Proceedings of the Twenty-Sixth International Conference on
Automated Planning and Scheduling (ICAPS 2016)

Scheduling Ocean Color Observations for a GEO-Stationary Satellite

Jeremy Frank* **and Minh Do*** **and Tony T. Tran*,+**
*NASA Ames Research Center
Mail Stop N269-3
Moffett Field, California 94035-1000
+Department of Mechanical and Industrial Engineering
Toronto, Ontario M5S 3G8

Abstract

The GEO-Stationary Coastal and Air Pollution Events (GEO-CAPE) mission plans to put a visible spectrum imaging instrument on a satellite in geo-stationary orbit to perform ocean color remote sensing. Two different instrument designs, Filter Radiometer (FR) and COastal Ecosystems Dynamic Imager (COEDI), with different shape for the imaged area and image acquisition time, are being evaluated. Scheduling observations for either instrument requires optimizing science objectives in the presence of predicted cloud cover and available daylight. We model this scheduling problem as both Mixed Integer Linear Program (MILP) and Constraint Programming (CP) problems, and compare these two formulations for FR and COEDI using real cloudiness data collected at different times throughout the year. Our results show that MILP is the more suitable technique, and the schedule quality metric shows FR is the preferred design. We have reported our results to the GEO-CAPE mission team to assist them making an informed decision for the next step in formulating this mission.

Introduction

GEO-CAPE is a proposed Earth-observing satellite mission capable of measuring many properties of Earth's air and water from space. These observations are to be achieved from a geo-stationary vantage point over the equator, in view of North and South America as well as the adjacent oceans. Multiple observations of the same area per day are required to explore the physical, chemical, and dynamical processes that determine atmosphere composition and air quality over (1) spatial scales ranging from urban to continental, and (2) temporal scales ranging from daily to seasonally. Likewise, high frequency satellite observations are critical to studying and quantifying biological, chemical, and physical processes within the coastal ocean and beyond (Fishman et al. 2012). Scheduling observations is complicated by several factors:

- Observations are constrained temporally by available daylight, which varies by location and time of year.
- Cloud cover over the regions of interest changes throughout the day.
- Different classes of science objectives have different ideal and acceptable intervals between consecutive observations (referred to as temporal separation), and the schedule quality depends on the set of separations achieved in the final schedule.

Given that the GEO-CAPE mission is in the planning and feasibility study stage, two different instrument designs are being considered: Filter Radiometer (FR) and COastal Ecosystems Dynamic Imager (COEDI). One of the main goals for our study is to assist the science team in selecting the final instrument design. Specifically, the quality of the observation schedules for different instruments will indicate which instrument can provide better science. For either of these instruments, the resulting scheduling problem is large, even when scheduling observations for a single day.

For each instrument, we model the scheduling problem in MILP and CP, create single-day problem instances utilizing real cloudiness data provided by weather satellites at different times of year, and solve them using the IBM ILOG CPLEX and CP Optimizer solvers. The key results are:

1. MILP produces better schedules than CP.
2. FR is more promising than COEDI in terms of achieving the desired science objectives.

We have reported our finding to the GEO-CAPE mission team to assist them making an informed decision for the next step in this mission.

In the next section we will provide some background on the various aspects of GEO-CAPE. We then follow with the description on how to formulate this problem in MILP and CP, two of the most popular scheduling approaches. Next, we compare MILP and CP formulations on different scenarios involving the two competing instruments (FR and COEDI) at different times of the year. We finish the paper with discussion on the related and future work.

Background

A satellite in a geo-stationary orbit is directly overhead the same spot on Earth at all times. We refer to a *scene* as the area captured by an imaging instrument on the satellite. The set of scenes the satellite can scan (i.e. take picture of) is fixed, and depends on the geometry of the instrument. The instrument can scan one scene at a time, and the scanning

Objective	Threshold	Baseline
Survey	2 hours	1 hour
Targeted Science	1 hour	0.5 hours

Table 1: Temporal separation for classes of science derived from GEO-CAPE Ocean Science Traceability Matrix.

	FR	COEDI
Nadir Pixel size (m)	250	375
Footprint (km)	512 x 512	768 x 168
Resolution (pixels)	4096 x 4069	2048 x 448
Scan / point time (seconds)	158	224
Scenes	19	38

Table 2: Instrument performance characteristics.

time is fixed[1] (and instrument-dependent). A scene can be scanned if it is illuminated and is sufficiently clear of clouds. Scenes will generally be scanned multiple times a day; the exact number depends on the science objectives. The instrument can be idle if no scene can be scanned.

Our satellite observation scheduling problem requires choosing, at each time, which scene to scan to maximize schedule quality, while respecting all constraints. The science objectives, preferences, and constraints are further elaborated in the remainder of this section.

Ocean Science

The GEO-CAPE Oceans Science Traceability Matrix (STM) (Fishman et al. 2012) describes the ocean science questions of interest, and the constraints and preferences on the observing strategies needed to answer these questions. Two complementary operational modes will be employed to meet different science objectives. The first is a *Survey* mode for evaluation of daily, monthly, and yearly variability of ocean chemistry and biology for river mouths, shallow water coast regions, and deeper ocean regions near large river outflows. The second is a *Targeted* mode, which employs high-frequency sampling for observing episodic events that feature daily variation on upper ocean constituents. These include hazards such as harmful algal blooms, which can cause red tides and other emergencies, as well as floods, oil spills, and other disasters. Both of these modes are used for observing the coastal waters of the United States.

The temporal separation constraints and preferences for the Survey and Targeted science modes are described in Table 1. The STM distinguishes between the *maximum acceptable* separation of repeated observations of the same scene, which is referred to as the Threshold separation, and the *ideal* separation, referred to as the Baseline separation.

