Operational Use of EPOS to Increase the Science Value of EO-1 Observation Data

Mark Abramson, David Carter, Brian Collins, Stephan Kolitz, John Miller, Peter Scheidler, Charles Strauss
The Charles Stark Draper Laboratory, Inc.
Cambridge, MA 02139

Abstract:

We have demonstrated, in an operational setting, that technology in Draper's Earth Phenomena Observing System (EPOS) can improve the overall quality of remote sensing data products by increasing the likelihood that scenes scheduled for imaging will be high quality scenes. We download, process and store cloud data generated by the Air Force Weather Agency's (AFWA) Stochastic Cloud Forecast Model (SCFM) and World-Wide Merged Cloud Analysis (WWMCA). The cloud cover data is used in EPOS to aid the tasking of the Hyperion instrument on EO-1, with the objective to increase the science value of the observation data. The SCFM global cloud cover data includes forecasts provided every six hours for 3-hour periods up to 84 hours in the future. The WWMCA data is an hourly report of the current global cloud cover, based on data from 9 geostationary and polar-orbiting spacecraft. We developed a methodology to fuse spatially and temporally correlated SCFM and WWMCA data and implemented a way to use this fused data to significantly increase the likelihood of EO-1 taking high quality scenes (with total cloud cover less than or equal to 20%). Visualization methods were developed and implemented that enable Human-Machine Collaborative Decision Making.

I. EO-1

The two instruments on EO-1 of interest for this work are ALI (Advanced Land Imager) and Hyperion. ALI provides image data from ten spectral bands. The instrument operates in a pushbroom fashion, with a spatial resolution of 30 meters for the multispectral bands and 10 meters for the panchromatic band. The standard scene width is 37 kilometers. Standard scene length is 42 kilometers, with an optional increased scene length of 185 kilometers. Hyperion collects 220 unique spectral channels ranging from 0.357 to 2.576 micrometers with a 10-nm bandwidth. The instrument operates in a pushbroom fashion, with a spatial resolution of 30 meters for all bands. The standard scene width is 7.7 kilometers. Standard scene length is 42 kilometers, with an optional increased scene length of 185 kilometers. EO-1 has the capability to perform off-nadir pointing by rolling the entire spacecraft. A comparison of the relative sizes and positions of typical ALI and Hyperion scenes are shown in Figure 1, along with the corresponding Landsat 7 WRS[1] path and row.

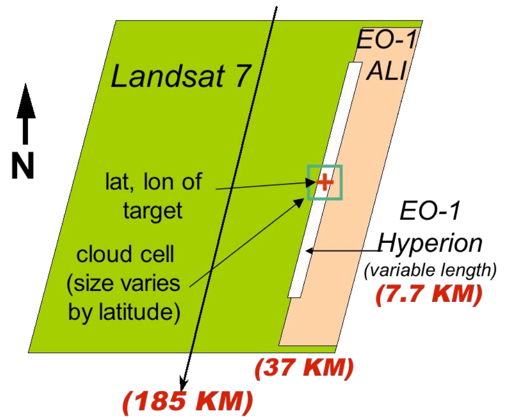

Figure 1: EO-1 ALI and Hyperion Scene Dimensions and Corresponding Landsat 7 WRS Path and Row

II. EPOS FOR EO-1

Variations of the basic EPOS concept of operations have been described in previous ESTC papers. The basic concept of operations is to use observation data gathered from one or more space-based sensors to cue the dynamic replanning and tasking of other space-based sensors. For EO-1 missions operations, we provide cloud coverage estimates, based on AFWA SCFM and WWMCA data, to improve the likelihood of getting high quality scenes.

A. EPOS FUNCTIONS FOR EO-1 OPERATIONS

The full functional architecture for EPOS has been described in previous ESTC papers. Figure 2 illustrates the functions in EPOS that are used for the application to EO-1 operations.

