first part of the editing procedure, a second part is conducted. This second part consists of reexamining the field of SFOVs and restoring any SFOVs that no longer satisfy the removal criterion. Restoration of some SFOVs can, in turn, cause some other SFOVs to newly qualify for removal; therefore, the restoration process is followed by another process of removal. The restoration and removal processes are repeated, iteratively, until the processes cease to change the status of any SFOVs. In the tests discussed in the following section, the number of iterations needed was typically two and was never greater than five.

3. GOES test

The edit method for composing a field of SFOVs was tested on 12 GOES sounder datasets that each covered

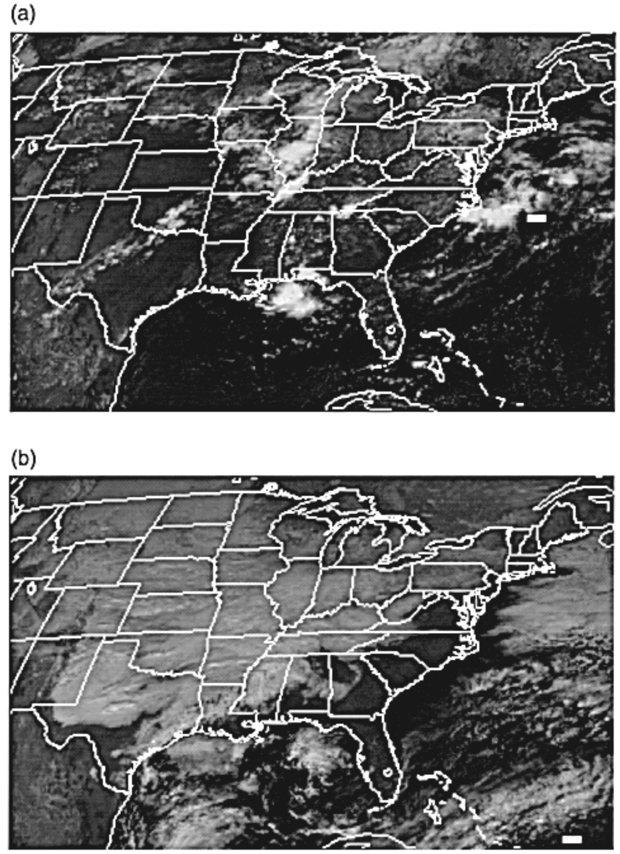

FIG. 2. Images created from visible channel data of the *GOES-8* sounder for (a) 22 July 1995 and (b) 5 January 1996. The white rectangles represent missing data.

the continental United States. The datasets were taken at 1750 UTC on or about the 5th and 20th of each month (depending on data availability), spanning from July 1995 to January 1996 (Table 1). The 1750 UTC time was chosen because the domain was well illuminated by the sun at that time, facilitating detection of clouds. Cloudy IFOVs were identified using a person–computer interactive cloud–clear discrimination program, which operated on images made with data from the visible channel and from the infrared channels at 3.7 and 11.0 μm. The data collection and processing were done with the Air Force Research Laboratory Interactive Meteorological System.

The edit method of generating SFOVs was applied to these data, testing various values of the isolation limit K. For comparison, SFOVs were also generated with the NESDIS version of the grid method. For both methods, the averaging areas were 5×5 in size, and any SFOV with fewer than nine clear IFOVs within its 5×5 area was omitted ($J = 9$). These criteria are the same as those used operationally at NESDIS at the time these tests were performed. For criteria where $J = 1$, the two methods provide equal SFOV coverage.

