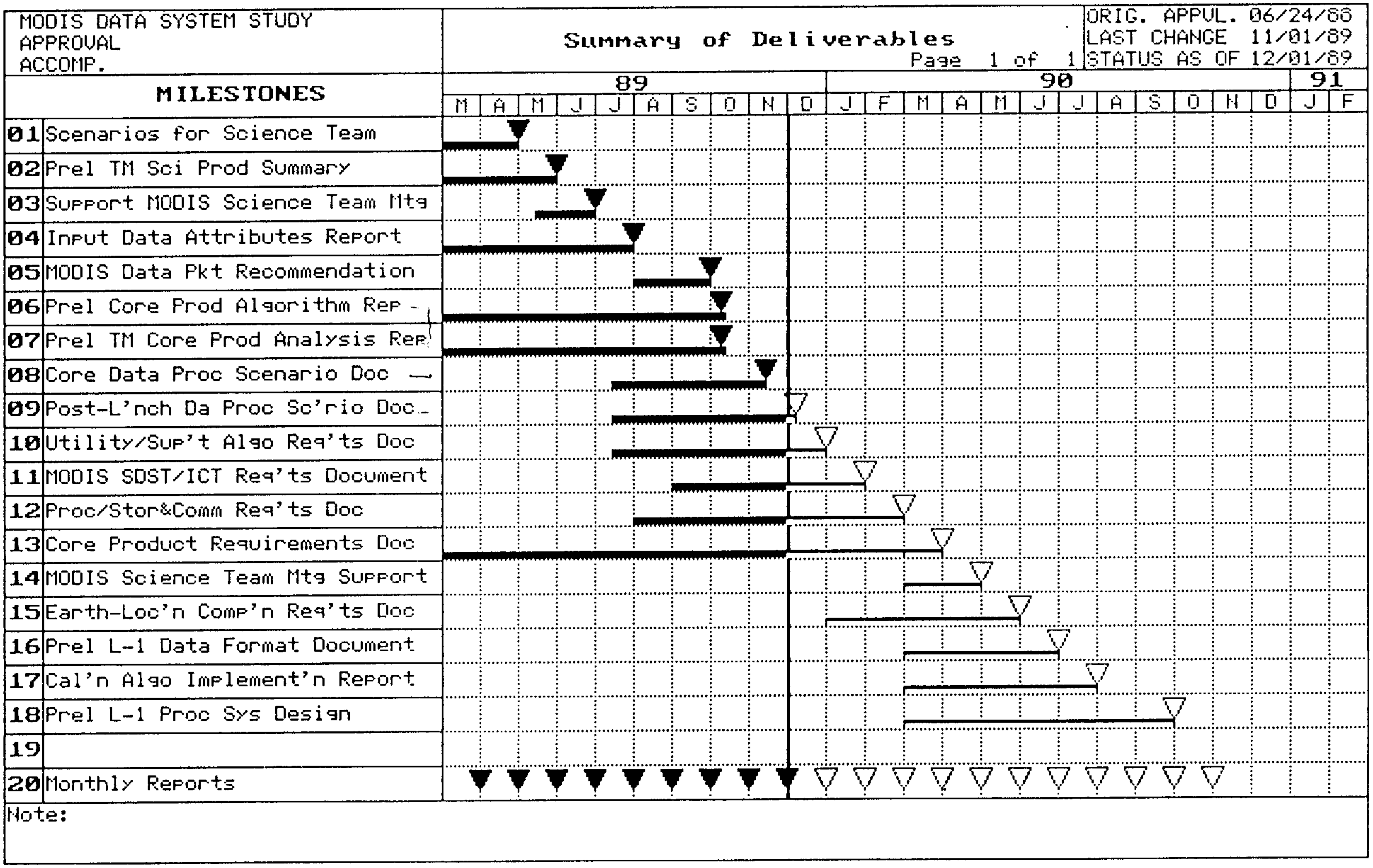
MODIS DATA STUDY TEAM PRESENTATION

December 1, 1989

AGENDA

1. Updated MODIS Data Study Team Milestone Chart
2. Accuracy Requirements for External Data Sets for Ocean Data Products: Wind Speeds and Total Ozone
3. Reprocessing of Data in the Post-Launch Processing Scenario
4. Atmospheric Correction Utility Algorithm for Terrestrial Products
5. Time and Space Averaging Utility Algorithm
6. Product Integration Issues



MODIS DATA SYSTEM STUDY
APPROVAL
ACCOMP.
Summary of Deliverables
Page 1 of 1
ORIG. APPVL. 06/24/88
LAST CHANGE 11/01/89
STATUS AS OF 12/01/89
MILESTONES
89
90
91
M A M J J A S O N D J F M A M J J A S O N D J F
01 Scenarios for Science Team
02 Prel TM Sci Prod Summary
03 Support MODIS Science Team Mtg
04 Input Data Attributes Report
05 MODIS Data Pkt Recommendation
06 Prel Core Prod Algorithm Rep
07 Prel TM Core Prod Analysis Rep
08 Core Data Proc Scenario Doc
09 Post-L'nch Da Proc Sc'rio Doc
10 Utility/Sup't Algo Req'ts Doc
11 MODIS SDST/ICT Req'ts Document
12 Proc/Stor&Comm Req'ts Doc
13 Core Product Requirements Doc
14 MODIS Science Team Mtg Support
15 Earth-Loc'n Comp'n Req'ts Doc
16 Prel L-1 Data Format Document
17 Cal'n Algo Implement'n Report
18 Prel L-1 Proc Sys Design
19
20 Monthly Reports
Note:

Accuracy Requirements for External Data Sets
For Ocean Data Products:
Wind Speeds and Total Ozone

Three data sets external to the MODIS processing environment have been identified as required in order to produce MODIS ocean data products. These are 1) atmospheric pressure to determine the Rayleigh scattering contribution to the total radiance received by the sensor, 2) wind speeds to determine the sun glitter and sea foam contributions to the total radiance, and 3) ozone optical depths, to determine the absorption of ozone. These data sets are required for the atmospheric correction of ocean data products, which allows the retrieval of water-leaving radiances. Water-leaving radiances, in turn, are required to obtain many ocean data products, including chlorophyll.

What is required is to know not only what data are needed, but from where the data will be obtained and at what resolution. This report examines the data requirements for ocean color for MODIS, and identifies sources for these data in the MODIS era. Final authority for deciding these accuracy requirements and external data set resolutions rests, of course, with the MODIS Science Team. This report is intended to be a preliminary analysis to facilitate understanding of the utility/support algorithm and processing/storage requirements.

Atmospheric Pressure

Atmospheric pressure requirements were presented in the MODIS Data Study Team Presentation of November 17, 1989. Briefly, we found that an error of $\pm$ 1 mb was sufficient to meet MODIS-N minimum detectable radiance requirements except under extremely unfavorable viewing geometries, where an accuracy of 0.5 mb was required. The 1 mb value will likely be obtainable for the MODIS era from National Meteorological Center (NMC) synoptic analyses (Wayman Baker, personal communication). These results were discussed with Wayne Esaias at GSFC, who pointed out that AIRS or MODIS oxygen soundings may provide an alternate source of atmospheric pressure data. However, such a procedure is not refined at this time and requires further research.

Wind Speed

Wind speed is required to determine the sun glitter and sea foam contribution to the total radiance received by the sensor. Wind speeds are known to relate to sun glitter by the theory of Cox and Munk (1954). The theory relating wind speed to sea foam is not presently known. Thus this discussion will concentrate on sun glitter, and results from a conversation with Dr. Wayne Esaias.

For CZCS processing, a 6 m s^{-1} global averaged wind speed was assumed. Then from orbital geometries and Cox and Munk (1954) wave slope distribution probability, a threshold sun glitter radiance was set, above which no processing was performed. This conservative scheme rejected for processing pixels that probably were useable, but at least provided a "hands-off", automated procedure for estimating sun glitter. MODIS, with its higher radiometric sensitivity, will require a better approach.

Sun glitter is considered a problem only between ≈ 40° N and 40° S latitudes. But it may be a considerable problem between these latitudes. In fact, in areas of extreme sun glitter, such as very near the solar specular point, imagery will be too contaminated to allow retrieval of any radiance signal from the water (Gordon, 1989), even with accurate wind speed information. It is in areas of mild sun glitter that the correction will be most useful.