Imager Specification and Scene Layout

The GEO-CAPE satellite will notionally be located over the equator at 94 degrees West longitude. When pointing the instrument straight down, or at nadir, each pixel is almost a square. However, when pointing the instrument elsewhere, these square pixels become distorted along the off-nadir axes, and also cover a larger area. The GEO-CAPE coverage requirement is to scan the ocean to a distance of 375 km offshore of the United State coastline, as well as the Great Lakes. The scene layout is assumed to be static (scene coordinates do not change day to day in response to predicted clouds or changing science needs). Given an instrument, the first task is to lay out the scenes to cover the coastal regions and the Great Lakes with the fewest scenes (to minimize the time taken to scan all scenes while meeting the temporal separation objectives from the STM). To accomplish this, we used the Satellite Tool Kit (STK) to manually lay out the scenes.

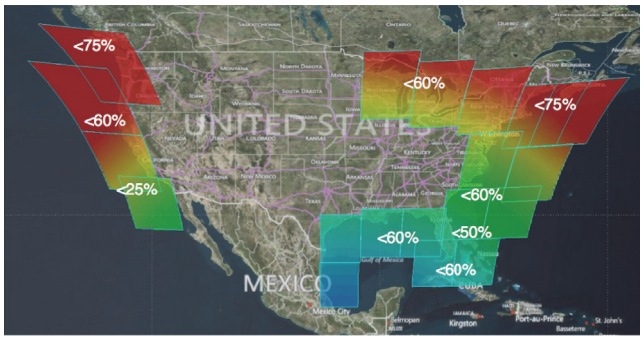

Figure 1: Scene layout for the Filter Radiometer covering the coastal regions. Pixels are 250m at nadir (equatorial); lighter pixels (bottom) are approximately 270m, while the darker pixels (top) are over 360m. The cloudiness thresholds for each scene are also shown.

Table 2 shows the specifications of the two instruments, FR and COEDI. One FR scene covers 512km $\times$ 512km at nadir. With a resolution of 4096 pixels $\times$ 4096 pixels, each pixel covers a square of 250 m $\times$ 250 m. The time taken to scan a scene and reorient to a new scene is 158 seconds. COEDI is a more complex imager design, consisting of two long parallel slits separated by 20 pixels. For the purposes of scheduling, we assumed that consecutive East-West scans are performed to assemble a 'pseudo-scene' with dimension 768km $\times$ 156km at nadir. In this case each pixel covers a 375m $\times$ 375m square at nadir. It takes 224 seconds to perform one scan and reorient COEDI to a new scene. Figure 1 shows the final set of 19 scenes for the FR imager that cover all the interested regions. The comparable scene layout for COEDI is 38 scenes.

Cloudiness Threshold

The picture of a given scene has no value if the acquired image is too cloudy. Specifically, a scene S is deemed to be *clear* if $Y\%$ of S is not covered by cloud. The threshold value Y depends on the average cloudiness over different parts of the coastline. Figure 1 shows the cloud threshold values Y for different FR scenes.

We utilize real cloud data provided by the Geostationary Operational Environment Satellites (GOES). GOES are geo-stationary satellites providing a large number of weather

[1] We have assumed this for our study.

forecasting and data products covering the U.S. and neighboring oceans. These satellites produce 4km resolution weather data every 30 minutes, 24 hours a day. Specifically, each GOES time-stamped image file is a grid of 2000×750 pixels. We use the following data: the *Latitude* and *Longitude* values that define the coordinate of that pixel in the Earth coordinate system, and the *Cloud Phase* value, which specifies if a given pixel is covered by a particular type of cloud.

The procedure to decide if a given scene S is *clear* (not-cloudy) enough to be imaged at time t is:

1. Find the GOES image F of size 2000×750 pixels (1.5M pixels) that was last collected before t (guaranteed to be within 30 minutes of t).
2. Find the polygon P within F that best represents S. This is done by mapping the $\langle latitude, longitude \rangle$ coordinates of S with the *Latitude* and *Longitude* values of the pixels in F.
3. Calculate the number N of pixels in P that are not covered by cloud using each pixel's *Cloud-Phase* value.
4. Take $C = N/|P|$ as the ratio of P that is clear. If C is smaller than the *cloudiness threshold* value predetermined for S (see Figure 1), then we decide that S is clear enough to be scanned.

This approach creates problem instances with 'omniscient' knowledge of the clouds throughout the day and assumes that they do not change significantly except on these 30 minute boundaries. However, it is sufficient to ensure that the resulting scheduling problems have similar characteristics to those that will be encountered in practice.

Land Masks

As shown in Figure 1, most scenes cover both water and land areas. However, given that we are only interested in the water pixels, we mask out the land pixels to avoid biasing the cloudiness count. Let's take as example a scene S covering part of the Great Lakes area (Figure 1) to see why this is important. For S, the water body W only makes up a small portion of the whole scene. Therefore, at a given time t, if most of the land area is covered by cloud but the water region W is not covered, using the procedure described in the previous section would conclude that S is too cloudy to take a picture – even though the region of interest W is clear.

To overcome this issue, we manually used Google Earth to create polygons covering only the water region W of each scene S. We then used the same procedure described in the previous section but only applied to W to determine if S is cloudy or not. Given that: (1) there can be multiple disconnected water regions within each scene; and (2) there can be land region totally enclosed within a water region (e.g., islands), isolating and reasoning only with water regions within each scene is decidedly more time-consuming but it gives us unbiased cloudiness assessments.

From now on, when we mention scene S, we mean the collection of water regions (polygons) within S.

Scene Illumination

FR and COEDI can only take images of scene S when it is illuminated. Sunrise and sunset times can vary by minutes over the locations within one scene; there are also seasonal differences in illumination throughout the year. Therefore, for each day that we need to schedule the GEO-CAPE satellite, we need to find the earliest and latest times when each scene S is fully lit. Specifically:

1. Find the point p_r and p_s within S that have respectively the latest sunrise time t_r and earliest sunset time t_s, among all points within S.[2]
2. Set $[t_r, t_s]$ as the interval during which S is under sunlight and can be taken pictures of.