[1] Worldwide Reference System; see http://landsat.gsfc.nasa.gov/documentation/wrs.html

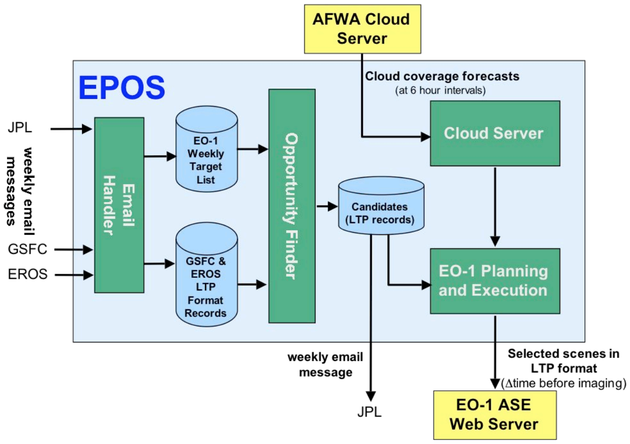

Figure 2: EPOS for EO-1 Functional Architecture

Weekly inputs include emails from JPL, with the following week's target list and pre-picked targets, and emails from GSFC and EROS[2], with LTP[3] records for targets of interest. The opportunity finder runs and identifies opportunities for replanning and candidate targets. Every six hours, new forecasts are obtained from the AFWA. When there are alternate candidate targets and the cloud forecast is such that the alternative target is picked to be imaged, EO-1 planning and execution sends a message to the ASE[4] web site, no less than 8 hours before imaging time.

We are currently automatically accessing cloud data (WWMCA and SCFM) from the AFWA server 24/7. We process the data and store in the EPOS Cloud Server. Queries by visualization (described later) and planning allow access to any of the current or forecast data sets

B. AFWA CLOUD DATA

WWMCA

The Air Force Weather Agency's WWMCA (World-Wide Merged Cloud Analysis) data is produced by the Cloud Depiction and Forecast System II. Each hour, cloud distributions are diagnosed from the imagery of five geosynchronous and four polar-orbiting satellites and merged into a single global cloud analysis. Cloud amounts and types are analyzed for four floating layers. Cloud amounts are expressed in percentages to the nearest 1% with layer tops and bases in meters above mean sea level. Cloud types include cumulonimbus, cirrus, cirrostratus, altostratus, altocumulus, stratus, cumulus, nimbostratus, and stratocumulus.

Cloud cover varies by location, as illustrated in Figure 3. It also varies by time of year.

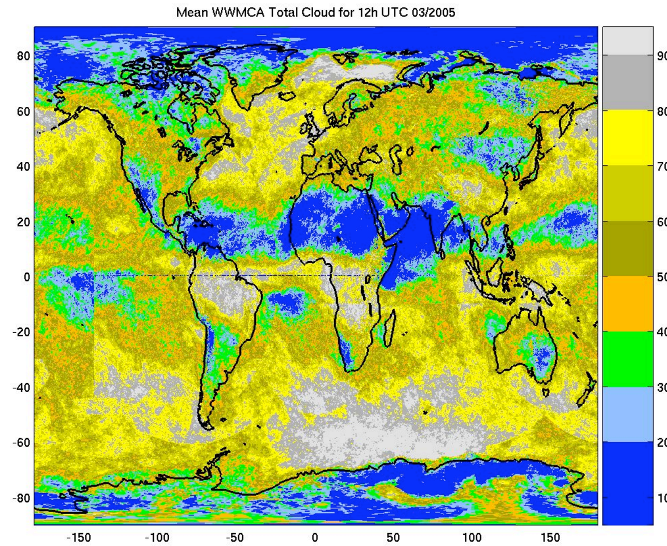

Figure 3: Average Cloud Cover: 1200 UTC (March 2005)

SCFM

SCFM forecasts are generated 4 times per day at the Forecast Generation Time (FGT): 0000, 6000, 1200, 1800 UTC (GMT/Zulu). Each forecast is distributed as 29 gridded binary (GRIB) files. Reception of forecast data starts 90 minutes after these times. Forecasts are given for 3-hour periods over an 84 hour time horizon into the future

Each file includes seven 721x1440 data matrices (0.25º latitude x 0.25º longitude), with predicted total cloud cover, predicted cloud cover at each of 5 pressure (altitude) levels, and a thunderstorm potential indicator. We use the predicted total cloud cover.