Wayne Esaias suggested that wind speeds at a spatial resolution of ≈ 50 km (0.5° by 0.5° latitude/longitude grids) were necessary for MODIS ocean processing. This compares with the 0.5 mb requirement for atmospheric pressure. However, as with atmospheric pressure, this resolution is probably not achievable with reliability from NMC synoptic analyses. Unlike atmospheric pressure, however, there may exist in the MODIS era remote sensors capable of obtaining surface wind speeds. These include SCAN-SCATT and passive microwave sensors. If SCAN-SCATT flies, a spatial resolution of 50 km is easily attainable with accuracy, although a question may arise as to the timeliness of the data, i.e., whether the data will be available within 48 hours to meet the MODIS Level 2 processing requirement. However, if these data are used for synoptic forecasts, as they probably will, the NMC's 6-hour requirement is much more stringent than MODIS', and suggests the availability of such data.

If passive microwave sensors are used to determine wind speeds, the timeliness requirement may not be met. Such sensors typically determine wind speeds by tracking observable clouds, a method that requires human intervention. In such case, Dr. Esaias suggested using forecast winds, running Level 2 processing for a quick-look output, then going back and re-running Level 2 a week or two later using the updated wind speeds. Such a scenario would likely produce very good results, but has a major impact on EosDIS, requiring two Level 2 runs and a doubling of storage requirements (quick-look Level 2 would be held in storage for use in comparison studies with the final Level 2 product).

Figure 1 illustrates the effect of sun glitter on the total radiance received by the sensor at a low chlorophyll concentration. For this illustration we used the bio-optical model of Sathyendranath and Platt (1988) to determine the optical properties of ocean water containing various concentrations of chlorophyll, and the radiative transfer model of Gordon et al. (1988) to estimate the radiance emanating from the water. Sun glitter reflectance was 0.0132, which resulted from solar and spacecraft zenith angles of 30°, and a wind speed of 14 m s^{-1}. These values were taken from Viollier, et al. (1980). All other assumptions were as stated for the analysis of atmospheric pressure in the November 17, 1989 MODIS Data Study Team Report.

Using the proposed atmospheric correction algorithm of Gordon (1989), the effect of uncompensated sun glitter on aerosol radiances was simulated (Figure 2). Virtually all of the uncompensated sun glitter was taken as aerosol radiance, and also changed the aerosol radiance spectral distribution. This change in the spectral distribution of aerosols affected the spectral normalized water-leaving radiance retrieved by the algorithm (Figure 3), decreasing that retrieved near 440 nm and increasing that received near 560 nm. Since the ratio of these values is used to compute chlorophyll, an increase in estimated chlorophyll is expected, and was in fact obtained (Table 1).

Table 1. Percent error in chlorophyll retrieval for uncompensated sun glitter at a wind speed of 14 m s^{-1}.

Chl	Pct Error
0.04	+30.9%
0.15	+43.5%
1.90	+48.7%
4.80	+58.5%
6.60	+56.8%

Clearly, uncompensated sun glitter will cause serious errors in chlorophyll retrievals by MODIS, and a sun glitter correction will help produce a higher accuracy product.

Ozone

Ozone optical thickness is required to correct incoming solar irradiance and outgoing water-leaving radiance. Ozone optical thickness has a distinct spectral influence, and thus affects most of the ocean core data products. The following discussion results from a conversation with Dr. Wayne Esaias.

Current CZCS processing uses low resolution TOMS data for ozone scale heights, with an accuracy of ± 25 Dobson units (DBU; ozone scale height x 1000). MODIS will require higher accuracy, on the order ± 10 Dobson units, which is the approximate accuracy of current remote sensing methods. Maximum ozone gradients are of the order 20-30 DBU per 100 km (Dr. Lanning Penn, personal communication). These represent gradients near the Antarctic ozone hole, so expected gradients under most operating conditions can be expected to be much less. It should be noted that the highest ozone gradients occur in the winter hemisphere, where MODIS observations of ocean color will be limited by insufficient light. To obtain an accuracy of ≈ 10 DBU will require observations on the order of 50 km (about 0.5° grids).

Figure 4 illustrates the problem in ocean color observations due to ozone. Plotted are normalized water-leaving radiances at low (≈0.03 mg m^{-3}), medium (≈1.0 mg m^{-3}), and high (≈6.0 mg m^{-3}) chlorophyll concentrations, along with ozone absorption coefficients (x 10). Since peak ozone absorption is near 560 nm and very low absorption is near 440 nm, errors in estimating the ozone optical thickness will dramatically affect the radiance ratio $L_w(443)/L_w(550)$ used to compute remotely sensed chlorophyll.

Again using the bio-optical model of Sathyendranath and Platt (1988) to determine the optical properties of ocean water containing various concentrations of chlorophyll, and the radiative transfer model of Gordon et al. (1988) to estimate the radiance emanating from the water, we simulated the effects of incorrect ozone optical thickness on the retrieval of chlorophyll by MODIS, using the proposed atmospheric correction algorithm of Gordon (1989). All other assumptions were as stated for the analysis of atmospheric pressure in the November 17, 1989 MODIS Data Study Team Report.

The simulations were performed for a solar zenith angle of 60°, and spacecraft zenith angle of 50°, a solar azimuth angle of 0°, and a spacecraft azimuth angle of 120°. This viewing geometry represents an extreme; a tilt of 20° and viewing at the edge of the swath for northern hemisphere winter at about 40° N latitude. This unfavorable viewing geometry can best assess the errors due to an incorrect ozone optical thickness.

These errors are shown in Figures 5 and 6 for ozone Dobson unit errors of ± 25 DBU, the accuracy value used in CZCS processing. Errors in ozone affect the computation of aerosols, because the algorithm uses the radiance at 665 nm, where ozone is strongly absorbing, to estimate aerosol characteristics. Thus an error in ozone "looks" to the algorithm as aerosol. This results in an incorrect Angstrom exponent determination, and produces an error in aerosols and water-leaving radiances that propagates into the short wavelengths, where ozone is only minimally absorbing.

These errors are reflected in the normalized water-leaving radiances (Figure 6) as oscillating spectrally between overestimates and underestimates. The crossover point in these errors (where the error is zero, at ≈ 500 nm) occurs where the ozone absorption coefficient is the same as at 665 nm, where the aerosol estimates were generated. Thus the aerosol is incorrect, and the water-leaving radiance attempts to compensate. By attempting to compensate, the water-leaving radiance is incorrect both where ozone is more strongly absorbing than at 665 nm, and where it is less, but the error reverses in sign for these two cases. Also plotted in Figure 6 is the minimum detectable radiance for MODIS-N, from the updated specifications report of September, 1989. Errors in water-leaving radiance exceed these minimum detectable radiances by as much as an order of magnitude for ± 25 DBU.

Errors in chlorophyll retrievals due to these ozone errors are shown in Table 2.

Table 2. Percent errors in chlorophyll due to over- and underestimating ozone scale height by ± 25 DBU.

Chl	+ 25 DBU	- 25 DBU
0.04	-26.7%	+31.5%
0.15	-27.9%	+34.8%
1.90	-22.8%	+24.4%
4.80	-19.5%	+19.7%
6.60	-16.7%	+16.6%

At an error of ± 10 DBU, errors in chlorophyll retrievals reduce to a maximum of ± 13% at this extremely unfavorable viewing geometry, now within the accuracy of the CZCS and probably very near the accuracy of MODIS.

Such an accuracy of ± 10 DBU is obtainable from current remote sensors, but the question arises, can remote sensing data be available for MODIS processing within 24 hours?

If so, then data from other sensors would be ideal for the MODIS ocean color data processing scenario. However, if not, then two other alternatives must be examined. These are: 1) obtaining ozone information from MODIS itself, and 2) processing on MODIS using forecast or previous day ozone estimates.

If ozone scale heights are obtained from the 9.37 μm band on MODIS-N, timeliness requirements for ocean data processing will be met easily, but it is likely that accuracies will be reduced from that obtainable from other sensors, according to Wayne Esaias. This option must stand as a secondary alternative. A third alternative is to use forecasts or previous day ozone values, perform Level 2 processing, and then update a week or two later using corrected ozone values. As for wind speeds, this option will produce high quality ocean color data but has a major impact on the processing scenario.