Science Objective

Translating the Baseline and Threshold temporal separations in Table 1 into formal criteria for schedule quality required extensive discussion with the science team. Below are the key assumptions that we made:

- Each scene is classified either as containing only Survey science or Targeted science.
- The overall schedule score is the sum of the values accrued for the full set of separations of each scene.
- Using the STM as guidance, values of consecutive scans are linear in the temporal separation between those consecutive observations.

Consider the Survey objective: the temporal separation Threshold value is 2 hours, and the Baseline value is 1 hour. This means two consecutive observations must be scheduled 2 hours (or less) apart to achieve the Threshold objective. Consecutive observations are more valuable if the separation is between the Baseline of 1 hour and the Threshold of 2 hours, but the STM provides no insight into the relative value of separations between the two extremes (e.g., separation of 1.5 hours). In consultation with the ocean science community, it was clear that schedules should ensure that the Threshold objectives are achieved before attempting to achieve the Baseline objectives. We started with a value of 0.9 for the Threshold separation of 2 hours, and a value of 0.5 for satisfying the Baseline separation of 1 hour. Thus, meeting the Threshold objective is *relatively more important* than meeting the Baseline. However, if three scans of the same scene can be scheduled in two hours (exactly one hour apart), the accumulated value of $1 = 0.5 \times 2$ is more valuable than only scheduling two conservative scans that are 2 hours apart (only worth 0.9).

The STM leaves undetermined the relative value of Survey and Targeted science objectives. Understanding how these objectives are prioritized is important when some scenes have targeted science, and others do not; the tradeoffs must be formalized to ensure that the schedules produced are consistent with the scientists' objectives. Since the Baseline Survey separation of 1 hour equals the Threshold Targeted

[2]The sunrise and sunset time for a particular location on a particular day can be calculated following the equations described at: http://users.electromagnetic.net/bu/astro/sunrise-set.php

science objective of 1 hour (see Table 1), and since Targeted science is more valuable than Survey science, it is evident that two Targeted science scans separated by 1 hour should be more valuable than two Survey scans separated by 1 hour. Consultation with the ocean science community led to the following objective value for the Targeted science: (1) an objective value of 1 for satisfying the Threshold separation, and (2) an objective value of 0.6 for satisfying the Baseline separation. Similar to the Survey objective, these values ensure scheduling three observations for Targeted science half an hour (Baseline separation) apart is preferred to scheduling two observations 1 hour apart (Threshold separation).

Finally, the STM does not explain how valuable consecutive observations separated by less than the Baseline should be. Allowing any value for intervals of less than the Baseline introduces schedules in which observing the same scene many times is equally valued to observing multiple scenes with Baseline separations, assuming a linear value function. To prevent the scheduler from considering such scenes, we assume there is no value to observing two scenes with a separation of less than the Baseline separation.

Summary: Let V be the value of two consecutive scans of the same scene S separated by a duration D such that: $S_B \leq D \leq S_T$, with S_B and S_T represent the Baseline and Threshold separations. Let V_B and V_T be the schedule quality values if the temporal separation is equal to the Threshold value S_T and Baseline value S_B. Then:

$$V = D \times \frac{V_T - V_B}{S_T - S_B} - \left(S_B \times \frac{V_T - V_B}{S_T - S_B} \right) + V_B \quad (1)$$

Note that $V = 0$ when $D > S_T$ or $D < S_B$. This equation applies for both Survey and Targeted science objectives. Thus, for Survey objectives: $V = 0.4 \times D + 0.1$ and for targeted science $V = 0.8 \times D + 0.2$.

Problem Formulation

To solve the GEOCAPE observation scheduling problem, we utilize Mixed Integer Linear Programming (MILP) and Constraint Programming (CP) frameworks. MILP and CP represent the current state-of-the-art techniques for scheduling problems with out-of-the-box solvers and is commonly used in the scheduling literature (Heinz and Beck 2012; Lam, Van Hentenryck, and Kilby 2015; Pesant, Rix, and Rousseau 2015). The two formulations will use the following shared notations:

- Ψ: the set of all scenes.
- H: the set of all timeslots in the schedule horizon and each observation takes one timeslot. For example, on January 1, 2015, the duration between the first time that any of the scenes in Ψ becomes observable and the last time that a scene can be observed is ≈ 13 hours. Since FR takes 158 seconds to scan a scene and reorient, the scheduling window for FR on January 1 contain $|H| = 13 \times 3600/158 = 296$ timeslots.
- S_B^i and S_T^i: the number of timeslots representing the Baseline and Threshold temporal separation for scene i. For example, if we use FR for a scene i in the Survey mode then: $S_B^i = 3600/158 = 23$.
- V_B^i and V_T^i: the value of performing two consecutive observations of scene i separated by a duration equal to the Baseline or Threshold separation.
- C^i: the set of timeslots within the scheduling horizon that scene i is observable (i.e., illuminated and not cloudy).

A Time-Indexed MILP Formulation

Our MILP encoding uses a time-indexed formulation (i.e., timeslots are indexed) and the full set of variables are:

- $o_{i,j}$: A binary decision variable equal to 1 if scene i is observed at time j.
- $t_{i,j}$ and $d_{i,j}$: $d_{i,j}$ is the separation time between an observation of scene i at time j and the last observation of scene i before j; $t_{i,j}$ is a proxy for $d_{i,j}$.
- $v_{i,j}$: The additional objective value accumulated if an observation of scene i is made at time j. Note that this value depends on the separation time $d_{i,j}$.
- $z_{i,j}$: A binary decision variable indicating whether or not an observation of scene i at time j will add value to the objective function.

