An Algorithm for the Reduction of Speckling and Striping Artifacts in OCTS GAC Data

Bryan A. Franz
John E. O'Reilly

1 August 2001

I. Introduction and Recommendation

The purpose of this analysis is to develop an algorithm to reduce speckling and striping noise in the OCTS GAC processing. The primary sources of this noise are the artifacts associated with cloud edges and the residual striping effects due to incomplete correction of the relative detector responsivities. Cloud-edge artifacts can result in highly deviant localized chlorophyll-a retrievals (e.g., 64 mg m^{-3} in oligotrophic waters), which will never be averaged-out in the binned products. Residual striping effects can result in systematic scan-to-scan variability in the chlorophyll retrievals of 100% or more. While it is not difficult to develop or simply apply existing image processing techniques to reduce these effects in the derived products, it would be preferable to minimize the problem prior to entry into the atmospheric correction process, so as not to magnify the error through removal of erroneous aerosol path radiances.

Since this data cleaning algorithm will be applied on the entire OCTS GAC mission archive, the method must be efficient in both processing time and memory usage. The algorithm will be operated within the Multi-Sensor Level-1 to Level-2 software (MSL12). MSL12 already includes a capability for buffering of multiple lines of input TOA radiance data with associated flags, path geometries, and pre-computable atmospheric quantities (e.g., cloud flags, Rayleigh path radiances), and it provides a mechanism for data filtering which allows for the replacement of TOA radiances in each pixel of a scan line with new values derived through evaluation of a user-specified filter function over an m-pixel by n-line sliding window.

It is recommended that we apply a 3x3 square dilation to the cloud and high radiance flags, followed by a 5x5 diamond interquartile-mean filter to the Rayleigh-subtracted TOA radiances in all OCTS bands. A description of these concepts and the basis for recommendation is provided in the sections which follow.

II. Co-registration Error

A major source of image speckling is associated with cloud edges. The cloud flag algorithm is essentially a threshold on the TOA reflectance measured in band 8. These cloud edge artifacts can occur on occasion for any sensor, when the radiance at the top of the atmosphere is just below the cloud threshold. The problem occurs more often for OCTS, due in part to the co-registration problem. Since band 8 is near one end of the relatively large OCTS focal plane and band 5 is near the opposite end of the focal plane, the two bands will observe the same spot on the earth at a slightly different path angle. For a sub-pixel cloud or cloud-edge at altitude, it is quite possible for the band 8 radiance to be contaminated in one pixel, while band 5 is contaminated by the same cloud, but in a neighboring pixel. One solution to this problem is to dilate the cloud mask such that, wherever a pixel is identified as a cloud in band 8, the 8 neighboring pixels are also flagged as clouds. Unfortunately, since we are working with GAC data which has been sub-sampled from the full resolution by every 6th pixel and every 5th line, it is still possible to have cloud contamination in the visible bands which is never observed in the band 8 GAC samples. However, as shown in Figure 1, the dilation removes the majority of cloud-edge artifacts.

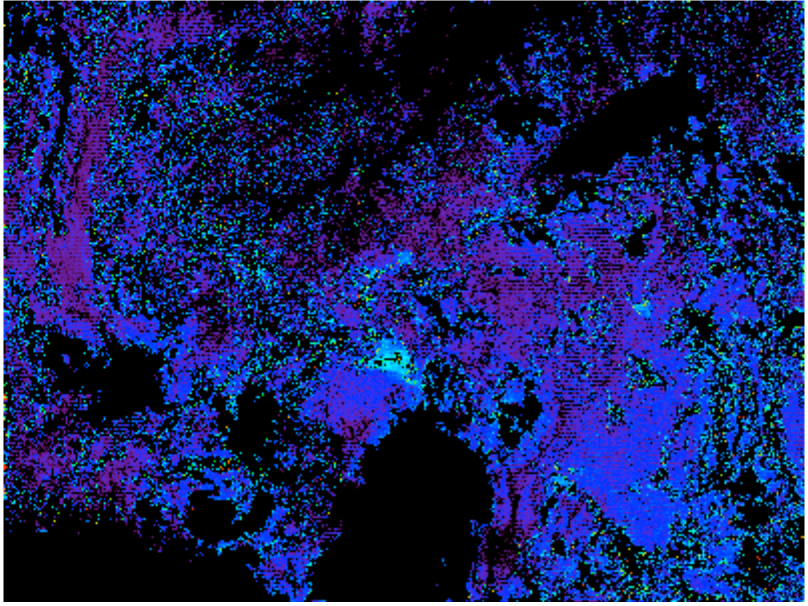
It is also suggested that we dilate the high radiance flag. For OCTS processing within MSL12, the high radiance flag is only set when one or more bands is saturated. In general, this should only occur over thick clouds and some land masses. It is possible that some of the deviant radiances are due to straylight effects, so this is just an added precaution to avoid the most extreme light levels.

The dilation approach is a very conservative scheme which will result in a significant reduction in the number of valid pixels available for higher-level processing. In many cases, these lost pixels may be perfectly valid. One alternative approach is to use a statistical measure, such as a spatial median filter, to reject outliers in each band. Such an approach will generally work, except in cases where there is so much variability that a nominal data range can not be determined statistically. The median filter is ideal for removing occasional cloud-edge artifacts, such as those in Figure 1b, and it will be discussed further in the next section.

