IEc

An Evaluation of Planning and Data Processes for Cleanup Milestones for the Federal Facilities Response Program

Final Report | November 28, 2008

prepared for:

U.S. Environmental Protection Agency

Federal Facilities Restoration and Reuse Office

prepared by:

Industrial Economics, Incorporated

2067 Massachusetts Avenue

Cambridge, MA 02140

617/354-0074

INDUSTRIAL ECONOMICS, INCORPORATED

TABLE OF CONTENTS

EXECUTIVE SUMMARY

INTRODUCTION AND PURPOSE OF EVALUATION

The U.S. Environmental Protection Agency's Federal Facilities Restoration and Reuse Office (FFRRO) works to improve the timeliness and cost-effectiveness of clean-up activities, and to promote the reuse of contaminated sites at federal facilities. This mission involves coordination with other agencies (e.g., Departments of Defense and Energy) and efforts to track and monitor site activities, primarily (but not exclusively) at sites on the National Priorities List (NPL or Superfund sites) and sites designated for Base Realignment and Closure (BRAC).

The remediation process at these sites involves significant communication among agencies, and the development of formal agreements governing selected remedies, cleanup goals, and schedules. Regional staff use and update the CERCLIS database to track and monitor site remediation activities and progress toward the cleanup goals and schedules that are incorporated into the Federal Facility Agreement for each site. While other agencies typically undertake the direct cleanup activities at federal sites, EPA is the regulatory agency responsible for cleanup oversight and environmental compliance, and ensuring that sites are effectively moving toward remediation.

To fulfill its oversight obligations, EPA Headquarters and regional staff participate in a work planning process in which each region commits to annual targets identifying the number of federal NPL sites that will achieve each of a range of clean-up milestones. The specific milestones that EPA tracks in this planning process are:

- RI/FS (Remedial Investigation and Feasibility Study)/RCRA Facility Investigation Starts
- Decision Documents
- Final Remedy Selected/Final Record of Decision (ROD)
- Remedial Action (RA)/Corrective Measure Implementation (CMI) Starts
- RA/CMI Completions
- Construction Completions
- Sitewide Ready for Anticipated Use (SWRAU)
- Five Year Review Completions

EPA regional staff identify the number of sites, and for some metrics, the names of NPL sites that will achieve each of the milestones outlined above in each fiscal year. HQ staff request and collect the data from regions for use in planning accomplishments and workload.

Both regions and Headquarters rely on the targeting process and CERCLIS information to report progress and guide workload allocation. To help ensure the validity of the process,

the Federal Facilities Response Program has identified two overarching objectives for this evaluation:

- Identify the reasons that certain Federal Facilities Response Program site cleanup milestone targets are not met under current planning and implementation processes.
- Identify processes that regions can put in place to more effectively establish and meet targets. If possible, this would result in a "preferred" work planning process, or "planning formula," that could be adapted to aid planning and implementation across all regions.

This evaluation is designed to focus specifically on the portion of the FFRRO remediation program that relates to the development, monitoring and tracking of site cleanup milestones and targets. The evaluation also focuses specifically on the following: NPL sites; the role of CERCLIS; and regional practices.

METHODOLOGY

To inform the objectives of this evaluation, IEc and EPA collaborated to develop a set of nine specific questions. The questions form the framework for the information collection and analysis. One question addresses CERCLIS data quality efforts, and four questions address issues external to the planning process, including specific site features and regional organization. The questions are:

1. What existing processes do regions have in place for planning and tracking FFRRO site remediation milestones and goals?
2. What factors do regions consider when developing their targets?
3. What quantitative or formulaic methods (if any) are used to develop targets (e.g., does the region have a system where certain calculations are used to evaluate the probability that a site will achieve a milestone)?
4. How is CERCLIS integrated into the FFRRO site remediation goal and milestone planning process?
5. What measures do the regions take to ensure high data quality in CERCLIS?
6. Do sites that present difficulty when meeting targets share any common features (e.g., contaminant types, specific federal agencies)?
7. Do certain specific performance measures present greater problems than others when developing and meeting targets?
8. Do any features unique to a region (e.g., type of sites, number of sites) affect the ability to plan and implement goals?
9. How does regional organization (e.g., how RPMs are assigned, how CERCLIS is maintained) affect what planning processes are in place?

This evaluation employs data from two key sources:

(1) Responses to detailed questions in a series of telephone interviews with selected staff in all EPA regions; and

(2) Data from CERCLIS addressing key issues raised by regional staff in the interviews as factors in the work planning process.

In addition, the evaluation used data from planning cycles for 2006, 2007, and 2008 to assess the accuracy of regional planning. See Appendix A for interview questions and Appendix B for planning cycle data for 2006 – 2008.

Regional Staff Interviews

The central data collection effort for this evaluation consisted of a series of phone and in-person interviews with regional staff. IEc interviewed at least one RPM, one manager, and one IMC from each of EPA's 10 regions. Regions have only one IMC, one or two Federal Facilities Program Managers, and multiple RPMs. Through 35 telephone interviews and two in-person group interviews, IEc interviewed 13 RPMs, 19 managers, and 10 IMCs for a total of 42 respondents.

CERCLIS and Headquarters Data Collection

To evaluate the impact of certain external variables (e.g., the number of active operable units (OUs) and sites in a region), IEc collected a range of data from the CERCLIS database and from the work planning records managed by EPA Headquarters.

The CERCLIS database contains background data on the FFRRO program as well as information regarding number of sites, key site features, and most recent milestones achieved at each OU. In addition, Headquarters maintains CERCLIS audit reports and tracking information about regional targets, achievements, and distribution of staff measured in full time equivalents (FTEs).