References

Cox, C. and W. Munk, 1954. Measurement of the roughness of the sea surface from photographs of the sun's glitter. J. Mar. Res. 44: 838-850.

Gordon, H.R., 1989. Algorithm development for ocean observation with EOS/MODIS. MODIS Science Team Member Report.

Gordon, H.R., O.B. Brown, R.H. Evans, J.W. Brown, R.C. Smith, K.S. Baker, and D.K. Clark, 1988. A semianalytic radiance model of ocean color. Journal of Geophysical Research 93: 10909-10924.

Sathyendranath, S. and T. Platt, 1988. The spectral irradiance field at the surface and in the interior of the ocean: a model for applications in oceanography and remote sensing. Journal of Geophysical Research 93: 9270-9280.

Viollier, M., D. Tanre, and P.Y. Deschamps, 1980. An algorithm for remote sensing of water color from space. Boundary-Layer Meteorology 18: 247-267.

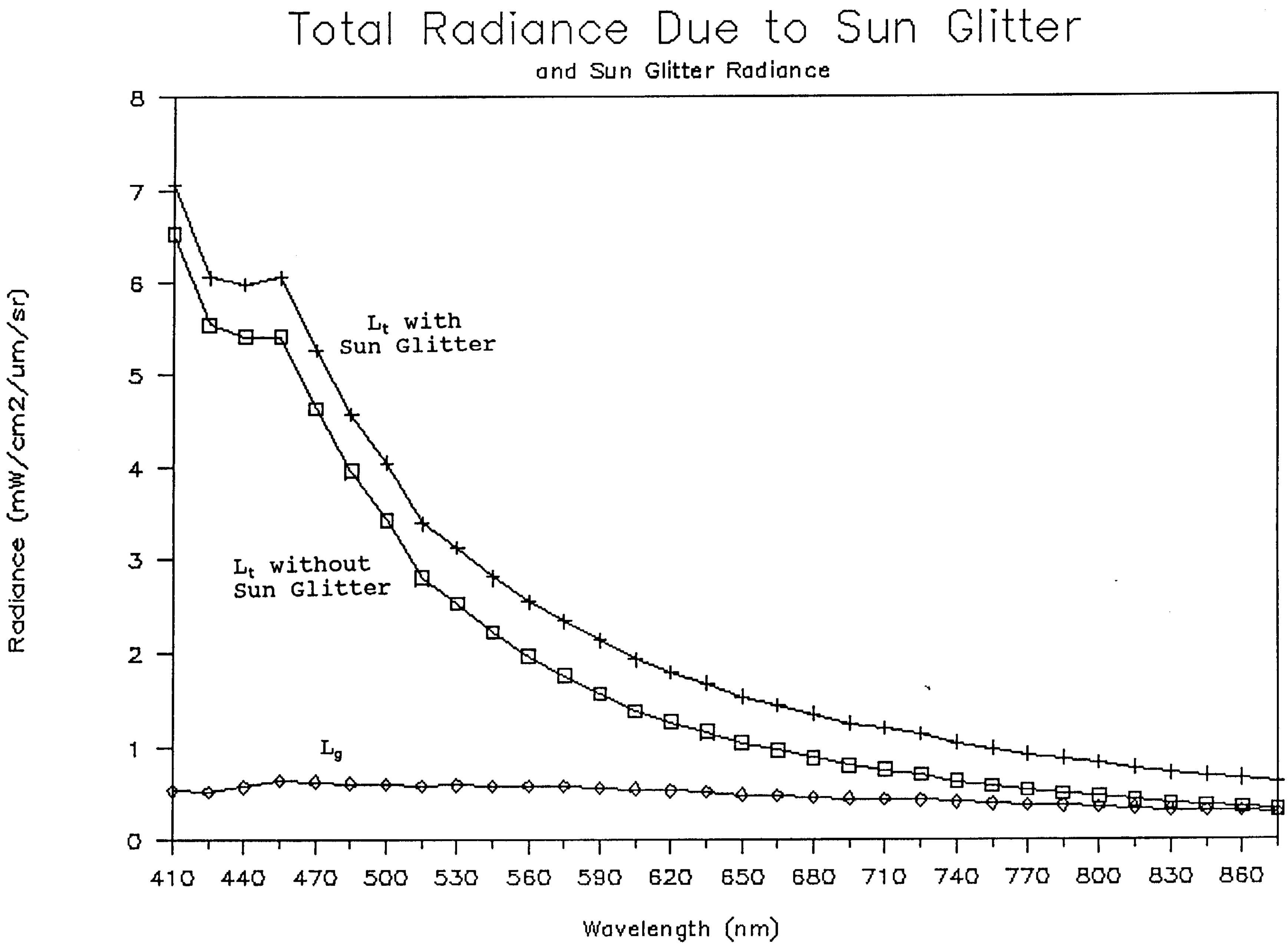


Figure 1. Total radiance L_t with and without sun glitter, and the sun glitter radiance L_g at wind speed of 14 m s^{-1}, and low (≈0.03 mg m^{-3}) chlorophyll concentrations.

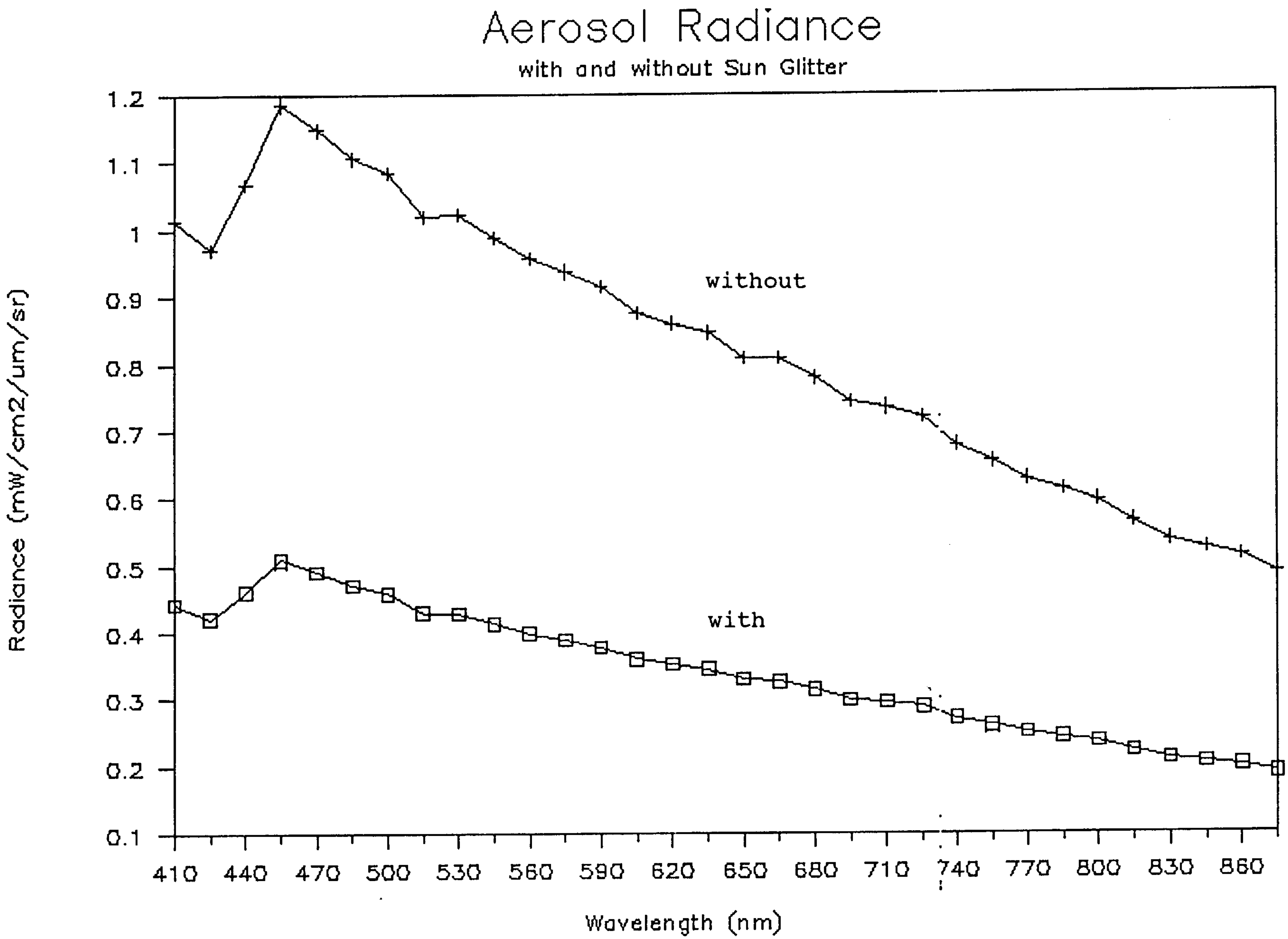


Figure 2. Aerosol radiances computed by the MODIS proposed atmospheric correction algorithm with and without a sun glitter correction.