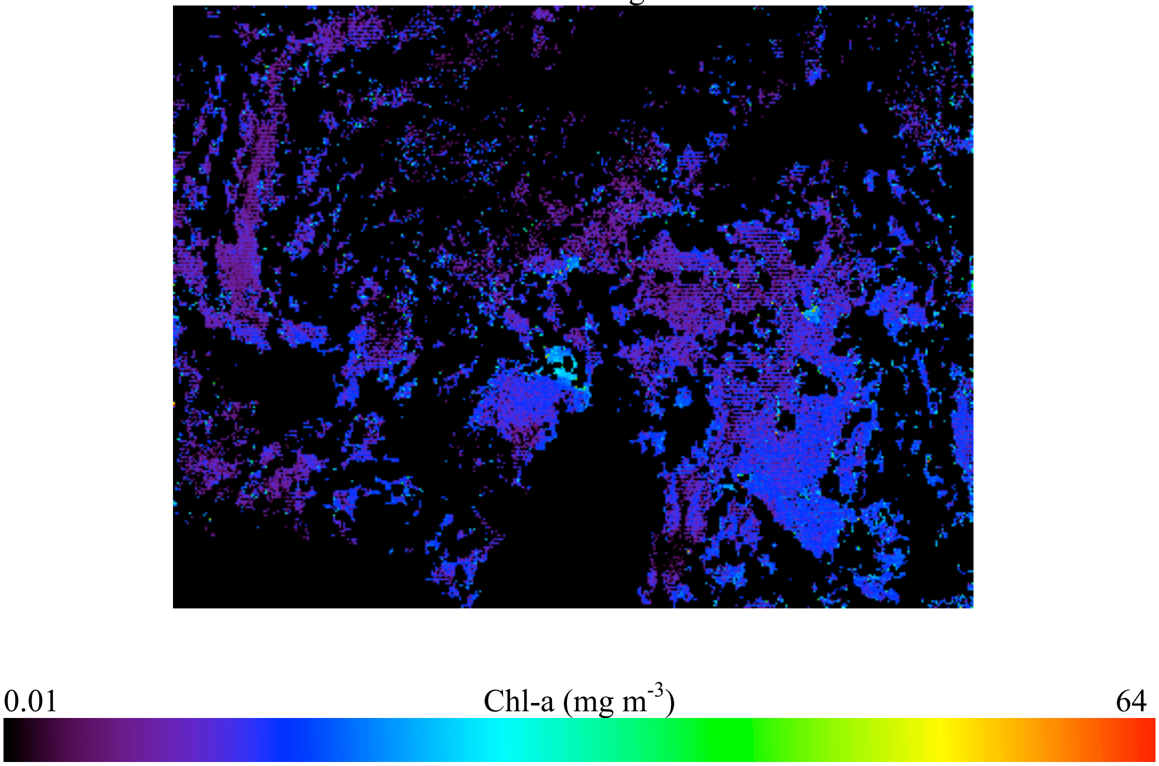
Perhaps the best approach to address the effects of misregistration around clouds would be to develop a multi-band cloud detection algorithm, or to use a test of spatial variability as an additional cloud flagging criterion. These ideas could be pursued in a future refinement of the OCTS GAC processing.

Figure 1: Chlorophyll-a concentration, before and after cloud flag dilation.

a. Original image



b. After cloud flag dilation



III. Residual Striping and Speckling

In the image of Figure 1b, it is still possible to find some highly deviant, single-pixel chlorophyll-a retrievals. A simple scheme to remove these artifacts would be to apply a median filter over the chlorophyll field, replacing each pixel value with the median in a 3-pixel by 3-line box, centered on that pixel. However, we would probably have to do the same for the K490 field, the aerosol optical thicknesses, and any other derived quantities of interest. Instead, it would be preferable to correct the problem in the top-of-atmosphere (TOA) radiances. Fixing the TOA radiances will also serve to stabilize the atmospheric correction process, which might otherwise be confused into selecting erroneous aerosol models or deriving anomalous aerosol concentrations, thereby magnifying the problem.

Horizontal striping is also evident in Figure 1b. The striping arises from the fact that each OCTS band is divided into ten individual detectors, and each detector has a slightly different responsivity. The effect is to produce a systematic, line-to-line variation in the TOA radiances. The instrument calibration includes a correction for relative detector response which largely removes this variability, however, the correction is imperfect and residual horizontal striping remains.

The characteristics of the striping in the GAC data are modified by the co-registration process and the GAC sub-sampling process. The GAC data is actually 0.85-km LAC data which has been sub-sampled by every 6th pixel and every 5th line. In the absence of co-registration, this would imply that the same pair of detectors is selected for every two lines, and thus we would expect to see a systematic difference in the TOA radiances between the odd-numbered lines and the even-numbered lines. However, the co-registration process reorganizes the detector samples such that, along any line of Level-1A data, the actual detector number will change periodically. The end result is that areas within the GAC scene will be represented by a constant detector pair, but that detector pair will change periodically as we move across the scan and along the swath. Thus, depending on the relative gain error between a detector pair and the influence of a given band on a particular retrieved field, we can expect to see regions of odd-even striping within the final Level-2 products.