2 Earth Resources Observation and Science; see http://edc.usgs.gov/about/background.html
3 Long Term Planning
4 Autonomous Sciencecraft Experiment; see http://ase.jpl.nasa.gov/

Figure 3 summarizes the steps in SCFM processing. SCFM processing for each pressure (altitude) level is as follows:

- Predicted pressure (P), temperature (T), relative humidity (RH), and vertical velocity (ω) are obtained on a 1° latitude x 1° longitude grid from NOAA's Global Forecast System (GFS)
- Use (P,T,RH,ω) to obtain the mean μ and variance σ^2 of cloud density from a lookup table
- The lookup table is constructed by AFWA using a historical database of World Wide Merged Cloud Analysis (WWMCA) data
- Separate mean and variance are obtained for each location
- Obtain predicted cloud cover at each location by taking a random draw from a beta distribution whose mean and variance are the lookup μ and σ^2 for the location
- SCFM predicted total cloud cover for a given location is the maximum of the cloud covers predicted for pressure (altitude) levels at that location

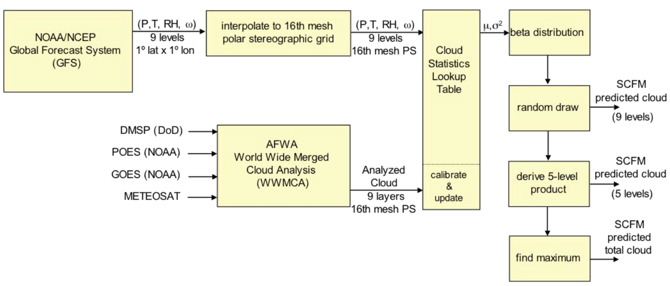

Figure 3: SCFM Overview

C. AUTOMATED TARGET SELECTION USING CLOUD DATA

WWMCA/SCFM Data Fusion Metric

We have evaluated an approach to data fusion for WWMCA and SCFM data for use in our scene selection process for EO-1. The approach takes advantage of the potential increase in the probabilities of taking high quality scenes if SCFM cloud forecasts are used in target selection.

The data illustrated in Figure 4 shows that using the SCFM cloud forecast can increase the probability (estimated using relative frequencies) of getting a high quality scene, i.e., one with total cloud cover less than or equal to 20%. The top picture presents a global view of the unconditional probabilities that the actual cloud cover (WWMCA) is less than or equal to 20% for March of 2005. The bottom picture presents a global view of the conditional probabilities that the actual cloud cover is less than or equal to 20%, given that the 12-hour (lead-time) SCFM forecast was less than or equal to 20% cloud cover. These probabilities vary by location. The green regions are those with a relatively high probability of having less than or equal to 20% cloud cover. The green regions are substantially larger for the conditional probabilities than for the unconditional probabilities.