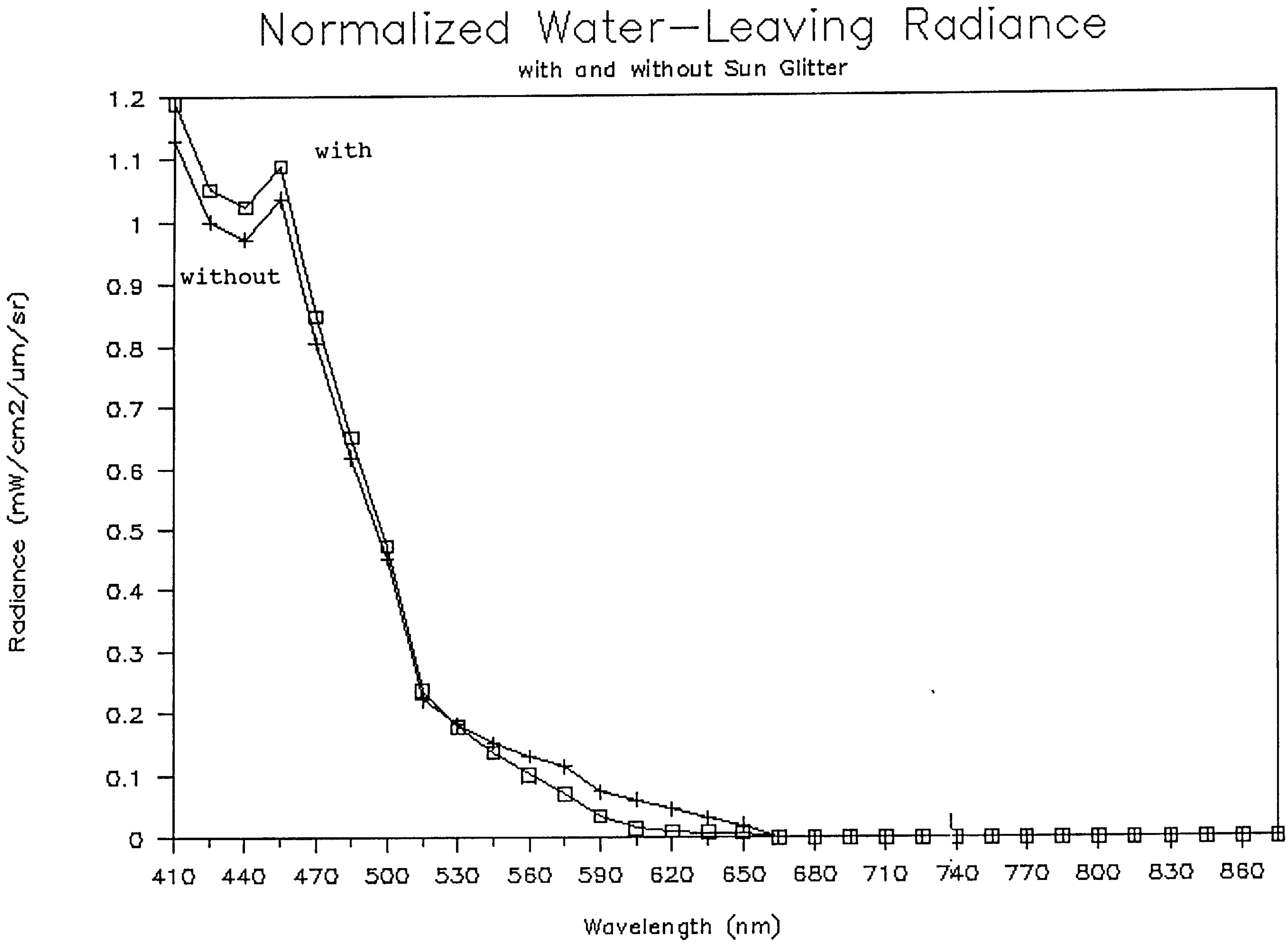


Figure 3. Normalized water-leaving radiances with and without a sun glitter correction.

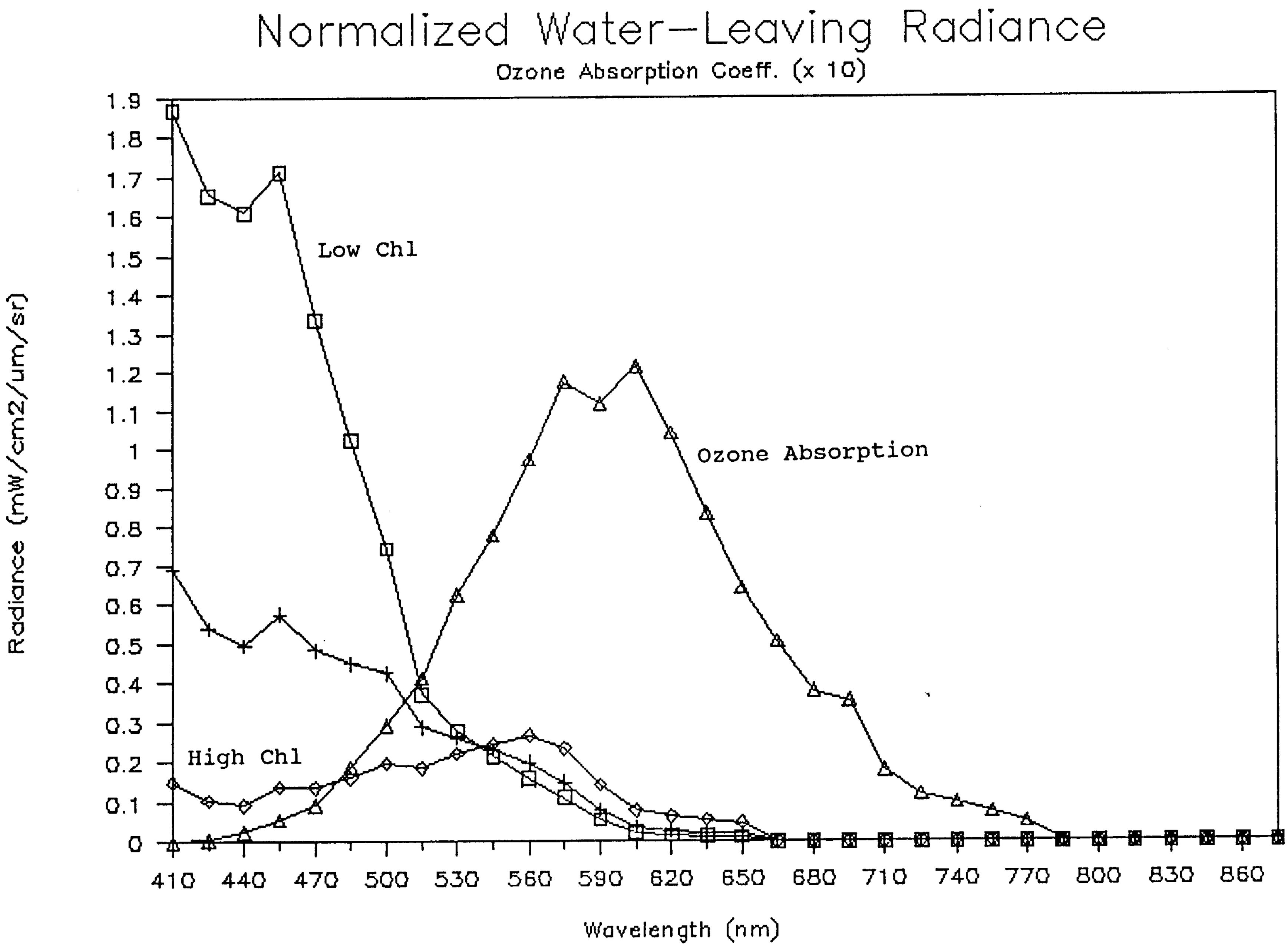


Figure 4. Normalized water-leaving radiance spectra at low (≈0.03 mg m^{-3}), medium (≈1.0 mg m^{-3}), and high (≈6.0 mg m^{-3}) chlorophyll concentrations. Also plotted are spectral ozone absorption coefficients (x 10).