To remove both the residual striping effects and the single-pixel artifacts, we have tested a variety of median filters which modify the TOA radiances. The median is used rather than the mean, as it provides a simple and systematic approach to avoiding highly deviant values. The filtering mechanism within MSL12 provides an m-pixel by n-line sliding window, within which we can make any variety of computations and replace the central pixel TOA radiance or other pre-computed quantities (e.g., flags) as needed to ensure a smooth field or to eliminate highly deviant values. We can use this framework to replace the TOA radiance in each pixel of each band with the median TOA radiance over an m x n window, centered on that pixel. However, it would be better to refine this slightly by first reducing the TOA radiance (L_t) by the Rayleigh path radiance (L_r), to minimize systematic spatial variability that might skew the median. The new TOA radiance, L_t', can then be computed as:

$$Lt' = Lr + \text{median}(L_t - L_r)_{mxn}$$

At this point, it has not been suggested as to what the window size should be. If the window is too small, we will not have a sufficient sample size to establish a nominal value, and thus we will not be able to avoid deviant values. If the window size is too large, we might be combining signals from regions of differing bio-optical populations and effectively reducing resolution in the Level-2 products. In deciding the window size, we must also consider that the number of available samples will be reduced by clouds, glint, land, and other masking conditions which we do not want to include in the median. To maintain statistical validity, we should fail any pixels for which the filter window is not at least 50% filled with unmasked observations. This has the effect of slightly dilating the masking conditions, and that dilation will increase with the filter window size.

Figure 2 shows a 4-to-1 zoomed region of a Level-2 GAC chlorophyll-a field where significant striping artifacts and single-pixel deviations are evident. The chlorophyll range of the image has been stretched between 0.0 and 0.2 mg/m^3 to better show these problems. The associated plot of Figure 2 provides a vertical slice through the image, indicating line-to-line variations of as much as 100% in these low chlorophyll waters. Figure 3 shows the same image after application of a 3x3 median filter on the Rayleigh-subtracted TOA radiances in all bands. The simple 3x3 median does a good job of removing the anomalous single-pixel values, but the residual striping is only partially corrected.

By careful inspection of Figure 3, and comparison with Figure 2, it can be seen that the position of the high-valued stripes and low-valued stripes has been shifted by one line. This effect is easily explained by the weighting characteristics of a square 3x3 median filter. Consider a radiance field of minimal natural variability or sensor noise, but with alternating lines of high and low values. If we center our filter window on one of the high-valued stripes, we will compute the median for a dataset of 6 low values and 3 high values, which will cause us to replace the high-valued center pixel with a low value.

To make the median filter more effective at removing alternating striping artifacts, we need to modify the approach to give more equality to the weighting of odd and even lines. This can be achieved by introducing the concept of a filter window kernel, which indicates which pixels within the window will be considered in computing, in this case, the median. Consider these two examples of a 5x5 filtering window, where the value of 1 indicates that the pixel at that location will contribute to the median:

Square 5x5						Diamond 5x5				
1	1	1	1	1		0	0	1	0	0
1	1	1	1	1		0	1	1	1	0
1	1	1	1	1		1	1	1	1	1
1	1	1	1	1		0	1	1	1	0
1	1	1	1	1		0	0	1	0	0

For a pixel centered on a high-valued line, the square 5x5 filter would compute the median over 15 high values and 10 low values. So, in the absence of random noise or natural structure, the square kernel would actually have no effect, as it would strongly favor the high values when centered on a high-valued pixel, and visa versa. Under the same conditions, the diamond kernel would provide 7 high values and 6 low values to the median computation, so the 5x5 diamond kernel would also favor the high values when centered on a high-valued pixel, but the ratio of highs to lows would be much closer to unity. Thus, the diamond filter will be more sensitive to the natural variability then will the 5x5 square kernel. In fact, Figure 4 illustrates that the high-to-low ratio achieved by a 5x5 diamond filter is equivalent to that of a much larger 13x13 square filter window. It should also be noted that, for the diamond kernel, only about 50% of the pixels within the window contribute to the median, so the effects of mask dilation and resolution degradation are reduced relative to the equivalent square filter.

The relative effect of the square and diamond window kernels can be seen by comparison of the chlorophyll-a images in Figures 5 and 6. Figure 5 shows the result of applying a 5x5 square filter to the Rayleigh-subtracted TOA radiances. The striping has been further reduced relative to the 3x3 square filter, but the artifacts are still clearly evident. In Figure 6, the 5x5 diamond filter has been applied, and we begin to see much more natural variability in the chlorophyll-a retrievals. There is still some evidence of striping artifacts in the 5x5 diamond-filtered data, but the visual effect has largely been removed relative to the square-filtered data, and a higher effective resolution has been maintained. However, the median actually just rearranges the stripes to make the error less coherent. This effect is quite evident in the region running vertically, just left of center in Figures 5 and 6, where the stripes have been transformed into a checkered pattern.

At this point we have only considered median filtering, primarily because the median is resistant to highly deviant values within the filtering window. It could be argued that the mean is a more appropriate measure for determining the "true" value when working with striped data, since we have know way to know whether the higher-valued stripe or the lower-valued stripe is closer to reality. Thus, we might consider taking the average over odd and even scans. Figure 7 shows the effect on chlorophyll retrievals of applying a 5x5 diamond mean filter to the Rayleigh-subtracted TOA radiances of the scene from Figure 2. The checkered pattern has now been mitigated, but the resulting image illustrates a major problem with simple averaging, as all spatial variability, natural or sensor-induced, has been reduced. The effect is to blur the image. In addition, some of the single-pixel high values from Figure 2, such as those near the left edge of the image, have now been spread over multiple pixels.

The blurring effect of the averaging filter can be minimized by reducing the number of pixels which contribute to the mean. To retain the outlier resistant benefits of the median filter and reduce the blurring effect, we could consider a multiple-pass approach in which we first median filter the data and then apply a smaller averaging filter to the result. An alternative approach which can be operated in a single pass is to compute the average of the inter-quartile range. This has the benefit of avoiding the highest and lowest values in the population, while averaging over any striping artifacts. Figure 8 shows the effect on chlorophyll of applying a 5x5-diamond inter-quartile-mean filter to the Rayleigh-subtracted TOA radiances. The resulting image shows minimal striping artifacts and no obviously deviant pixels. Furthermore, the dynamic range of the apparent natural variability retained by the inter-quartile mean is similar to that of Figure 6, the 5x5 diamond median.