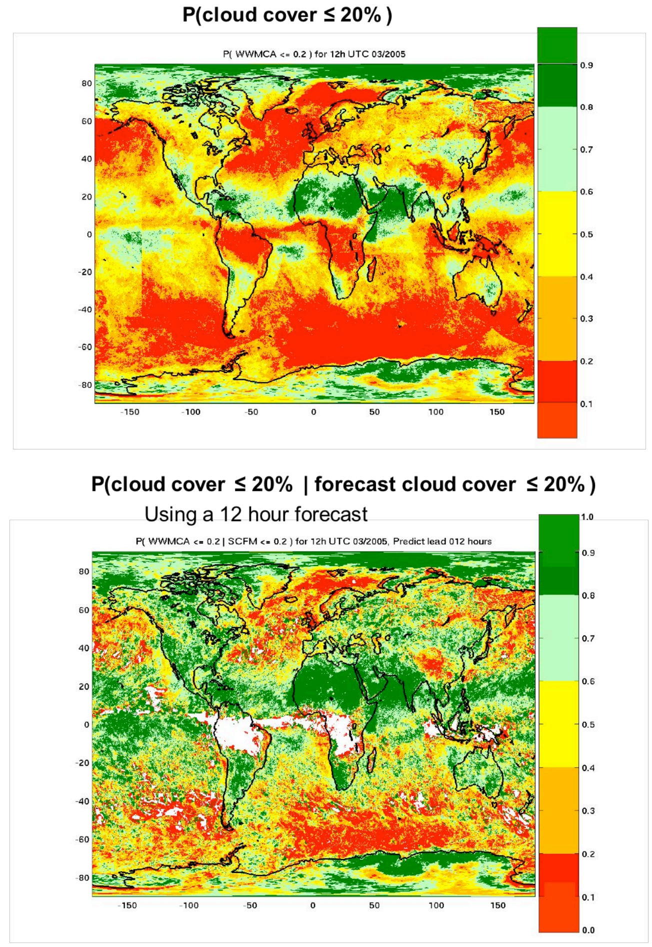

Figure 4: Comparison of the Unconditional and Conditional Probabilities of Cloud-Free Scenes (March 2005)
White = No Data

We developed and evaluated a metric that measures the value of the SCFM forecast for a given location and time of year. The metric is the increase in the predicted probability of getting a high quality scene, conditioned on the SCFM forecast. Figure 5 illustrates this metric calculated on March 2005 data using a 12-hour SCFM forecast.

The red regions are those for which the metric is negative, i.e., the conditional probability is less than or equal to the unconditional probability. In red regions, the metric is not useful. The yellow and green regions are those in which the metric will be useful for improving imaging performance.

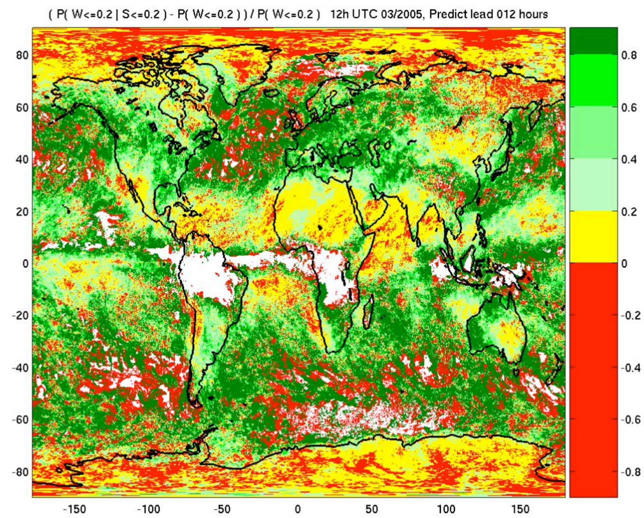

Figure 5: Metric: Predicted Increase in the Probabilities of High Quality Images (March 2005)
White = No Data

Figures 5 and 6 (which illustrate the metric for June and September 2005) show that the metric is a function of time of the year as well as location on Earth.

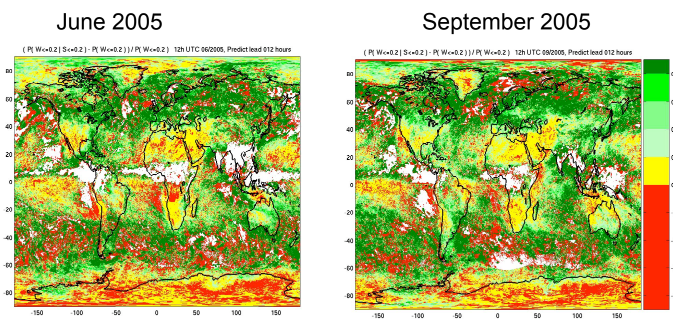

Figure 6: Predicted Increase in high quality Imaging is a Function of Month and Location
White = No Data

We evaluated the value of using the metric to aid our EO-1 scene selection process and found that it will increase the relative frequency of high quality scenes, resulting in a reduction in the expected number of images that need to be taken to get a high quality scene. Details are in the next subsection.

WWMCA/SCFM Metric Evaluation

We evaluated the metric by examining the increase in the yield of high quality scenes. We currently consider a scene to be high quality if it is less than or equal to 20% obscured by cloud. However, the 20% value is an input parameter and can be adjusted if needed in the operational system. "Yield" is the fraction of scenes imaged that are high quality in this sense.