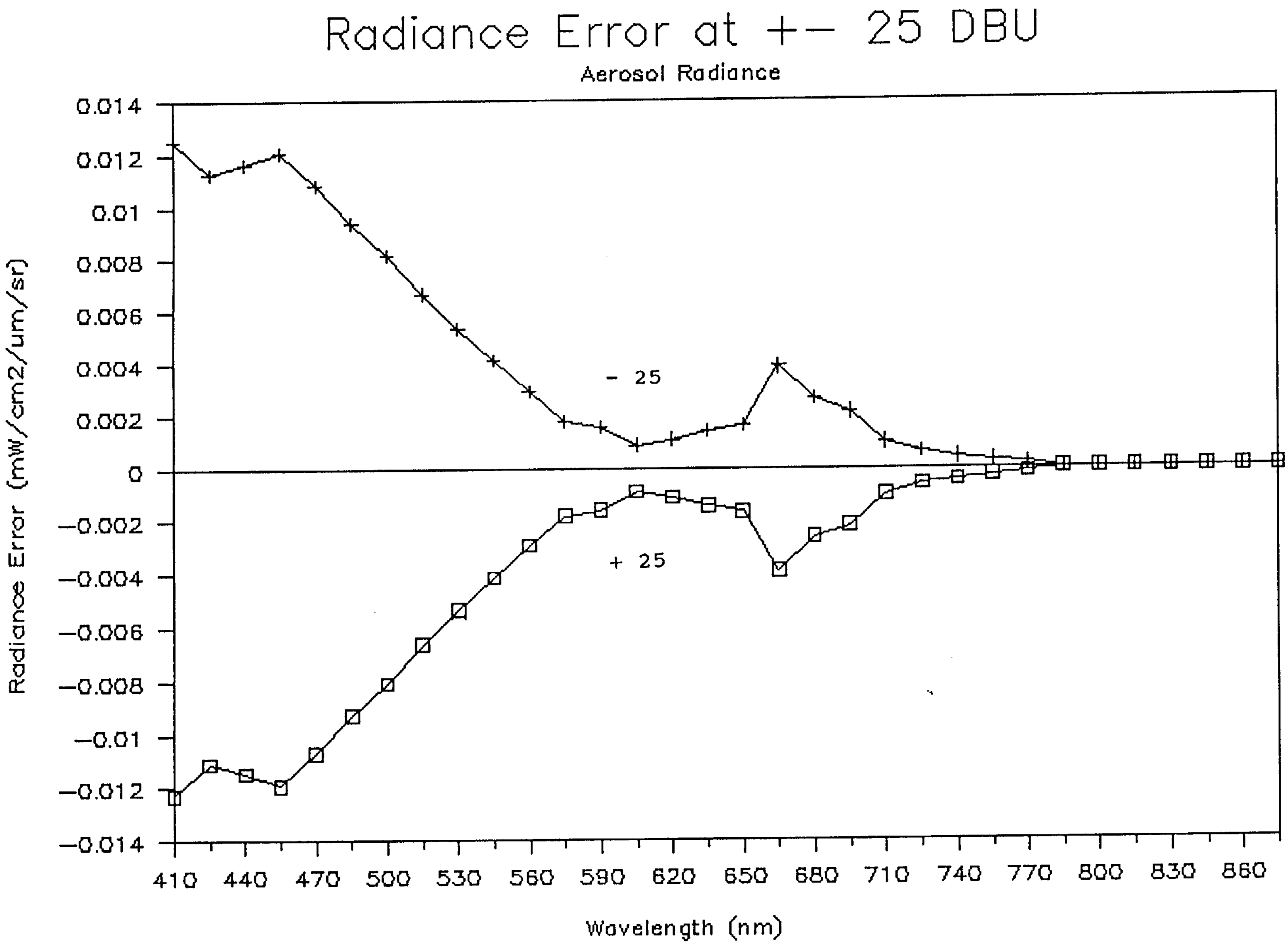


Figure 5. Radiance error in aerosols using the proposed MODIS atmospheric correction algorithm when the ozone scale height is under- and overestimated by 25 DBU.

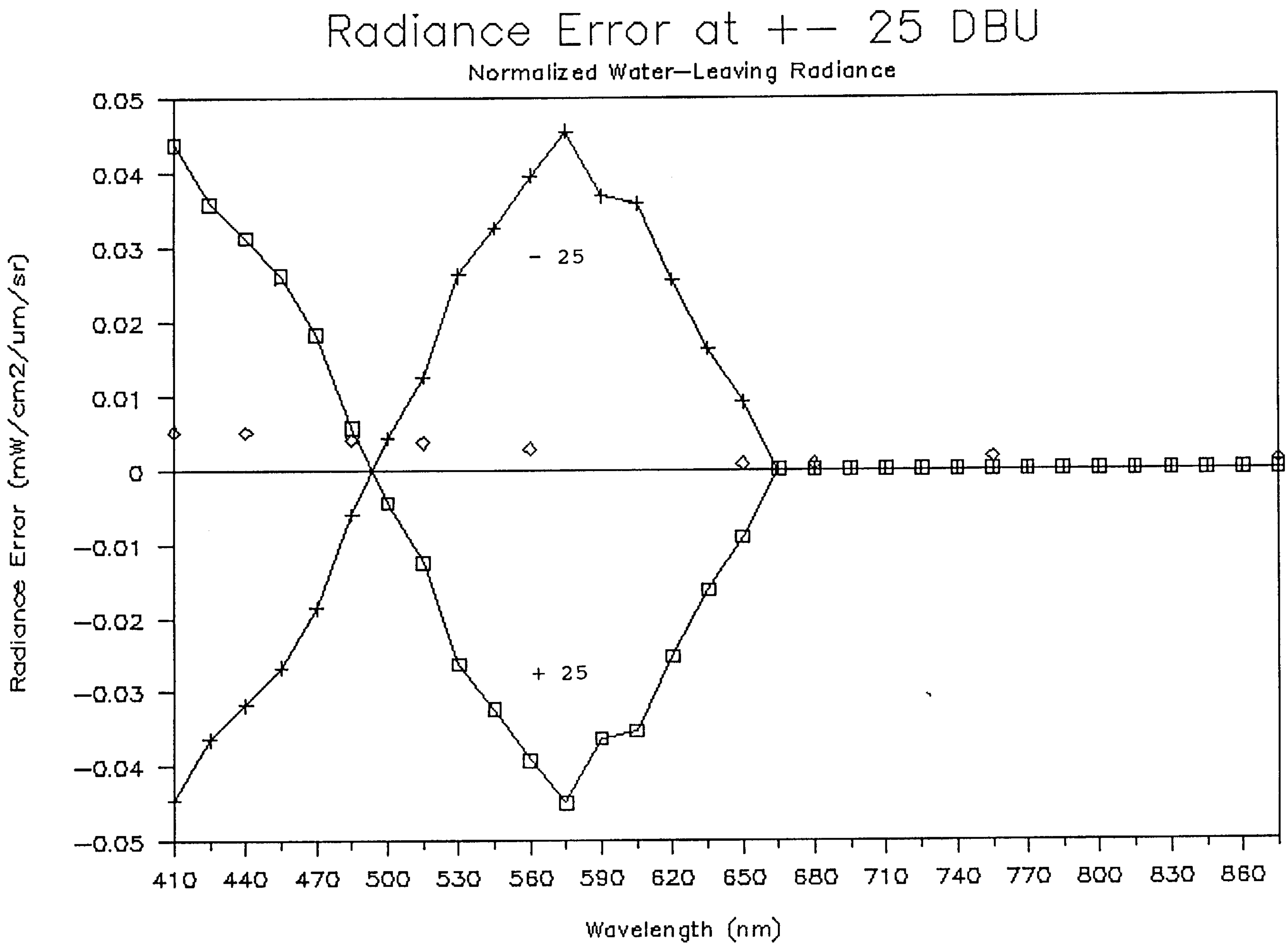


Figure 6. Radiance error in normalized water-leaving radiance using the proposed MODIS atmospheric correction algorithm when the ozone scale height is under- and overestimated by 25 DBU.