Figure 2: OCTS GAC chlorophyll-a, no filtering applied.

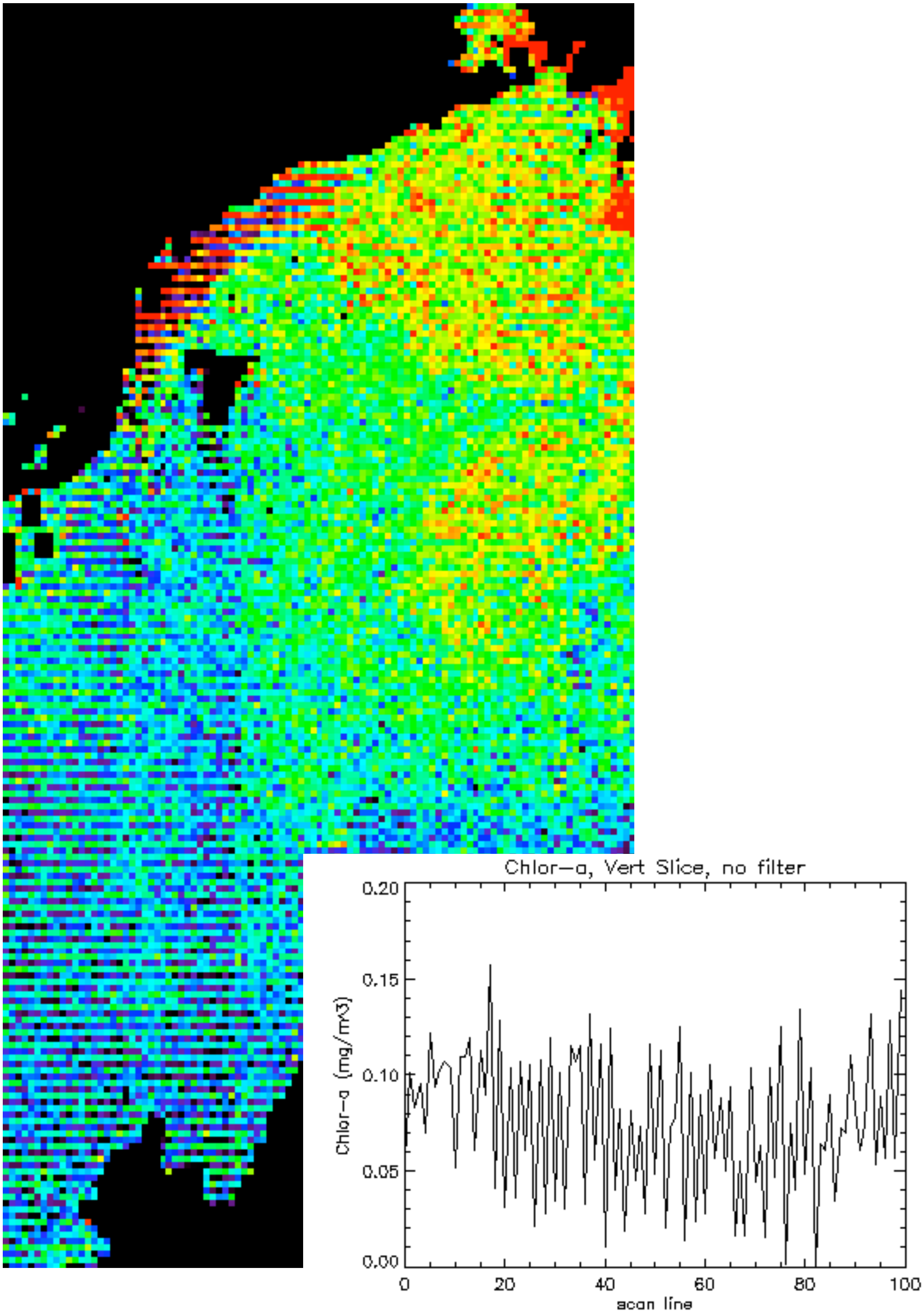


Figure 3: OCTS GAC chlorophyll-a, 3x3 square median filter applied.