As described previously, the likelihood that a scene will be high quality if target selection / scene scheduling is done *without* the use of cloud forecasts depends on the target location to be imaged, time of year, and time of day. To estimate this likelihood, we use our database of World Wide Merged Cloud Analysis (WWMCA) data. We have a continuously growing database of WWMCA and SCFM data, currently containing approximately 1 TB of data, from which this and other relevant cloud statistics are calculated.

In particular, we determine the historical relative frequency for the event "WWMCA total cloud cover ≤ 20% for the given location and time". When determining this relative frequency in an operational mode, we consider historical data for the 15-day period prior to the given date and also for the 31-day period from the previous year which is centered on the given month and day. We write $P(W_{x,t} \leq 20\%)$ to denote the probability that WWMCA total cloud cover is ≤ 20% for location x and time (date + time of day) t.

If we choose to image a location only when its predicted cloud cover is less than or equal to some threshold a given number of hours in advance, we find that the yield of high quality images will be improved. To investigate this for a particular location x and time t, we determine the historical relative frequency (conditional probability) for the event "WWMCA total cloud cover ≤ 20% for location x and time t" given that "SCFM total cloud cover for x,t is ≤ C". Although our discussions throughout use a value of C = 20%, the threshold C need not be 20%. We write $P(W_{x,t} \leq 20\% \mid S_{x,t} \leq C)$ for the conditional likelihood; it's natural to guess that this is a decreasing function of C. Note that $P(W_{x,t} \leq 20\%) = P(W_{x,t} \leq 20\% \mid S_{x,t} \leq 100\%)$.

We found that there are certain locations and times for which $P(W_{x,t} \leq 20\% \mid S_{x,t} \leq 20\%) < P(W_{x,t} \leq 20\%)$. The SCFM, in other words, is not always a reliable predictor of total cloud as measured by the WWMCA, as shown by the red regions of Figures 5 and 6.

We have developed a capability for studying how well the SCFM forecasts WWMCA (which we equate with actual cloud cover) when the forecasts are used to schedule imaging operations for EO-1. This capability makes use of a list of historical EO-1 targets that we compiled from the target candidates files sent to Draper Laboratory in support of EO-1 operations between July 20, 2005, and April 20, 2006. There are 389 targets in the list.

Using historical EO-1 ephemeris data obtained from the Air Force SpaceTrack web site, we can quickly determine all times in a specified period that each of these 389 targets could be viewed by EO-1. Note that EO-1 can roll to view targets that are within ±2 WRS paths of its ground track. Once the (large) list of possible realistic viewing opportunities is obtained, we associate both the historical SCFM total cloud forecast that would have been available at least 8 hours (the lead time for target selection currently being used in our operations with EO-1) before each viewing opportunity, and the corresponding actual cloud cover, as determined later by WWMCA.

We generate the value of the conditional probabilities at four times during the year. If we don't have the previous year's data we use the month surrounding the given date for the analysis. There are situations in which there were not enough data points to calculate conditional probabilities, e.g., there is no data when the forecast and actual cloud cover values were under the threshold C = 20%. Once this has been transitioned into the operational system, the data will continue to accumulate over time and these gaps will tend to fill in. As an example, looking at the period March 1 and 2 of 2005, there were 265 targets visible and 30 of the targets did not have sufficient cloud data for analysis at C = 20%. When this is the case, we do not include them in our analysis.

The metric, the predicted increase in the relative frequency of high quality scenes, is calculated by:

$$K = (P(W_{x,t} \leq 20\% \mid S_{x,t} \leq 20\%) - P(W_{x,t} \leq 20\%)) / P(W_{x,t} \leq 20\%).$$

If $K > 0$, then the forecast improves the likelihood of getting a high quality scene.

Of the 235 targets with sufficient data for calculation of K, we found the results given the Table 1. The specified level (SL) percentage is a candidate for the specified level used in the proposed new rule given above. Note that 7.2% of the time, the forecast resulted in a reduction in the likelihood of getting a high quality scene.