REPROCESSING

This section presents our current understanding of reprocessing. The reasons for doing reprocessing (calibration changes, algorithm changes, and ancillary data changes) are presented. A number of issues and questions on the data management required for and associated with reprocessing are discussed. The final subsection of this report considers the implications of reprocessing on the capacity required for the CDHF.

Calibration Changes

If and when it is necessary to change or correct the calibration of the MODIS data, extensive reprocessing will be required. There are two ways in which the calibration can be changed. The simpler method is to use different calibration coefficients. The more complex method would use a different calibration procedure with a different number of coefficients and/or different equations for calibration. The data management and processing capacity issues will be discussed in a separate section.

New Coefficients

The simple type of recalibration would be done by using revised calibration coefficients. In the scenarios which we have developed, it has been assumed that the calibration coefficients are stored in the header of the Level-1A data cube. If only the values of the coefficients change, reprocessing would require that the entries in the header be revised and the calibration could be redone using the same procedure as used in the initial processing.

This type of reprocessing is expected to proceed as follows. (See data management section for additional discussion.) The Level-1A data will be recovered from the DADS or permanent archive. The revised calibration coefficients will replace the old values in the header. The reprocessed Level-1A data could then be sent through the standard processing which would regenerate all of the products which were produced during the prior processing. (This scenario assumes that the standard processing software has not changed since the previous processing.) The Level-1, 2, 3, and 4 data products generated would replace the previously stored values. The reprocessed data would have volume and format identical to the previous version. Any archived data would be overwritten.

It is likely that a change in calibration coefficients would not affect all of the channels. This would imply that not all of the data would need to be recalibrated. There could be some higher level products which would not be changed as a result of recalibration. It is our view that the data management will be so complicated that it is recommended that all products be recalculated. To attempt would be made to determine which products needed to be recalculated.

In summary, when calibration coefficients are changed the header in the Level-1A data will be changed and all of the higher level products calculated. The reprocessed data would replace the previous stored data and data products.

New Calibration Procedure

It is possible that reprocessing of data will be required due a change in the calibration procedure. This would involve the implementation of new calibration equations. A revised set of calibration coefficients would be used and the number of coefficients could change.

The Level-1A data would be recovered as the starting point for reprocessing. If the number of coefficients changes, it would be necessary to modify or replace the existing header. The algorithm used to recalibrate the data will be different from the that used in the initial processing. The processing control software may have been changed to properly read the new header format.

The reprocessed data will again replace the previous version of the data. However, if the header has been changed significantly, the data may no longer fit into the area where it was stored. Any change in the volume of the data during reprocessing may require significant changes in how the data are stored in the DADS. This is an important issue. Certain types of reprocessing may require that the data archive be reformatted.

Algorithm Changes

Algorithm changes will be the second major reason for reprocessing. There are three types of algorithm changes presented in this discussion: error correction, algorithm updates, and new algorithms.

Correction of Algorithm Errors

With the volume of computer code that will be required to generate the MODIS data products, it is certain that there will be some undetected errors in the code as implemented on the CDHF (e.g., X*2 instead of X**2). The correction of these errors will require that data be reprocessed.

The reprocessing will begin with the data that is required as input to the corrected algorithm. As an example, the input to a weekly composite product would be seven days of Level-2 data. The standard processing sequence would be started at the point of the corrected algorithm. The corrected product and all the products calculated using that product as input will be generated. The reprocessed products will replace the previous version in the archive.

This type of reprocessing will require that it be possible to start the standard processing sequence at the position of the corrected algorithm. It would also require the ability to overwrite selected products in the archive. This implies a set of process control software that is substantially different than that used in the standard processing.

Updated Algorithms

There will be changes made to algorithms. For example, the resolution of a weekly composite map might be changed from 8 to 4 kilometers. This would double the volume of the particular Level-3 product. A algorithm might also be changed to use a different source of ancillary data.

This type of reprocessing would require more significant changes in the data processing control software. In the second example, reprocessing would require that ancillary data be used which was not stored in the header of the data cube. The control software would need to identify the location of and obtain the required ancillary data. (Reprocessing could not be done if the ancillary data were not available.)

In the first example above, the data volume of the particular Level-3 product would be doubled. The reprocessed data would not fit in the storage location of the original data. (See the section on data management for a further discussion of this issue.)

New Algorithms

There will be additional MODIS algorithms developed and implemented after launch. A majority of the Team Members' proposed data products are not included among the core data products. Furthermore, it is expected that some of these products will be much more complicated than the core data products.

New algorithms will be developed and installed in the standard data processing. It is likely that some or all of the previously collected data will be reprocessed to generate this newly installed product. The data management problem will be particularly difficult in this situation (see data management section). The new product will not replace a previously archived product.

Ancillary Data Changes

Ancillary data will be used in generating most of the geophysical data products. A primary source of ancillary data will be other satellites and other Eos instruments. The values of the ancillary data may change due to reprocessing done on those experiments. When there is a significant change in the ancillary data, it will be necessary to reprocess MODIS data to regenerate products with the corrected values of non-MODIS input data.

This reprocessing will be done by retrieving the appropriate level of data from the archive, replacing the ancillary data with the corrected values, and regenerating the affected product and those products that follow. This type of reprocessing will require that the standard processing be started at the appropriate point in the processing cycle. Process control software similar to that required for algorithm correction will be required.

Data Management Issues

There are a number of unresolved issues on how the data will be managed during reprocessing. These issues are presented in this section beginning with a brief overview of how the data will be managed in standard processing.

The standard processing will be fully automated and controlLed by a process management function in the CDHF. The raw MODIS data will be received, calibrated, and combined with any needed ancillary data. The assumption has been made that any non-MODIS data needed to generate standard products will be written to and retained in the header of the MODIS data and/or data products. The process control software will drive the proper sequencing of the data processing and ensure that all of the inputs to a given algorithm are available.

It is assumed that the raw data, all of the derived data products, and all of the required ancillary data products will be archived together. Metadata, browse data, and catalogue data will be generated and sent to the IMC.

Reprocessing will require different process control software. How much different will be determined by the exact type of reprocessing to be done.

Here is a specific yet hypothetical example. In standard processing, the atmospheric correction algorithm used as part of ocean data product generation requires surface pressure with an accuracy of 1 mbar. The pressure is obtained with this accuracy by analysis of AIRS/AMSU data. The standard processing control software waits until the pressure is available before proceeding with the atmospheric correction.

Four years after the launch of MODIS, an improved atmospheric correction algorithm is perfected and the decision is made to reprocess all the ocean products. Only the ocean data products will be reprocessed. The processing control software will be required to obtain the data from the DADS and permanent archive, start the processing at the point where atmospheric correction is done, and recalculate only the ocean products.

The needed AIRS/AMSU data may be stored as ancillary data in the MODIS data record. If not, the ancillary data must be recovered from the DADS and/or permanent archive where it is stored. This will require that the process control software interact with the IMC to locate and request the desired data. The control software will also generate only ocean data products.

This example makes it clear that the process control software for reprocessing will be fundamentally different from the standard processing control software. It is possible that each reprocessing task will require modifications to the reprocessing control software.

In the above discussion, the assumption was made that the reprocessed data would replace the previous version of the products. It has not been decided that this procedure will be followed. However, the are several reasons why replacement should be done. First, reprocessing will be done to correct errors in either the calibration or product generation. There does not appear to be any reason to retain products that were wrong.