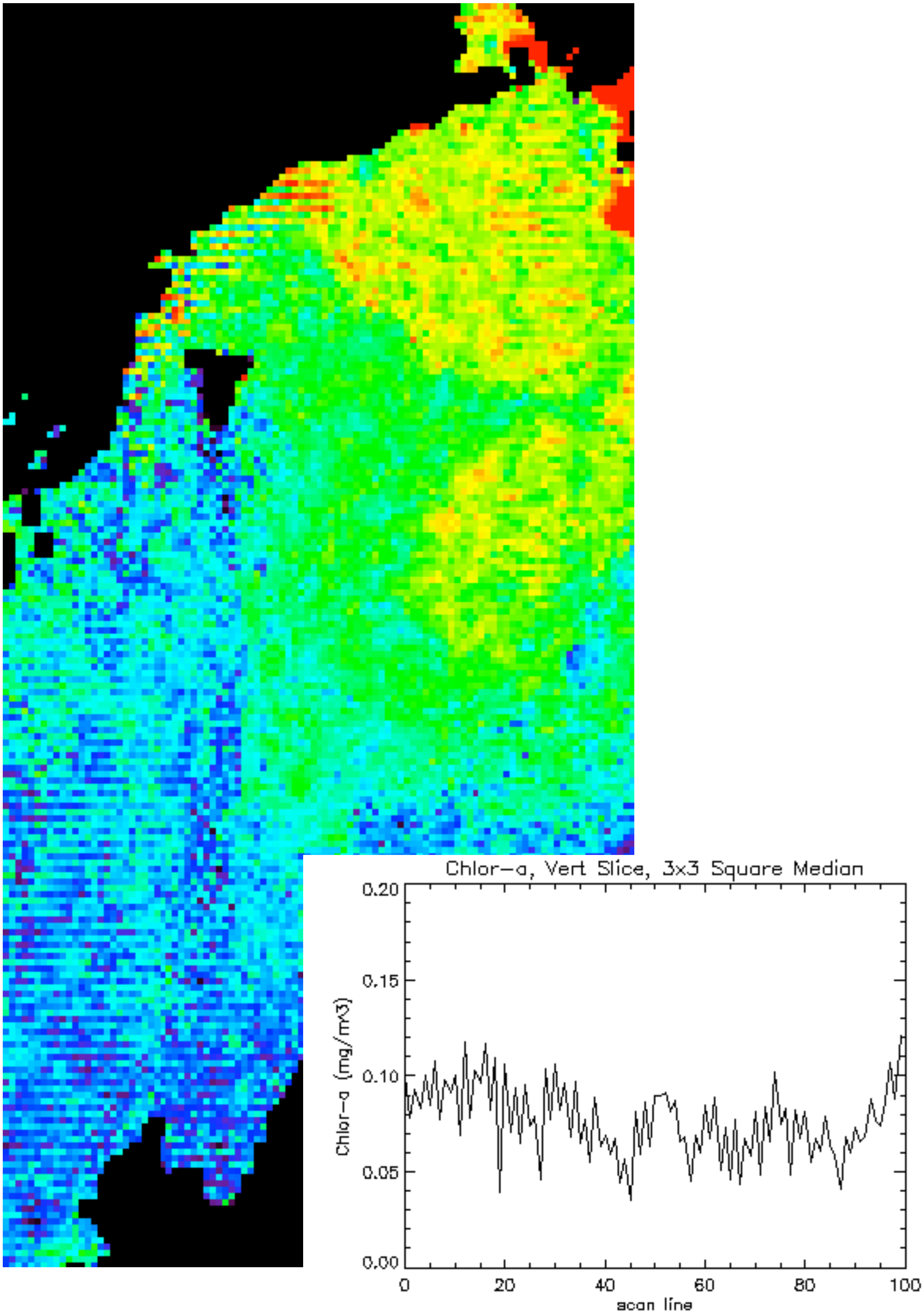


Figure 4: Comparison of high-to-low ratio for square and diamond filter kernels.

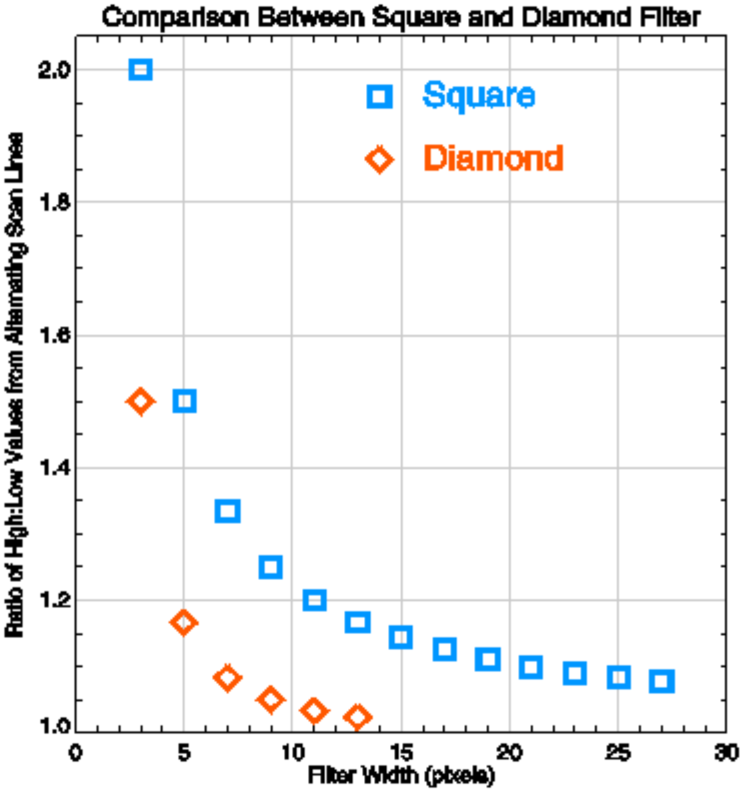


Figure 5: OCTS GAC chlorophyll-a, 5x5 square median filter applied.