Specified level = SL	Number of targets with K ≥ SL	Percent of targets with K ≥ SL
0%	218	92.8%
10%	194	82.6%
20%	160	68.1%
30%	142	60.4%
40%	120	51.1%
50%	105	44.7%
60%	89	37.9%
70%	80	34.0%
80%	69	29.4%
90%	59	25.1%
100%	49	20.9%

Table 1 – Simulated EO-1 Views with Cloud Data

These percentages illustrate the benefit of using the forecast cloud data for target selection. One way to interpret a positive K value is as a reduction in the expected number of images that need to be taken of a target in order to get a high quality scene.

We are testing the following rule and are planning to go operational with it by July 1, 2006. When imaging is scheduled to occur on an orbital revolution of EO-1, there is a primary target that has been pre-selected for imaging and one or more alternate targets that could be selected instead.

For the time and location of the target scenes being considered:

- If the metric, *the predicted increase in the relative frequency of high quality scenes*, is more than a specified level (SL), then use the current NASA-supplied tasking rule, otherwise the default target is imaged
- The current NASA-supplied tasking rule is:
  - Select the primary (pre-picked) target if its cloud cover is forecast to be ≤ 20%
  - Select the primary if the cloud cover is forecast to be ≥ 80% for all targets
  - Select the primary if there is less than 20% difference between the forecast for the primary's cloud cover and any alternate target's cloud cover
  - Otherwise, select the alternate

D. Visualization for Support of Target Selection

We have developed mixed-initiative visualization capabilities for Human-Machine Collaborative Decision Making to aid in the EO-1 target selection process. We can watch evolving cloud locations and movement through the visualization cloud data (both WWMCA and SCFM) over the targets as two key event milestones are approaching – 1) the decision time at which either the pre-picked or an alternate target is selected, and 2) the imaging time over the targets. A human is able to visually assess complex cloud patterns, increasing the chances of getting a high quality scene. We are currently developing how the mixed-initiative concept of operations will be used with our existing operations.

The figures below illustrate timelines that begins with the earliest forecast available prior to imaging (84 hours in advance) the target continuing in imaging time. The figures show in the left half the evolving cloud forecast over the target, and in the right half the actual cloud cover at the time of the corresponding forecast.

Pre-picked target left unchanged

The first case shows a situation where the weather over the pre-picked target was superior to that of the alternate, and so the planner did not send in the alternate to be imaged.

The first picture in Figure 7 is early in the process, 83 hours before imaging.

Figure 8 shows the forecast at the time (11 hours before imaging) that a decision had to be made as to whether to use the pre-picked or alternate target. At decision time, the pre-picked target was predicted to be significantly less cloudy than the alternate target (3% vs. 88%). The alternate was not sent.

Figure 9 shows the forecast and actual cloud cover at the imaging time. The actual cloud cover at imaging time confirmed that the pre-picked target had the better weather.