$$\max \quad \sum_{i \in \Psi} \sum_{j \in H} v_{i,j}$$

$$\text{s.t.} \quad \sum_{i \in \Psi} o_{i,j} \leq 1 \qquad \forall j \in H \quad (2)$$

$$o_{i,j} = 0 \qquad \forall i \in \Psi; j \in H \setminus C^i \quad (3)$$

$$\sum_{t=j}^{j+S_B^i} o_{i,j} \leq 1 \qquad \forall i \in \Psi; j \in H \quad (4)$$

$$t_{i,j} = j \cdot o_{i,j} - \sum_{k=1}^{j-1} d_{i,k} \qquad \forall i \in \Psi; j \in H \quad (5)$$

$$d_{i,j} \geq t_{i,j} \qquad \forall i \in \Psi; j \in H \quad (6)$$

$$d_{i,j} \geq 0 \qquad \forall i \in \Psi; j \in H \quad (7)$$

$$d_{i,j} \leq |H| o_{i,j} \qquad \forall i \in \Psi; j \in H \quad (8)$$

$$d_{i,j} \leq t_{i,j} + |H|(1 - o_{i,j}) \qquad \forall i \in \Psi; j \in H \quad (9)$$

$$z_{i,j} \leq \sum_{k=j-S_T^i}^{j-S_B^i} o_{i,k} \qquad \forall i \in \Psi; j \in H \quad (10)$$

$$v_{i,j} \leq V_T^i \cdot z_{i,j} \qquad \forall i \in \Psi; j \in H \quad (11)$$

$$v_{i,j} \leq V_B^i + |H|(1 - z_{i,j}) + \frac{(d_{i,j} - S_B^i)(V_T^i - V_B^i)}{(S_T^i - S_B^i)} \qquad \forall i \in \Psi; j \in H \quad (12)$$

$$v_{i,j} \leq V_T^i \cdot o_{i,j} \qquad \forall i \in \Psi; j \in H \quad (13)$$

Model 1: MILP Formulation.

The full MILP encoding is presented in Model 1. The objective is to maximize the total value obtained from making observations of the different scenes. Constraint (2) enforces that at most there is a single observation at each timeslot. Constraint (3) enforces that an observation is not made when a scene is cloudy. The minimum Baseline separation of a scene is guaranteed by constraint (4). Constraints (5 - 9) are used to calculate the actual separation time $d_{i,j}$ between two successive observations. Constraint (10) ensures that an observation will add value only if there was a previous observation that would result in a separation between S_B^i and S_T^i time units. Finally, constraints (11 - 13) calculate the added value $v_{i,j}$ of an observation. For our particular parameters where $V_T^i \leq 1$, using $|H|$ in constraint(12) is sufficient for guaranteeing that $v_{i,j}$ will not be forced to be less than 0 since $H > S_B^i \times \left(\frac{V_T^i - B_B^i}{S_T^i - ST_B^i}\right)$. However, it is not generally the case that this will be true if the input parameters V_T and V_B are changed to have larger rewards. Thus, scaling the second term to some sufficiently large value may be necessary.

A CP Formulation

The CP model makes use of interval variables that are defined by a start time x, end time y, and duration $d = 1$ unit of time. Each interval variable represents an observation of a scene and thus there are $|\Psi|$ sets of interval variables, one for each scene. Prior to obtaining a schedule, it is not known how many observations of a scene are made. However, it is necessary to impose an upper bound on the number of interval variables considered per scene in order to formulate the problem as a CP. Given the set A_i of interval variables for scene i, a valid upper bound is $|A_i| = |H|$, one observation per timeslot. This upper bound is equivalent to the potential observations in the MILP model since one $o_{i,j}$ variable is created for each timeslot and for each scene. For the CP formulation, it is possible to use a tighter upper bound to reduce the size of the model. The bound we use is $|A_i| = \lfloor \frac{\max C_i - \min C_i}{S_B^i} \rfloor$, since we know that at most one observation can occur every S_B^i time slots.

The CP model is presented in Model 2. Here, the decision variables are:

- $a_{i,j}$: An interval variable indicating the jth observation of scene i.
- $x_{i,j}$: The start time of an interval variable $a_{i,j}$.
- $d_{i,j}$: Separation time between observations $a_{i,j}$ and $a_{i,j-1}$
- $z_{i,j}$: A binary variable indicating that separation $d_{i,j}$ will contribute to the objective function.
- $v_{i,j}$: The added value of separation $d_{i,j}$ to the objective function.

We also use additional parameters:

- A: Set of all interval variables (observations).
- A_i: Set of all observations of scene i.
- $\bar{A}_i$: Set of all observations of scene i except the first observation.

$$\max \quad \sum_{i \in \Psi} \sum_{j \in \bar{A}_i} v_{i,j}$$

$$\text{s.t.} \quad \text{forbid}(a_{i,j}, C^i) \qquad \forall i \in \Psi; j \in A_i \quad (14)$$

$$d_{i,j} = x_{i,j} - x_{i,j-1} \qquad \forall i \in \Psi; j \in \bar{A}_i \quad (15)$$

$$d_{i,j} \leq S_T^i + |H|(1 - \text{pres}(a_{i,j})) \qquad \forall i \in \Psi; j \in \bar{A}_i \quad (16)$$

$$d_{i,j} \geq S_B^i - |H|(1 - \text{pres}(a_{i,j})) \qquad \forall i \in \Psi; j \in \bar{A}_i \quad (17)$$

$$\text{pres}(a_{i,j}) \leq \text{pres}(a_{i,j-1}) \qquad \forall i \in \Psi; j \in \bar{A}_i \quad (18)$$

$$x_{i,j} \geq x_{i,j-1} \qquad \forall i \in \Psi; j \in \bar{A}_i \quad (19)$$

$$v_{i,j} \leq \text{pres}(a_{i,j}) V_T^i \qquad \forall i \in \Psi; j \in \bar{A}_i \quad (20)$$

$$v_{i,j} \leq V_B^i + |H|(1 - \text{pres}(a_{i,j})) + \frac{(d_{i,j} - S_B^i)(V_T^i - V_B^i)}{(S_T^i - S_B^i)} \qquad \forall i \in \Psi; j \in \bar{A}_i \quad (21)$$

$$\text{NoOverlap}(A) \quad (22)$$

Model 2: CP Formulation.