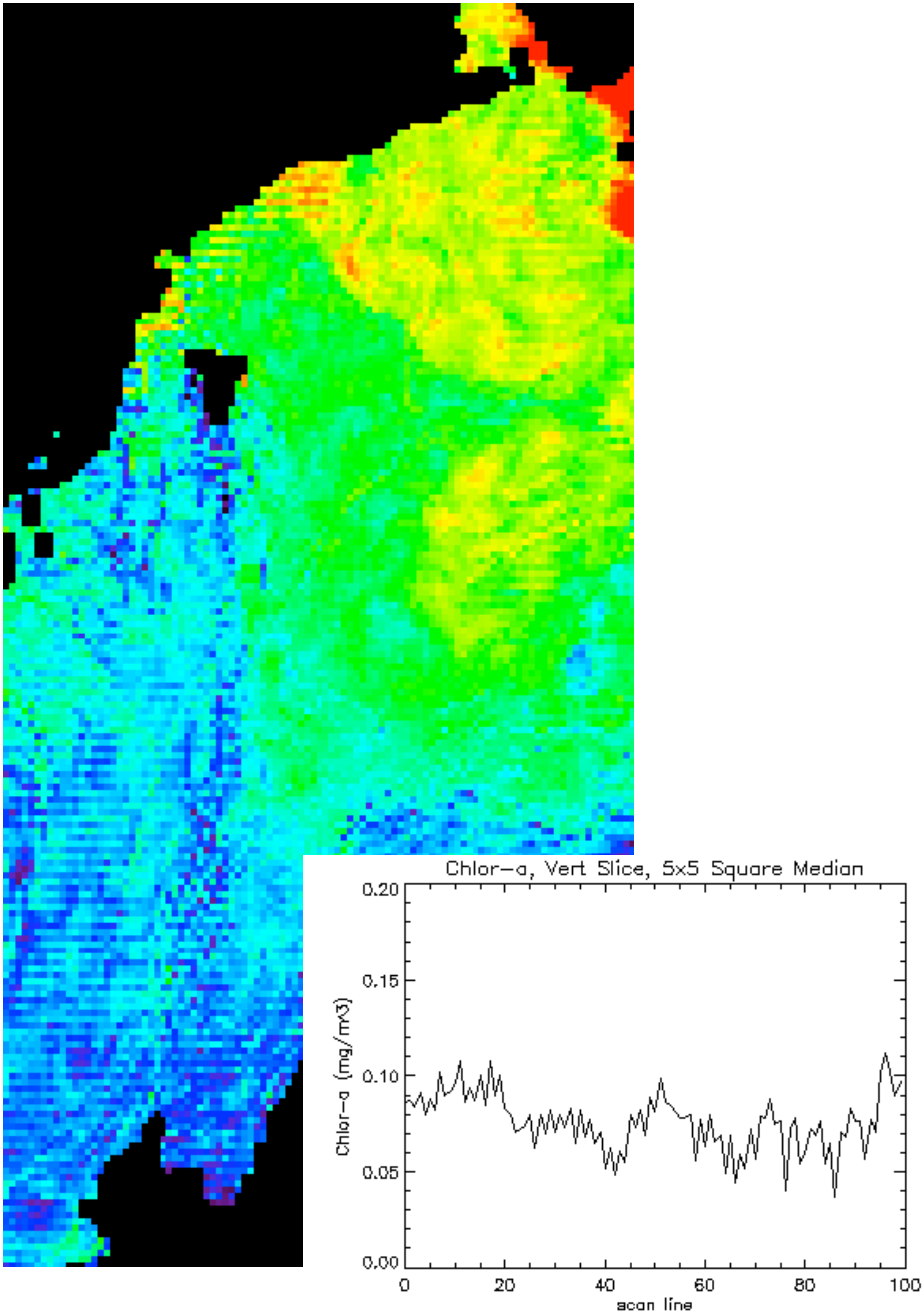


Figure 6: OCTS GAC chlorophyll-a, 5x5 diamond median filter applied.