Results of the methods are plotted for the 22 July

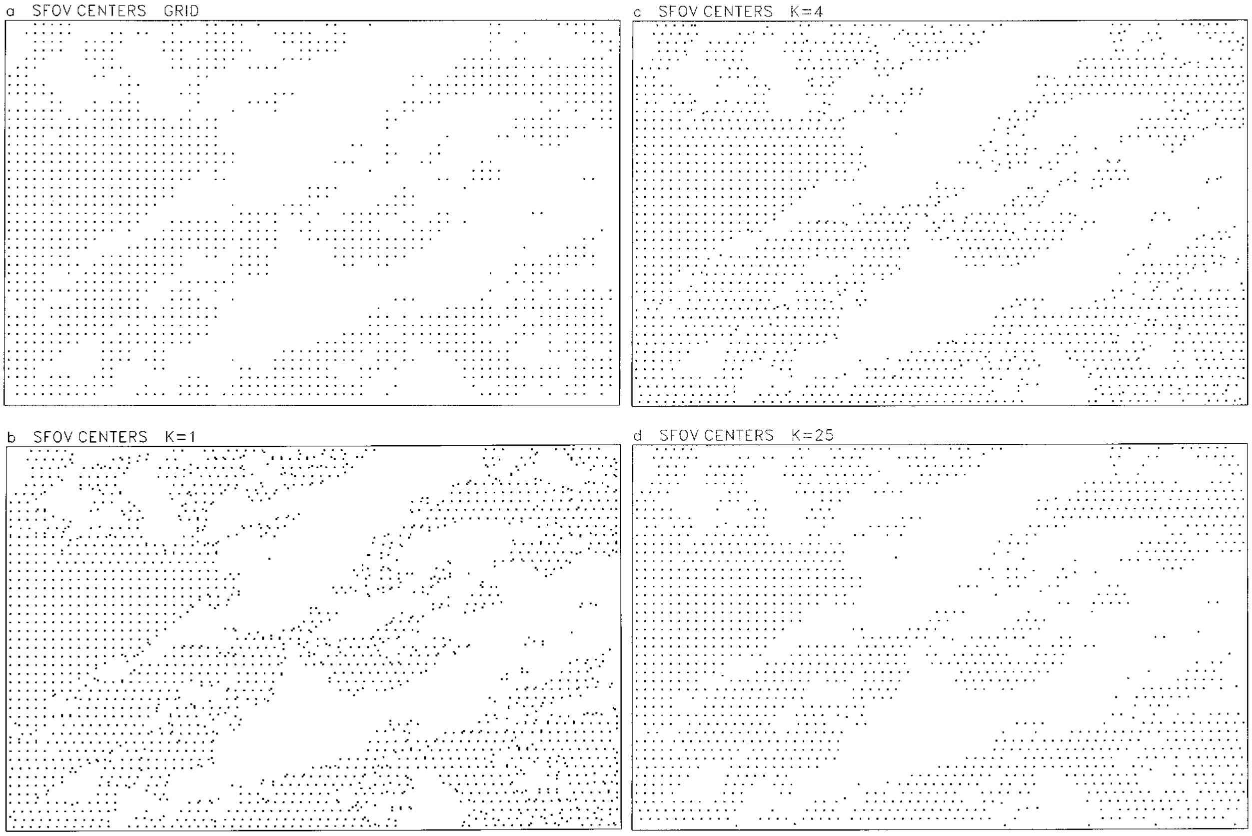

FIG. 3. Locations of the centers of SFOVs for 22 July 1995. The area shown is the same as in Fig. 2a. Each marker covers the IFOV at the center of the 5 × 5 averaging area of an SFOV. The four frames are for (a) the grid method and the edit method with isolation limit values of (b) 1, (c) 4, and (d) 25.

1995 dataset as an example. A visible-channel image for that dataset (Fig. 2a) is shown for reference. The grid method yielded a uniform pattern of SFOVs, except in cloudy areas, where there were gaps in the grid (Fig. 3a). The edit method, with $K = 1$, produced a somewhat irregular grid of SFOVs in the large cloud-free regions and many occurrences of closely spaced SFOVs along the edges of cloudy areas (Fig. 3b). The close spacing resulted from the phenomenon illustrated in Fig. 1. Plots are also shown for $K = 4$ and $K = 25$ (Figs. 3c, d). As the isolation limit increases, the crowding of SFOVs around cloudy areas is diminished. With $K = 25$, the SFOVs are no more crowded than with the grid method, but there are more SFOVs (1690 versus 1530) and the areas without SFOVs are smaller.

Figure 4 illustrates the data field coverage and the degree to which it can depend on the method used to create SFOVs. With each method there are large cloudy areas without coverage, but the coverage is considerably broader using the edit method (with $K = 1$) than with the grid method in this situation.

To assess the performance of the SFOV generation methods over a range of cloud conditions, statistics regarding the methods' performance were averaged over the 12 datasets (Fig. 5). The results indicate that the number of SFOVs generated and the coverage provided by those SFOVs are nearly constant for values of the isolation limit higher than 9 ($=J$). For lower values of K, both statistics vary substantially with K. A change in K from 1 to 4 has a relatively large impact on the number of SFOVs generated but has a more modest impact on the coverage. In comparison with the grid method, the edit method provided significantly greater coverage even when K was set to its largest value (25). The average improvement for the test cases was a factor of 8% for $K = 25$ and 20% for $K = 1$.