The objective is the same as the MILP model. Constraint (14) enforces that if an observation is made, it is made during a time slot where the scene is observable. Constraint (15) compute observation separation value $d_{i,j}$. With $\text{pres}(a_{i,j})$, a function that returns 1 if observation $a_{i,j}$ is scheduled and 0 otherwise, constraints (16) and (17) enforce that an observation will only ever contribute to the objective function if the observation is made, and the separation is within the Baseline and Threshold values for the science objectives of that scene. Constraint (18) and (19) enforce that $a_{i,j}$ will represent the jth observation of scene i if at least j observations are made. The added value of observations are calculated with constraints (20) and (21). Finally, constraint (22) is a global constraint that ensures none of the observations are performed at the same time.

Discussion: An equally common alternative to the time-indexed model using MILP is a disjunctive model. However, our study was built around designing a scheduler to provide insights to the GEO-CAPE mission team regarding their instrument design. Due to the application nature of the problem and the rapidly changing requirements of the mission team, the time-indexed model was chosen so that we could easily adapt the formulation to their requirements. Problem characteristics such as optional observations (require modeling more observations than desired) and the observation separation are less straight-forward to deal with in a disjunctive model.

While both MILP and CP have been heavily used for scheduling problems, each has different strengths and weaknesses. A more expressive formulation in CP allows more straight-forward representations of problem characteristics through not just linear, but also logical and non-linear constraints. CP can also take advantage of specialized inference

	FR											COEDI										
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov
Max	6	6	7	11	9	11	9	11	11	13	9	12	14	16	27	22	24	18	21	26	21	16
Avg	3.3	4.0	4.2	7.6	5.7	7.0	4.9	7.7	7.3	8.4	5.7	7.5	8.14	9.7	19.0	14.1	14.9	9.0	14.0	18.1	15.6	9.8
Never	10	6	7	1	6	3	4	6	4	5	5	20	13	10	2	6	6	7	11	5	10	17

Table 3: Cloudiness data for scheduling problem instances for the scene layouts of FR and COEDI. The 'Max' row shows the maximum number of scenes visible at any time slot during the day. The 'Avg' row shows the average number of scenes visible. The 'Never' row shows the number of scenes never visible during that day.

algorithms to assist in finding solutions. CP makes it easier to model complex constraints, and is able to perform well in searching for feasible solutions. MILP on the other hand is more restrictive in the modeling capability and does not have as many inference algorithms tuned to assist the solver to find feasible solutions quickly. However, MILP uses duality and problem relaxations to drive the solver towards an objective. When using complex and irregular objective functions, CP can be a poor performer due to its constraint focused approach rather than MILP's objective function driven strategies.

Empirical Evaluation

The main purpose of our evaluation is to find and analyze the best observation schedule for FR and COEDI at vastly different weather conditions throughout the year. Since only one of those instrument candidates can be installed on the GEOCAPE satellite, our results can give an objective and informed advice to the mission design team on which instrument is advantageous (and in which way) over the other.

In order to evaluate the candidate instruments more thoroughly, we created a suite of daily problem instances. We picked the first day of each month of 2015 and for each day we (1) first analyzed sunrise and sunset times to create the set of observation timeslots for each instrument's scenes, then (2) used the cloudiness data to determine observable timeslots for each scene, and finally (3) decided whether each scene contains Survey science or Targeted science objectives. Next, we modeled and solved each problem instance using both the MILP and CP formulations described in the previous section.

Cloud Cover Effect on Observation Schedule

Assuming no clouds, FR can scan all scenes of the U.S. coastal waters in 50 minutes (scan time of 158 seconds $\times$ 19 scenes). This is less than the Baseline temporal separation for Survey science, and less than the Threshold temporal separation for Targeted science (both are 1 hour as shown in Table 1). By contrast, COEDI can scan U.S. coastal waters in 142 minutes (scan time of 224 seconds $\times$ 38 scenes), which does not even meet the Threshold temporal separation for Survey science (2 hours). Only 16 COEDI scenes can be scanned in 1 hour (Baseline for Survey) and 32 scenes can be scanned in 2 hours (Threshold for Survey). On the surface, COEDI appears uncompetitive compared to FR. However, this analysis optimistically assumes all scenes are viable throughout the day. Taking cloudiness into consideration will generally reduce contention for scenes, and may affect the decision regarding instrument choice, motivating our empirical evaluation.

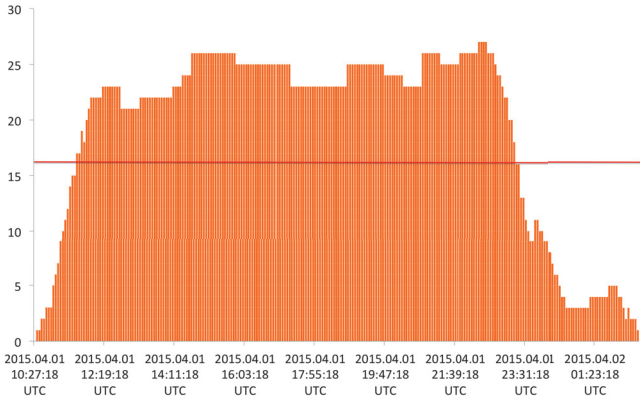

Figure 2: Number of visible and cloud free scenes in April 1 for COEDI. The red line shows the 16 scene limit of the number of scenes that can be scanned once per hour, achieving the Baseline separation for Survey science.

Table 3 shows the cloudiness statistics for the two instruments (by analyzing the GOES satellite data as described earlier). It is apparent that cloudiness drastically reduces contention for the instruments, and that the specific effect of the reduction depends on the time of year. Of special interest is the last row, showing the number of scenes that is not visible at any time of day. This demonstrates that the simple instrument performance analysis assuming no clouds is too optimistic. It is notable that in January, more than half of the scenes cannot be observed at all. By contrast, in April only one or two scenes are not visible throughout the day. The difference between the maximum and average number of scenes visible shows that cloudiness varies considerably over a single day as well. While some of this variation is explained by illumination (in the evening, for instance, the only scenes visible are in the Western United States), other variations depend on daily cloud variation. Figure 2 shows the number of visible scenes per timeslot for April 1, 2015 for COEDI where the number of visible scenes varies between 21 and 27 (out of 38) throughout the middle of the day.