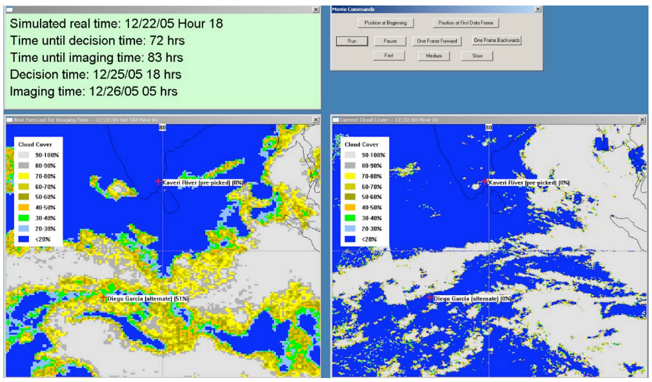

Figure 7: Earliest Forecast: Kavari River vs. Diego Garcia

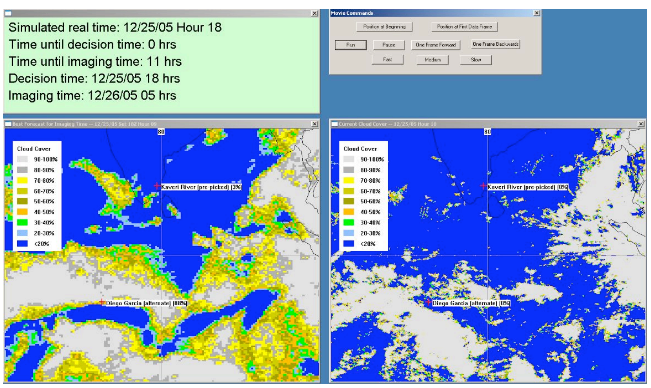

Figure 8: Decision time: Kavari River vs. Diego Garcia

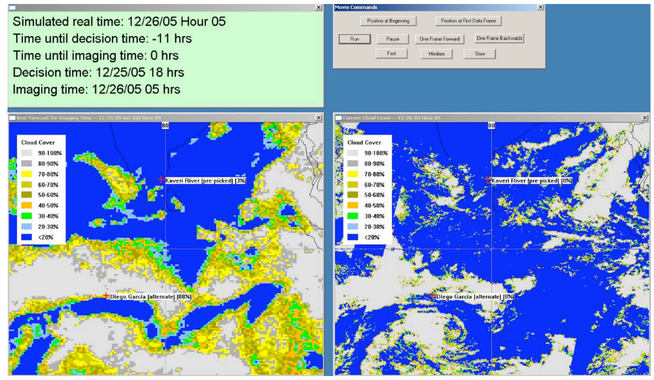

Figure 9: Imaging time: Kavari River vs. Diego Garcia

Alternate target selected

The second case shows a situation where the weather over the alternate was determined to be better than that of the pre-picked target, and so the planner sent in the alternate to be imaged. Figure 10 shows the earliest edge of the forecast, 86 hours before imaging.

Figure 11 shows the forecast at the time that a decision had to be made as to whether to use the pre-picked or alternate target. At decision time, the alternate target was predicted to be significantly less cloudy than the pre-picked (0% vs. 89%). The alternate target was sent in by the planner to be imaged.

Figure 12 shows the forecast and actual cloud cover at the imaging time. The forecast was correct, and the alternate target was clear at imaging time.