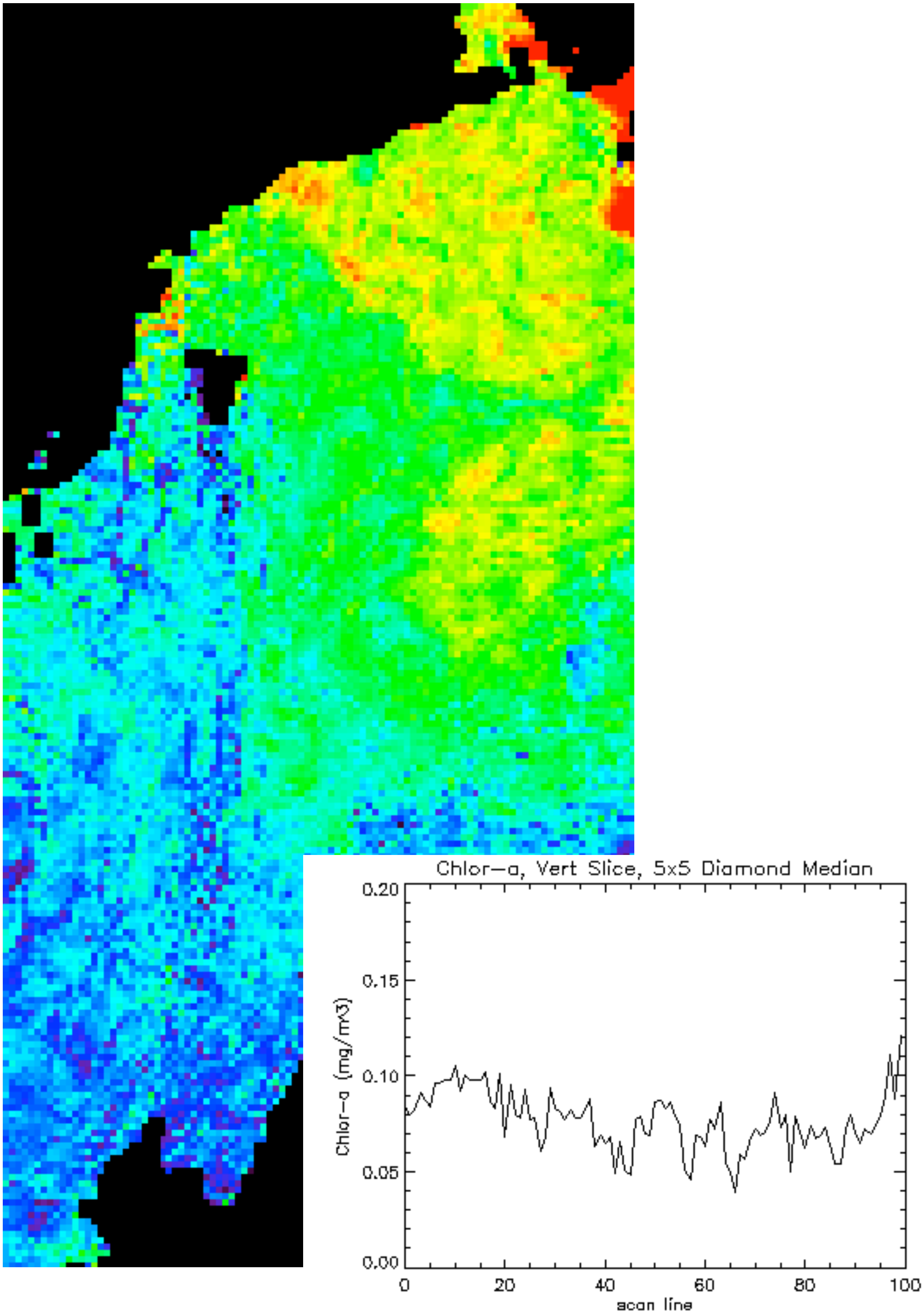


Figure 7: OCTS GAC chlorophyll-a, 5x5 diamond mean filter applied.

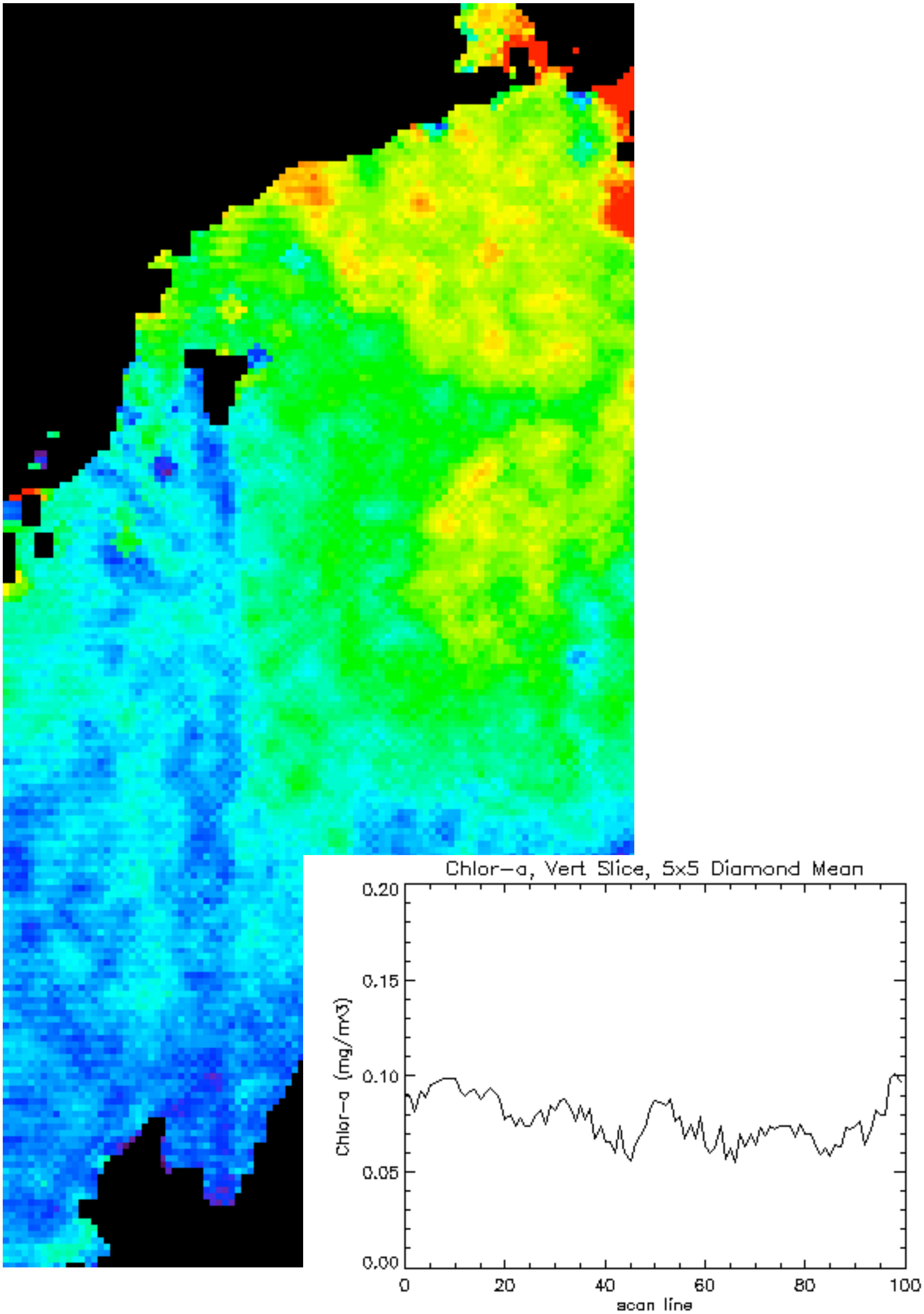
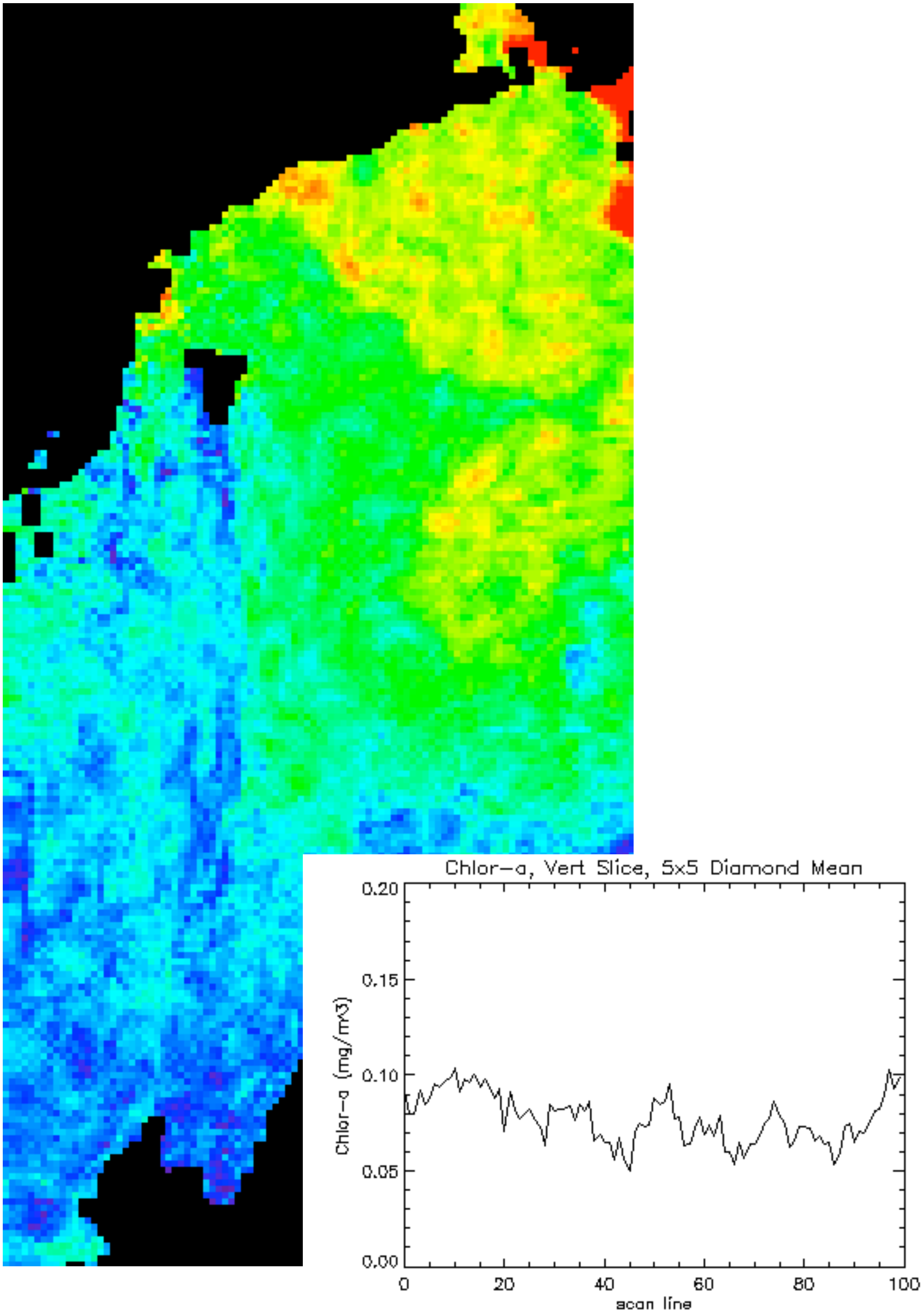