Classes of Science

In our discussions with the ocean science community, even though values of Baseline and Threshold separation were provided for Survey and Targeted science, the tradeoffs between classes of science were not yet fully settled for this mission. Instead of creating arbitrary mixes of Survey and

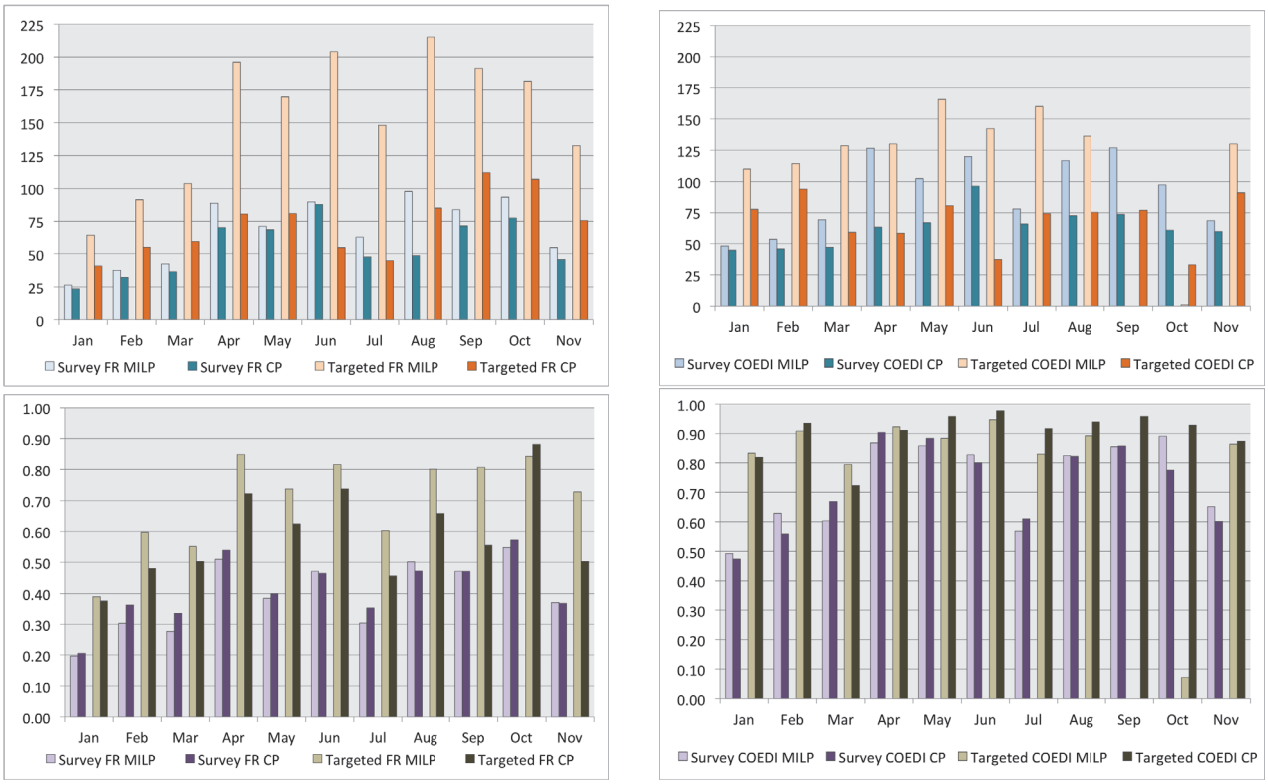

Figure 3: Schedule quality and efficiency comparisons for FR (left graphs) and COEDI (right graphs). The top graphs compare schedule quality for both Survey and Targeted science, and for MILP and CP. The bottom graphs compare schedule efficiency (percentage of timeslots during which the instrument scans a scene).

Targeted science and using the existing prioritization to assess the schedules, we opted to create two problem instances for each day: one consisting solely of Survey science, and the other consisting solely of Targeted science.

Scheduler Results

Each problem instance was solved by running the appropriate solver (either CPLEX or CP Optimizer) for an hour on an Intel Core i7 3.00 GHz CPU (in 64 bit mode) with 2 MB cache per core, 12 GB of main memory, running Linux.

Figure 3 compares the schedule quality and efficiency for FR and COEDI problem instances. The quality measure was described in the earlier part of the paper and efficiency is the percentage of the available timeslots where the instrument was scanning a scene. On balance, the MILP formulation performs better than the CP formulation for the FR problem instances. The MILP schedule score difference is quite large (see July and August in particular for both Survey and Targeted science), while there is little difference in schedule efficiency.

The COEDI problem instances show a different pattern. While the MILP solver creates poor quality schedules for some COEDI problems (specifically the September and October Targeted science problem instances), CP creates at least acceptable schedules for all problems. However, most of the time, the MILP formulation is still better at creating schedules of higher quality than the CP formulation, with quality differences as high as a factor of 3. The CP formulation is as good or slightly better at creating efficient schedules for targeted science problems.

Discussion: Comparing the schedule efficiency score with scene availability (see Table 3) indicates that both MILP and CP are effective in scheduling observations when scenes are available; however, neither solver was able to prove that the optimal solution was found for any of the problems. As noted earlier, the MILP solver failed to find a reasonable solution to two COEDI instances, and the CP solver was unable to find good quality solutions. As mentioned in the previous section, MILP is objective driven and can better handle the complex objective function that we have, while CP in general finds poorer quality schedules as it cannot handle the complex objective function adequately. In fact, other than for two COEDI instances, the MILP dominates the CP in schedule quality, often by a very significant amount. When even finding a schedule is difficult, MILP can perform very poorly, but finding feasible schedules is much easier for CP. The fact that MILP consistently returns schedules with higher overall quality even with similar efficiency scores, compared to CP, indicates that it was able to space out observations much closer to the optimal separation range between the Baseline and Threshold values.