It does not seem appropriate to store multiple versions of the same product. The "best" version of the product should be retained. If the product that results from reprocessing is not clearly better than the previous version, it is probably not appropriate to reprocess the data.

Another reason for replacing data on reprocessing is to reduce the volume of data to be stored. The system is being designed to allow for 2 reprocessings of the data. If all of the results are retained, the volume of the data archive would be tripled.

It may not always be possible to simply replace old data with reprocessed data. If the volume of the reprocessed product is larger than the old version, the new data will not fit in the place of the old data. This would require either than the archive be reformatted or that techniques be developed to manage data when the set of products is not stored together.

When the standard products are generated, the raw data and all of the products generated from that data will be archived in a contiguous storage area. This will be an efficient storage method and will greatly simplify the data management problem (e.g., it would be possible to locate any data product in the archive by some simple criteria such as observation time). If products from a single set of observations are in a variety of locations, the data management becomes much more complex. It is likely that the data management issues are so difficult that the archive will be reformatted when necessary to keep data sets together.

Metadata will be generated during all reprocessings and in some cases browse data will be generated. This data will be used by the IMC and is assumed to replace the previous versions. The record of all processing done on any data will be retained as part of the data history.

There are a couple of issues on data distribution and reprocessing. There will be data in the archive that is to be reprocessed but for which the reprocessing is not completed. Will this data be available for distribution. As an example, it may be necessary to reprocess to correct calibration errors. The archived data will be inaccurate at best. It might be reasonable to hold this data until it is reprocessed, i.e., don't send out inaccurate data. On the other hand, the reprocessing could take a long time (one year to recalibrate two years worth of data) and the scientists really shouldn't be forced to wait to do their science.

The second issue regards data distribution. The IMC will retain a record of what data has been sent and to which scientists. It will be necessary to at least inform the scientists who have received the data when it is reprocessed and redistribution of the data may be necessary. It is unclear how the distribution of reprocessed data will be done.

The final issue on data management can be referred to as the "processing tree". There will be a complex set of dependencies on the various data products. This will include not only the dependencies among MODIS data products but also the products which MODIS requires from other sources and the MODIS products used by other instruments and/or science teams. The reprocessing of a low level MODIS product will require the recalculation of higher level MODIS products plus all of the non-MODIS products that use the new MODIS products as input. There will also be reprocessing of MODIS products as a result of changes in ancillary products. The interdependence of the various instruments must be given careful attention.

In summary, there are a number of issues on the data management required for reprocessing. They are:

1. Process control software.
2. Data replacement.
3. Archive reformatting.
4. Data access during reprocessing.
5. Data distribution after reprocessing.
6. Processing tree.

Processing Capacity/Requirements

This section contains a discussion of the needed capacity for reprocessing and the requirements on the CDHF needed to accomplish this task. Unfortunately, the topic of reprocessing is not well developed so quantitative estimates cannot be made.

The Level-1 documents require that the CDHF be designed to allow all data to be reprocessed twice. This requirement has been included in the preliminary sizing estimates by multiplying the standard processing estimate by three. This is a very rough estimate. The several types of reprocessing will require varying levels of computing capacity.

Only calibration changes will require that most of the initial calculations be done during reprocessing. Even in the case of calibration changes some of the calculations (e.g., Earth location or cloud detection would not be repeated). The reprocessing of higher level products will start well along the processing sequence. There may be much less processing involved in recalculating a level-3 product then was done in the initial calculation. The factor of three used in sizing the CDHF may be a large overestimate.

The capacity required for reprocessing is a major driver on the system design. An accurate estimate of reprocessing is necessary to generate a defensible estimate of the required CDHF capacity. The reprocessing concepts must be developed and reprocessing scenarios generated.

It will not be possible to do reprocessing without "good" justification. There must be a formal mechanism for submitting and approving reprocessing requests. Reprocessing requests will only be done with adequate scientific justification. The effects of any reprocessing requests on the CDHF must be considered.
Even the simplest reprocessing may have significant impact on the performance of the CDHF. The prioritization of reprocessing requests can only be done for the Eos project as a whole.

The CDHF resources used for reprocessing will be in an operating environment separated from the resources used in the standard data processing. The data flows will be different, different process control software will be used, and different types of processing may be done. The operations concepts and scenarios which have been developed for the standard product generation do not apply to reprocessing.

There are many questions and issues concerning reprocessing that need to be answered. Obtaining these answers will be difficult since a reliable estimate of many calibration and algorithm errors will be made is required. A defensible estimate of the capacity required for reprocessing absolutely necessary if the CDHF is to be properly sized.

UTILITY/SUPPORT ALGORITHMS

Atmospheric Corrections for Terrestrial Data Products

Correction for atmospheric effect can produce remote sensing signals that are better related to the surface characteristics (Kaufman, MODIS Proposal), as opposed to uncorrected data. Atmospheric correction of MODIS data will result in land leaving radiance as a Level 2 data product. Land leaving radiance is a measure of the energy reflected from a terrestrial surface, and may be better related to surface characteristics than the top of the atmosphere radiances (i.e. Level 1, radiometrically corrected data) that contain energy contributions from a variable atmosphere. Land leaving radiances will allow for research with quantitative goals such as, computing energy exchanges at the surface, required in research concerned with modelling energy flows in ecosystems, or in studying changes in surface characteristics (e.g. vegetation) over time.

The problem of how to correct for atmospheric effects in terrestrial studies is a problem open to debate at the present time. Unlike remote sensing studies of the oceans, where atmospheric correction algorithms have been developed and are commonly applied, the application of atmospheric corrections in terrestrial remote sensing is not common. It is apparent from MODIS team member proposals that there will be a demand for data that have been corrected for atmospheric effects. Development and implementation of algorithms for atmospheric correction for terrestrial studies is an issue that requires more input from the MODIS science team members.

Many of the MODIS team members who submitted terrestrial study proposals identified the need, or desire for atmospherically corrected data i.e. land leaving radiances, for analysis in some phase of their studies either for core data products or proposed research and development products. In Table 1 is a list of those land core data products that will require atmospheric correction, or be enhanced by the use of land leaving radiance in their calculation. MODIS team members with expressed need or desire for atmospherically corrected data in their proposals are; V. Vanderbilt, A. Strahler, J. Muller, A. Huete, C. Justice, and S. Running. Those who have proposed developing atmospheric corrections as part of their objectives are; V. Vanderbilt, Y. Kaufman, D. Tanré,

At this point in planning, the issue of how atmospheric corrections for terrestrial studies are going to be accomplished is unresolved. The atmospheric correction problem must be further addressed by the MODIS team members and potential solutions presented so that atmospheric correction algorithms can be planned for in MODIS data processing flows and system requirements. It seems that at this time, that one or more of the team members should assume the responsibility of providing algorithms for atmospheric corrections to be implemented at the CDHF for production of land-leaving radiances.

TABLE 1.	MODIS land core data products that require, or may be enhanced by using data that has been corrected for atmospheric effects.
	Reflected land-leaving radiances
	Emitted land-leaving radiances
	Vegetation indices
	Polarized vegetation indices
	Spectral bidirectional reflectance distribution function
	Snow and ice coverage

Four atmospheric effects proposed to be corrected for are; Rayleigh scattering, aerosol scattering, ozone absorption, and precipitable water. The methods of attaining corrections for these atmospheric effects vary among the disciplines of atmosphere, oceans, and land, because their goals are somewhat different and regions of spectral interest vary. But, the processing level that the corrections should be made at is similar in all disciplines; directly following radiometric calibration at Level 1. The atmospherically corrected data may then be identified as Level 2 land-leaving radiance, a core land data product.

Each of the four atmospheric effects proposed to be corrected for are discussed in brief, concerning the spectral region they affect, and possible ways to correct for them, in the following sections.

Rayleigh scattering

Is a dominate atmospheric effect in visible wavelengths and out to approximately 1.0μm; (scattering intensity is dependent on the inverse of wavelength to the fourth power) a spectral region employed in vegetation studies. Rayleigh scattering may be accounted for in fairly simply and direct relationships. Data needed for correction are; optical depth/pressure corrected air mass, ground measurement of spectral optical depth, from Langley plots, and barometric pressure. Rayleigh optical depth can be calculated from barometric pressure and wavelength (Slater).