It is generally undesirable to have a large number of SFOVs because each SFOV represents some cost in computer resources. An example of such resources is the time required to perform a thermodynamic retrieval or to assimilate the radiance data in a numerical weather prediction model. On the other hand, it is desirable to maximize the SFOV coverage of data fields so that meteorologically important features are represented as well as possible. A way to balance these two considerations is to consider the number of SFOVs in relation to the amount of coverage provided. Such a cost–benefit ratio was computed for the test results by dividing the number

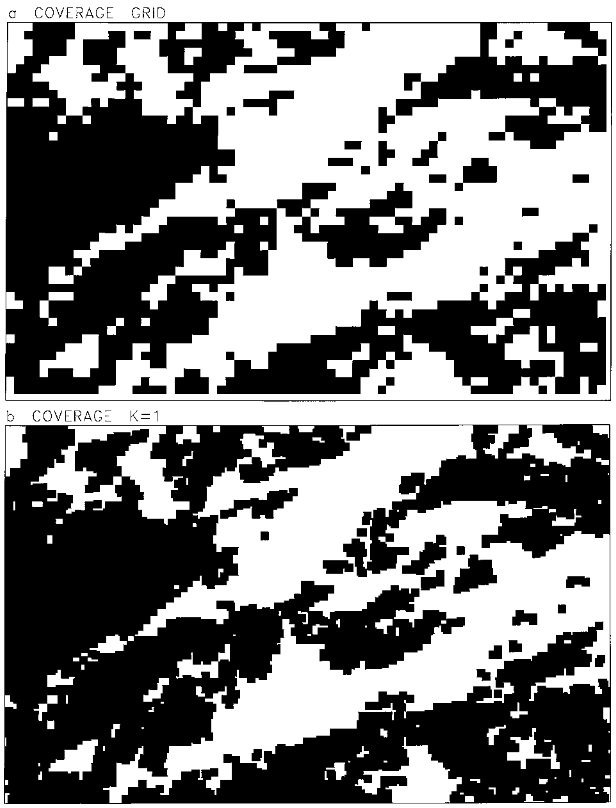

FIG. 4. Coverage of the 22 July 1995 dataset by SFOVs according to the definition of coverage given in the text. Black areas are covered and white areas are not. The area shown is the same as in Fig. 2a. The frames are for (a) the grid method and (b) the edit method with an isolation limit value of 1.

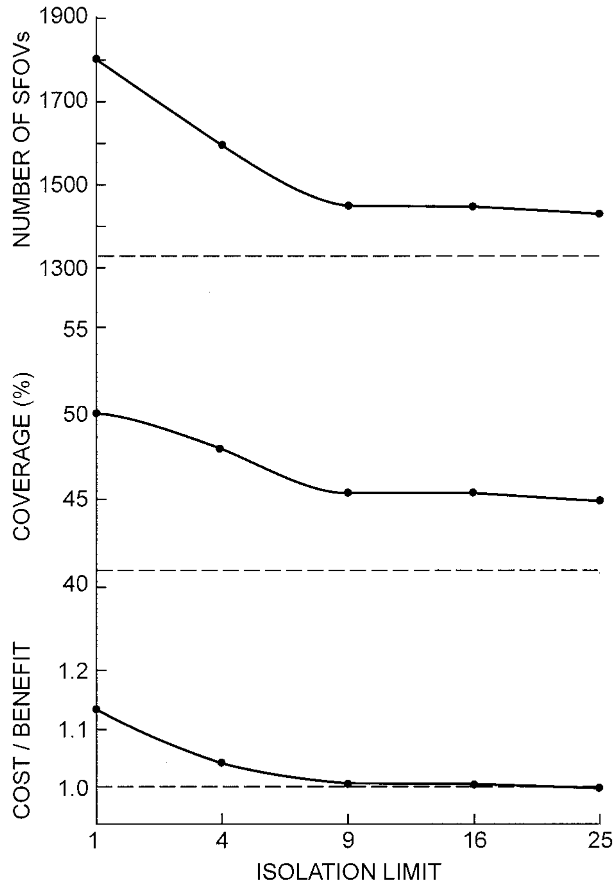

FIG. 5. Performance statistics, averaged over the 12 datasets, for the methods of generating SFOVs. The curves are for the edit method, plotted as a function of the isolation limit K (scaled to the square root of K). The dashed lines mark the statistics for the grid method. At the top is the number of SFOVs within a dataset, in the middle is the proportion of the dataset area covered by at least one SFOV, and the bottom is the cost–benefit ratio defined in the text.

of SFOVs by the number of covered IFOVs and normalizing so that the ratio is 1 when each SFOV covers 25 IFOVs. By definition, a cost–benefit ratio of 1 occurs for the grid method and for the edit method when $K = 25$. The ratio changes little for values of K ranging from 9 to 25 but increases sharply for lower values of K (Fig. 5). These ratios could be used to choose an appropriate value of K for any particular application of the sounder data.