Figure 8: OCTS GAC chlorophyll-a, 5x5 diamond inter-quartile mean filter applied.



IV. Additional Analyses

The various filter forms discussed above have been applied to several full scenes of OCTS GAC data during Level-1A to Level-2 processing. The resulting chlorophyll-a retrievals, as well as angstrom coefficients, epsilon, aerosol optical thicknesses, and normalized water-leaving radiances are posted here:

http://orca.gsfc.nasa.gov/octs/gac/filtering.html

Global results: *to be added*.

V. Conclusions

The primary focus of the NASA/NASDA collaboration on OCTS GAC reprocessing is to produce a set of time and space binned OCTS data which is as compatible as possible with current SeaWiFS products. In that regard, the first priority should be to minimize the possibility that highly deviant values will be included in the spatial and temporal averages, since such anomalous values may never average-out over the OCTS mission life-span. Thus, it is recommended that we take a conservative approach, applying both a 3x3 square dilation of the cloud and high radiance flags, followed by a band-by-band statistical outlier rejection scheme. The suggested method is to replace any previously unmasked TOA radiance with the sum of Rayleigh path radiance and the mean of the Rayleigh-subtracted TOA radiances within the interquartile range of 5x5 diamond spatial area. By limiting our consideration to the data in the interquartile range, we effectively drop the most deviant high and low values.

Another goal of the OCTS GAC reprocessing is to produce the highest quality Level-2 products possible. In the context of this analysis, that means reducing striping artifacts. The choice of a diamond window kernel, as originally suggested by J. O'Reilly, has the desirable property of near equal weighting of odd and even scans, which makes it highly compatible with the characteristics of OCTS GAC striping . The mean is suggested as the best measure of truth within the window, as we have no way of determining which stripe value (high or low) is closer to reality. Finally, since the data distribution going into the mean is limited to the interquartile range, the likelihood that highly deviant values will be included in the average is very low.