MODIS-N Bands identified for correction are; 1,2 (214 m IFOV); and bands 13,14,16,17,18,19 (856 m IFOV). These are all bands within the red 0.6-0.7 μm and near infrared 0.75-1.1 μm, spectral regions. These regions cover those bands that have been proposed for use in MODIS team member proposals concerned with land.

(Also, these are the spectral regions recommended in the literature for remote sensing of vegetation.)

Mie (Aerosols) scattering
Aerosols in general have a smaller contribution than Rayleigh scatter and are much more difficult to correct for. Data need for corrections; aerosol optical thickness, can be obtained from ground measurements (Kaufman) or, aerosol optical thickness can be derived from satellite radiances over selected ground targets (Kaufman).

Ozone Absorption
Correction for ozone is not well defined in the proposals.

Precipitable Water (Water vapor) Absorption
Though the MODIS bands of interest in terrestrial vegetation occur in spectral regions where the atmosphere is highly transparent to water vapor, there may be a desire by some users to account for water vapor absorption. Data needed to correct for precipitable water may be obtained from; ground relative humidity, temperature measurements, and radiosonde data (Slater).

Data required to correct for these atmospheric effects may come from sources other than the MODIS itself. Possible sources of data for atmospheric correction algorithms were given above. Depending on the source(s) of required data, the timeliness of producing data products at the TMCF or CDHF may be affected.

The question also arises of on what scale must these atmospheric corrections be made. Can they be calculated and applied at a low resolution, or are they required to be calculated for each pixel. Part of this answer may depend on the variability of air masses.

Time and Space Averaging

Purpose of averaging.

Several motivations for generating time and space averages of MODIS products can be identified. Possible motivations include:

1. The identification of major data trends. Averages suppress the individual variability of data and may allow the identification of trends not apparent from an examination of the discrete data items. Averages computed to identify trends may sometimes be presented in a conventional "map" type presentation where each pixel represents an average value for a region and time period.

However, averaging to identify trends may also include the generation of a single number that represents the summation of all activity for a domain under investigation. An example of such a single-number average might be the Normalized Difference Vegetation Index (NDVI) averaged over a particular terrestrial biome, say Coniferous Forests of the Pacific Northwest. The average NDVI for such a region might be displayed as a single number superimposed on an outline map showing the boundaries of the defined region.

Averages in the time domain may be particularly useful to a user not equipped to do EosDIS kinescope presentations. For some types of analysis, the data reduction associated with high-level spatial and temporal averaging may be essential to creating a doable analysis.

2. The reduction of spatial and temporal variability of data. When spatial data is presented on a visual display, the variability between adjacent pixels is apparent in the image, and with suitable selection of color display thresholds, the variability characteristics of various regions may be discernable as color or tone variations across the display. Nevertheless, the data "granularity" characteristics of a scene may not be a useful topic for display in all presentations, and spatial averaging is a natural way to reduce data variability, if needed.

If the EosDIS design includes support for a kinescope or "movie" data display mode, then the time variability of data will again appear as a certain "graininess" in the presentation, and, as with spatial variability, the eye can serve as a data "integrator" that discerns the general trend of the data. While such a maximum time resolution display is inherently useful, as in the case of maximum spatial resolution, the information in such a display may not be particularly relevant to the topic at hand, and averaging can again be used to reduce data variability.

3. The presentation of synoptic views of data products at reduced spatial resolution. The CRT displays used with data systems are

normally capable of displaying only a very limited number of pixels. Since the resolution of the MODIS instruments is about 1 km, the number of pixels generated for a regional or global scene will exceed the display capabilities of a terminal. Spatial averaging is a natural way to reduce the number of pixels displayed while retaining and displaying the general character of the data.

Algorithm requirements.

Time and space averaging algorithms for MODIS Data Products will probably be included in a set of general-purpose utility/support algorithms provided for use by all MODIS Science Team Members. Time and space averaging will be done during Level-3 processing after Earth-gridded, full spatial resolution data product images are generated.

Several granularities for time and space averages have been proposed by MODIS Science Team members; examples are listed in Tables 1 to 3. No consensus on required product granularity is apparent from Team Member proposals, and the utility algorithm will doubtlessly need to support a variety of product granularities. Time-averages might be generated daily, weekly, monthly, and annually. Spatial averages will doubtlessly be generated for various subsets and supersets of the regions defined in the Eos standard Earth reference system (not yet agreed upon by the Science Team).

To support the generation of single-number averages for terrestrial biomes, the algorithm will need to allow the specification of irregular regions over which the appropriate quantities are to be averaged. Potential methods of defining regions for the computation of averages include the specification of geodetic coordinates, mouse-drawn outlines on suitable "map" displays, threshold filters working on the parameter to be averaged or other geophysical parameters, and all unions, intersections, and complements of the above-defined regions.

Dr. Robert Evans suggests that the sum of the variable in question, the sum of the squares of the variable, and other intermediate statistical variables should be preserved for each averaging process completed. With these data, one can evaluate the variance as well as the mean of the data, and averages and variances for larger regions can be computed from the data for subregions within the larger region when these data are available.

The averaging process may need to support the generation of weighted averages, i.e. certain values may "count" more heavily in determining desired averages than others. Certainly the system should support the weighting of observations by the size of the area to which the observation or observations applies. Other potential weighting schemes might be based on assessed data quality.

Table 1. Examples of time and space averages proposed by MODIS Science Team members for global products.

Space	Daily	Weekly	Monthly
1 km	Coastal Water-Leaving Radiance Chlorophyll (Evans)		
4km	Open Ocean Water-Leaving Radiance Chlorophyll (Evans)	SST (Brown)	
10 km		Snow Cover (Salomonson)	
50 km		SST (Barton)	Cloud Fractional Area (Menzel)
$1^\circ \times 1^\circ$		Aerosol Optical Depth Aerosol Size Dist. Single Scattering Albedo (Tanre)	Aerosol Optical Depth Aerosol Size Dist. Single Scattering Albedo (Tanre)

Table 2. Time averages proposed by the MODIS Science Team with no spatial requirement stated.

5-Day	Oceanic Primary Production (Esaias)
Weekly	Oceanic Primary Production, Fluorescence Yield (Abbott) Evapotranspiration, Net Photosynthesis (Running) (North America Only)
Monthly	Oceanic Primary Production (Esaias, Abbott), Fluorescence Yield (Abbott)
Seasonal	Oceanic Primary Production, Fluorescence Yield (Abbott)
Annual	Oceanic Primary Production (Esaias) Net Terrestrial Primary Production (Running) (North America Only)

Table 3. Spatial averages proposed by MODIS Science Team with no temporal average stated.

$0.5^{\circ} \times 0.5^{\circ}$	Aerosol Mass Loading, Single Scattering Albedo (Kaufman)

Justice: Vegetation seasonally and annually, no spatial requirement

PRODUCT INTEGRATION

Definition of Product Integration.

Investigators working independently and responsible for only one or a few final data products may implement algorithms and intermediate procedures similar or identical to those already implemented by other investigators. We shall use the term product integration to mean a series of procedures and data system design features that reduce data system redundancy. System hardware and software requirements may both be affected by the elimination of redundant data system features.

Goal and Value of Product Integration.

One obvious reason to reduce data system redundancy is to reduce hardware resource requirements. Specifically, requirements for processing capability and data storage will both be reduced by the elimination of redundant processing and products. Software initial implementation complexity and maintenance requirements will also be affected. However, since product integration may require considerable coordination and agreement among investigators, it is not obvious that product integration will result in overall simplification of the software implementation process. Indeed, one could perhaps argue that the generation of integrated software may require a higher level of software support from EosDIS than that needed for strictly independent product generation by individual team members.

QUESTIONS:

Is the cause of science best served by a single set of consistent, integrated products or by a multiplicity of perhaps inconsistent products and analytic approaches?

Why reduce system redundancy? Is the cost of redundancy primarily an increase in hardware requirements? Do increased software requirements just affect the implementing researchers? Or are software support personnel or other people affected? What about the ultimate data user, is he affected when multiple or inconsistent products exist?