The cost–benefit ratio is defined in terms of computational costs, but it also carries information regarding the amount of redundancy among the SFOVs in a dataset. The isolation limit controls how much redundancy is allowed at any particular location, while the cost–benefit ratio describes how much redundancy actually occurred overall. With a cost–benefit ratio of 2, for example, the SFOVs could be considered to each contain half-units of information.

It is expected that the edit method would have its greatest impact on SFOV coverage in situations where cloudy areas are relatively small and scattered, such as in conditions with patches of convective clouds. This is expected because the method relies on optimally positioning SFOVs in relation to edges of cloudy areas, so the method will be most advantageous when there are many edges. The data in Table 1 provide evidence in support of this expectation. The four datasets in which the edit method had the largest positive impact on SFOV coverage were the four that were taken in the summer season (the first four), when clouds tend to be more convective than in the fall or winter. The images in Fig. 2 illustrate the differences in cloud field structure between the first and last datasets listed in Table 1. The summer dataset in Fig. 2a primarily had convective clouds, while the winter dataset in Fig. 2b had broad areas of stratiform cloud cover.

4. Conclusions

The proposed ''edit'' method for creating and editing sounding fields of view can be used to enhance coverage of usable satellite sounding data in fields containing clouds when averages of cloud-free data samples are required. Substantial increases in coverage are possible, relative to the coverage provided by averaging data over a regular grid of rectangular areas. The method offers some flexibility, through the isolation limit parameter, to balance the interest of maximizing field coverage against a possible requirement to constrain costs involved in thermodynamic parameter retrieval or assimilation. An isolation limit of about four would likely be appropriate for most GOES sounder applications, assuming that the standard averaging area for sounding is 5×5 IFOVs.

While data coverage is an important factor in the use of satellite sounder data, data quality is of even greater importance. It would be straightforward to alter the grid method to account for data quality. Any available quality measure could be incorporated into the procedure, including any generated as a byproduct of the cloud–clear discrimination procedure (Hayden and Schmit 1994). The edit method could be initially executed using only SFOVs that pass the strictest quality criteria to create a preliminary field of edited SFOVs. Then the SFOVs with lower quality ratings could be selectively added. An SFOV of suspect quality would be added only if it would extend the coverage enough that K clear IFOVs would be converted from isolated to covered, where K is the isolation limit.

An additional line of possible future research would be to perform retrievals and analyses from sets of SFOVs generated by each of the methods discussed here. Comparisons of the analyses with ground truth data would make it possible to measure the importance of SFOV coverage to the quality of analyses derived from sounder data. Differences in the sizes of areas that lack SFOV coverage could be related to the errors involved in interpolating across those areas. As with other satellite sounding evaluation studies, it would be difficult to obtain appropriate ground truth data.

Acknowledgments. The author thanks Frank H. Ruggiero and Donald W. Hillger for reviewing a draft of the manuscript and the anonymous reviewers for helpful suggestions toward revisions. Thanks go also to Anthony S. Lisa and John D. Pickle of Atmospheric and Environmental Research, Inc. (AER), Michael K. Griffin of AFRL/GPOR for assistance with GOES data aquisition, and Marjorie Paulsen for assistance with figure preparation. The interactive cloud–clear discrimination program was written by Daniel C. Peduzzi and Gary B. Gustatson of AER.

REFERENCES

Fuelberg, H. E., P. A. Rao, and D. W. Hillger, 1995: Clustering of satellite sounding radiances to investigate intense low-level humidity gradients. *J. Appl. Meteor.,* **34,** 1525–1535.

Hayden, C. M., 1988: GOES-VAS simultaneous temperature–moisture retrieval algorithm. *J. Appl. Meteor.,* **27,** 705–733.

——, and T. J. Schmit, 1994: *GOES-I* temperature moisture retrievals and associated gradient wind estimates. Preprints, *Seventh Conf. on Satellite Meteorology and Oceanography,* Monterey, CA, Amer. Meteor. Soc., 477–480.

Hillger, D. W., and J. F. W. Purdom, 1990: Clustering of satellite sounding radiances to enhance mesoscale meteorological retrievals. *J. Appl. Meteor.,* **29,** 1344–1351.

Lipton, A. E., G. D. Modica, S. T. Heckman, and A. J. Jackson, 1995: Satellite-model coupled analysis of convective potential in Florida with VAS water vapor and surface temperature data. *Mon. Wea. Rev.,* **123,** 3292–3304.

Menzel, W. P., and J. F. W. Purdom, 1994: Introducing *GOES-I*: The first of a new generation of Geostationary Operational Environmental Satellites. *Bull. Amer. Meteor. Soc.,* **75,** 757–781.

——, W. L. Smith, G. S. Wade, L. D. Herman, and C. M. Hayden, 1983: Atmospheric sounding from a geostationary satellite. *Appl. Opt.,* **22,** 2686–2689.