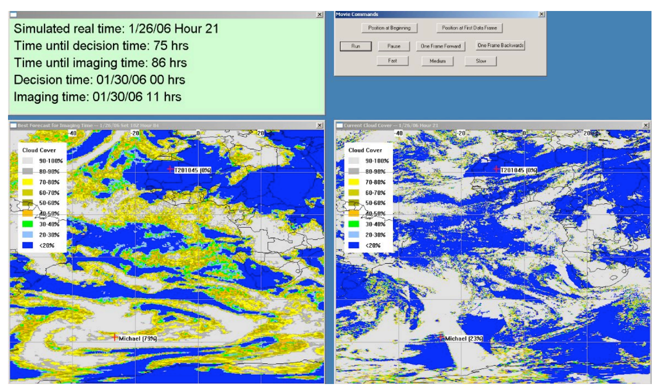

Figure 10: Earliest Forecast of Michael vs. T201045

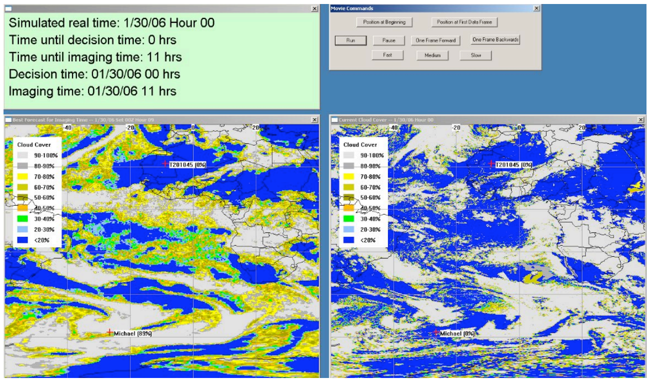

Figure 11: Decision time of Michael vs. T201045

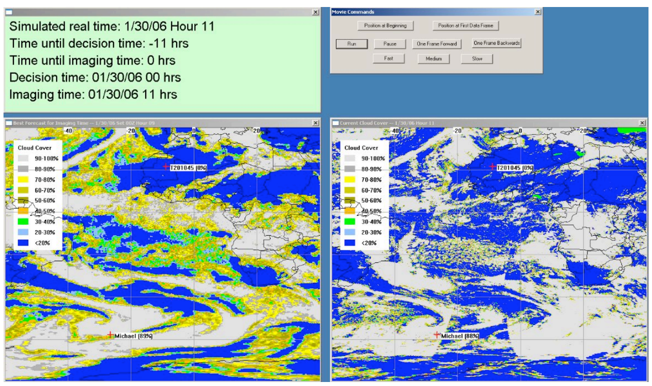

Figure 12: Imaging time of Michael vs. T201045

We also have the capability to show the scene area (swath) on the Earth's surface, along with the cloud cover, either actual or forecast. The following figures contain EO-1's ground tracks, forecast cloud cover and a legend, along with targets taken from the EO-1 target list from December 26, 2005. The swaths for EO-1's ALI and Hyperion sensors are both shown, along with the target location, which is marked with a +. The start (S) and end (E) time of the imaging are shown on EO-1's "displaced ground track" (from rolling), in dark pink; this is an artifact used in the computation of the swath and has been used in verification of the software. The actual ground track is given in red in the larger windows. The satellite ephemeris and sensor swath parameters are input values to the visualization, so other satellite's sensors footprints superimposed on cloud data could be easily visualized using this capability. Figure 13 illustrate the setup and selection of the orbit and the target file for the orbital revolutions, targets and swaths to be generated; forecast set 0 Zulu, hour 09. Figure 14 illustrates the situation for the fourth orbital revolution.

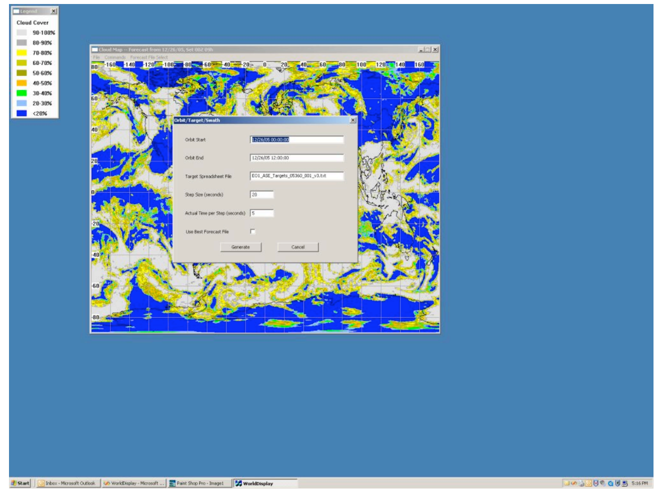

Figure 13: Setup Panel

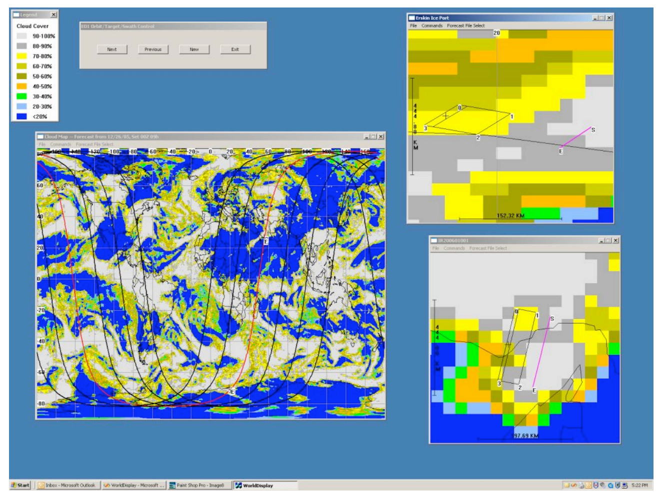

Figure 14: Fourth Orbital Revolution with Targets and Swaths

ACKNOWLEDGMENTS

This material is based upon work supported by the National Aeronautics and Space Administration's AIST program, under Contract #NAS5-03121, through the Goddard Space Flight Center. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Aeronautics and Space Administration