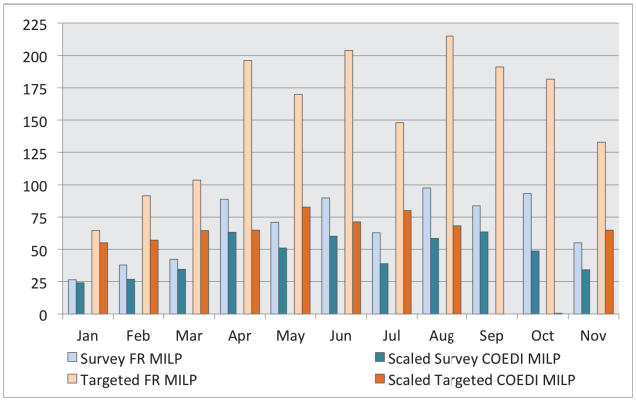

Figure 4: Comparison of schedule scores for FR and COEDI. COEDI scores are halved to account for the larger number of scenes needed to cover the U.S. Coastal waters.

FR vs. COEDI: For the remainder of this section, we will use the MILP results to evaluate the relative tradeoffs between FR and COEDI. The schedule quality for the two instruments is not directly comparable because (1) there are twice as many scenes for COEDI as there are for FR (38 small scenes vs 19 large scenes) and (2) the schedule quality is a function of the temporal separation between consecutive observations of the same scene, without accounting for the larger number of scenes. For this reason, in Figure 4 we compare the FR score to the COEDI score divided by 2. This figure shows that FR outperforms COEDI by a modest amount on most of the Survey science problems, achieving almost a factor of 2 improvement in August and October. The margin is much higher on the Targeted science problems, reaching almost a factor of 3 improvement on the June and August instances. These results show that, based solely on the schedules that can be generated, FR dominates COEDI. However, the dominance is not complete. In the presence of significant clouds, the two instruments are more closely matched, and the gap for the Survey science problems is not as high as it is for the Targeted science problems. Overall, our detailed analysis not only gives more confidence to the instrument designers that FR is currently superior, but also may drive design changes to COEDI to improve its performance, and make it more competitive.

We observe a significant increase in the schedule score between the Survey and Targeted mode. This is because both preclude scheduling consecutive observations with less than the Baseline temporal separation. The score increase reflects slack in the Survey schedules that can be filled in the Targeted science schedules. Since the Targeted Baseline separation is half the Survey Baseline separation, we expect to see the score roughly double between the Survey and Targeted science problems. This pattern is observed for the FR problems, but the increase is not as large in the COEDI problems. The explanation is that the longer scan times for COEDI fill the schedule for the Survey problems, leading to less slack, and less opportunity to improve the score.

Finally, we noted earlier in this section that some scenes were too cloudy over the entire day to observe. Cloudy periods can also exceed the Threshold temporal separation, imposing an implicit cost to the scheduler to add observations and 'restart' observing a scene. Additionally, the scheduler may simply focus on one scene at the expense of others. Table 4 compares the number of scenes never observed in MILP schedules with the number of scenes that are too cloudy to observe all day. We see FR schedules include most (but not all) scenes that are visible at least once. However, COEDI schedules are missing many observable scenes, again pointing to increased contention and insufficient time to meet the mission objectives.

Related Work

Our problem differs from scheduling low-earth observing satellites, e.g. (Pemberton 2000), where the orbital dynamics dominate scene visibility considerations. It shares some aspects with cyclic scheduling (Draper et al. 1999) since each observation is ideally repeated at regular intervals daily. However, the combination of cloudiness, variable visible light, and competition between scenes for the instrument, breaks the regularity otherwise present in the problem. In (Verfaillie and Sebbag 2012), the authors discuss planning image acquisitions of multiple satellites for ocean surveillance mission. While the overall objective is similar to GEO-CAPE, the main differences are: (1) it employs several low-earth-orbit satellites; (2) it is a planning problem with scheduling components; (3) it employs local-search techniques. Also, while this paper does briefly mention 'regularity' of coverage in the optimization criteria it is not quantified formally, unlike the criteria for the GEO-CAPE mission.

A unique aspect of this problem is the use of separation between observations as a metric of performance. Although separation between tasks is not new to the scheduling literature, our understanding is that no other work has looked at the separation in the objective function. All other works only deal with separation constraints, with many looking at either the minimum or maximum separations rather than both simultaneously. An analogous representation of minimum separation that is well studied in the scheduling literature is setup time, as some amount of time must pass between execution of two tasks (Allahverdi et al. 2008). On the other hand, maximum separation problems are more sparse, but do exist. Studies have looked at restricting the maximum separation of jobs in a cyclic schedule (Han and Lin 1992). Works that look at both minimum and maximum separation in the constraints can be found in simple temporal networks (Dechter, Meiri, and Pearl 1991) and resource constrained scheduling problems (Brucker et al. 1999).

Producing schedules with temporal flexibility (Nicola Policella and Oddi 2014) is a viable mean to provide robustness against actual real-time cloud coverage that differ from predictions. We could take the schedules produced by either the CP or MILP approaches and make them temporally flexible, or use a first-principles means to produce temporally flexible schedules, including ensuring that two observations of the same scene are never more than S_b apart. At this time, this is premature since the real-time rescheduling policy for

	FR											COEDI										
Schedule	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov
Survey	13	10	9	4	9	7	9	6	5	5	7	26	23	15	9	16	14	17	14	9	14	21
Targeted	12	8	7	4	7	7	8	6	4	5	6	24	19	15	14	17	17	17	19	38	37	22
Not-visible	10	6	7	1	6	3	4	6	4	5	5	20	13	10	2	6	6	7	11	5	10	17

Table 4: Number of scenes never observed (vs not-visible) in schedules generated using the MILP formulation.

this mission has not been settled by the ocean science community. Furthermore, classical temporal flexibility does not easily extend to preserving schedule quality.

Conclusion and Future Work

We have described GEO-CAPE, namely, scheduling the observations of a visual spectrum instrument whose job is to scan parts of the ocean adjacent to the United States during daylight hours. Since instrument designs are still being assessed, one objective of this work is to provide feedback to the GEO-CAPE mission concerning the two instrument designs; the second objective is to evaluate automated scheduling techniques for solving this problem. We formalized the problem as an MILP and as a CP and used CPLEX and CP Optimizer to solve it. Scheduling GEO-CAPE is amenable to automated scheduling, but the performance on the models suggests that improvements are needed; the problems are too large and complex to be solved optimally in an hour's runtime. The MILP model appears to give better performance than the CP model with respect to schedule quality. With regard to instrument selection, it is clear that schedules for FR have higher quality than those for COEDI. It is difficult for the slower COEDI instrument to meet the science objectives, even when clouds reduce contention for the instrument.

There are several avenues for future work. In discussions with the GEO-CAPE Oceans community, it is plain that the instrument performance characteristics, especially for COEDI, are still in flux. The scene layout was performed manually, and the science community suggested that some scenes could be eliminated from the current COEDI scene layout, although perhaps not enough to offset the slow scan time. The science community has identified a number of different science classes not considered in this paper, including 'target of opportunity' modes, which in general will lead to a more complex scheduling problem. Investigating problems with mixes of Survey and Targeted science will be needed once tradeoffs between the different science types are available. We are also looking at relaxing the assumption that the scanning time for each scene is fixed. Finally, it is notable that there is considerable interest in re-scheduling throughout the day, leading to problems of scheduling a few hours into the future, but with regard to the prior schedule and minimum disruption of customers' expectations. The preliminary results suggest that heuristics or greedy search may be needed to solve the problems more effectively. Alternatively, changes to the MILP and CP models may improve the final solution quality.

Acknowledgement: We acknowledge the contributions of our collaborators in this work: Sasha Weston, Fan Yang Yang, Jason Swenson and Andres Dono Perez of the NASA Ames Research Center Mission Design Center; Clarissa Anderson of the University of California Santa Cruz and Joseph Salisbury of the University of New Hampshire for consultation on GEO-CAPE Ocean science; our high school interns Christina Ho and Umesh Tanniru; Rabindra Palikonda and Patrick Minnis of NASA Langley Research Center for assistance with GOES data; and Antonio Mannino and Karen Moe of NASA Goddard Space Flight Center, the leads of the GEO-CAPE mission and GEO-CAPE mission operations study team.

References

Allahverdi, A.; Ng, C.; Cheng, T. E.; and Kovalyov, M. Y. 2008. A Survey of Scheduling Problems with Setup Times or Costs. *European Journal of Operational Research* 187(3):985–1032.

Brucker, P.; Drexl, A.; Möhring, R.; Neumann, K.; and Pesch, E. 1999. Resource-Constrained Project Scheduling: Notation, Classification, Models, and Methods. *European journal of operational research* 112(1):3–41.

Dechter, R.; Meiri, I.; and Pearl, J. 1991. Temporal Constraint Networks. *Artificial intelligence* 49(1):61–95.

Draper, D. L.; Jonsson, A. K.; Clements, D. P.; and Joslin, D. E. 1999. Cyclic Scheduling. In *The 16th International Joint Conference on Artificial Intelligence*.

Fishman, J.; Iraci, L. T.; Al-Saadi, J.; Chance, K.; Chavez, F.; Chin, M.; Coble, P.; Davis, C.; DiGiacomo, P. M.; Edwards, D.; Eldering, A.; Goes, J.; Herman, J.; Hu, C.; Jacob, D. J.; Jordan, C.; Kawa, S. R.; Key, R.; Liu, X.; Lohrenz, S.; Mannino, A.; Natraj, V.; Neil, D.; Neu, J.; Newchurch, M.; Pickering, K.; Salisbury, J.; Sosik, H.; Subramaniam, A.; Tzortziou, M.; Wang, J.; ; and Wang, M. 2012. The United States' Next Generation of Atmospheric Composition and Coastal Ecosystem Measurements: NASA's Geostationary Coastal and Air Pollution Events (GEOCAPE) Mission.

Han, C.-C., and Lin, K.-J. 1992. Scheduling Distance-Constrained Real-Time Tasks. In *Real-Time Systems Symposium, 1992*, 300–308. IEEE.

Heinz, S., and Beck, J. C. 2012. Reconsidering Mixed Integer Programming and MIP-Based Hybrids for Scheduling. In *Integration of AI and OR Techniques in Contraint Programming for Combinatorial Optimzation Problems*. Springer. 211–227.

Lam, E.; Van Hentenryck, P.; and Kilby, P. 2015. Joint Vehicle and Crew Routing and Scheduling. In *Principles and Practice of Constraint Programming*, 654–670. Springer.

Nicola Policella, Stephen F. Smith, A. C., and Oddi, A. 2014. Generating robust schedules through temporal flexibility. In *ICAPS*.

Pemberton, J. 2000. Towards scheduling over-constrained remote sensing satellites. In *Proceedings of the 2d International Workshop on Planning and Scheduling for Space*.

Pesant, G.; Rix, G.; and Rousseau, L.-M. 2015. A Comparative Study of MIP and CP Formulations for the B2B Scheduling Optimization Problem. In *Integration of AI and OR Techniques in Constraint Programming*.

Verfaillie, G., O. X. P. C. R. S., and Sebbag, I. 2012. Planning acquisitions for an ocean global surveillance mission. In *International Symposium on Artificial Intelligence, Robotics and Automation in Space*.