Assessment of Regional Performance Using CERCLIS Data

To compare performance across regions, IEc examined four different measures of accuracy by examining the "variance," or differences between the number of targets that regions set for each milestone and the number that they achieved in FY 2006, 2007, and 2008. The types of variance we examine are:

1. **Total variance:** The absolute value of the difference between the total number of achievements and the total number of targets. This metric identifies how many targets were both missed or exceeded; a low total variance is a general indicator of accuracy.
2. **Negative variance**: The total number of targets "missed" across all milestones. This metric isolates the targets that regions failed to meet and does not consider cases where regions exceed targets.
3. **Percent total variance:** The total variance expressed as a percentage of total targets for each region. This metric considers the overall accuracy of a region in meeting

targets, but also considers the number of targets; a region with a large number of targets could therefore have a higher total variance and still report a relatively low percent total variance.

4. **Percent negative variance:** Negative variance expressed as a percentage of total targets. Again, this metric addresses targets "missed" and accounts for the number of targets set.

No single metric is appropriate for identifying high (or low) performing regions in all cases. We therefore do not attempt to develop an overall ranking of regional performance. We consider total variance and total percent variance throughout our analysis, and consider negative variance and percentage of total targets except in cases where we are considering normalizing variables such as number of sites, number of operable units, or FTEs.

Analysis Of The Impact Of External Factors

In addition to the planning process, a range of external factors, such as specific site features or management practices of specific federal agencies, affects the ability of regions to establish accurate targets for achieving cleanup milestones. For example, other federal agencies sometimes have differing requirements for finalizing milestones such as remedial actions. Our analysis of data from interviews and CERCLIS sought to determine whether some or all of the variation in planning performance may be related to external factors outside the control of regional EPA staff.

Note On Limitation Of Statisical Analyses

In the following sections, we discuss the results of our analysis. The two most common forms of results are counts and discussions on correlation tests. Counts are straightforward and are used to provide EPA with an understanding of the frequency with which practices are employed or issues arise. A correlation is used to identify a relationship between two variables. It is important to note that a correlation shows whether a relationship exists between two variables, such as links between specific contaminants at a site and the length of time remedial actions last. It does not determine whether one variable causes change in the other (i.e., causality).

For variables with sufficient quantitative data (e.g., in CERCLIS) we sought through our correlation tests to determine where relationships exist between variables potentially contributing to a region's overall performance. After plotting the two data points for each region, we used a linear regression line to estimate the significance of the relationship between the two variables. A highly-sloped regression line suggests that one variable may have a large impact on the other, but only when the data points are located close to the regression line does the test suggest a high level or correlation between the two variables.

RESULTS

This section of the report summarizes the results of our analyses as outlined in the methodology. We first summarize our assessment of recent regional work planning performance in achieving the targets established during the planning process, and then describe the results of our analysis of each of the nine evaluation questions.

For each evaluation question, we identify the specific interview questions that we use to frame the assessment, and then describe the relationships between regional responses or data and regional performance. Finally, we provide a brief summary of the general insights that our analysis provides into each question.

In addition to considering both total and negative variance, Exhibit 1 documents regional performance both in terms of "total" variance (i.e., actual number of targets missed) and variance expressed as a percentage of total targets. This approach accounts for regions with higher number of targets and identifies cases where low absolute variance is potentially a result of setting limited targets.

EXHIBIT ES-1. TOTAL AND NEGATIVE REGIONAL VARIANCE FROM 2006-2008 PLANNING CYCLES

REGION	TOTAL TARGETS	TOTAL ACCOMPLISHMENTS	TOTAL VARIANCE[A]	PERCENTAGE TOTAL VARIANCE	NEGATIVE VARIANCE	PERCENTAGE NEGATIVE VARIANCE
A	106	126	30	28%	5	5%
B	45	60	15	33%	0	0%
C	132	205	85	64%	6	5%
D	19	23	8	42%	2	11%
E	208	265	81	39%	12	6%
F	23	21	10	43%	6	26%
G	65	54	17	26%	14	22%
H	22	15	15	68%	11	50%
I	52	66	28	54%	7	13%
J	74	50	44	59%	34	46%

Total variance includes all variance. Thus, negative variance is counted in both total variance and negative variance.

In general, two regions perform well in all categories, in that both their total and negative variance is very low (i.e., they achieve the targets they set with some accuracy), whether expressed in raw numbers or as a percentage. When looking only at total variance numbers, four regions have low variance, but they also tend to have fewer sites and/or OUs, and therefore set fewer targets. The number of targets is typically related to both the number and age of regional sites, but is an important factor in considering performance.

If the number of targets is taken into account, then regions with larger total variance can demonstrate relatively strong performance if variance is expressed as a percentage (see Region E, for example). Four regions (Regions A, B, C, and E) show relatively small numbers of targets "missed" and therefore low negative variance, though two of these regions (E and C) have the highest total variance due to "overshooting" many targets. Nevertheless, the number of targets missed for these regions is less than 10, and the percentage of negative variance is also below 10 percent.

Two regions record high percentages "missed" for both total variance and negative variance. Region G shows the lowest total variance (26%) but the 4th highest level of negative variance, suggesting that when the region's planning is wrong, the result is missed targets. Three regions are "in the middle" in terms of both total and negative variance when number of targets is considered.

QUESTION 1: INSIGHTS AND CONCLUSIONS

Responses to questions about the planning process suggest that:

- The overall work planning process is similar across regions, and most regions emphasize that the process is in part developed "organically" as RPMs and managers work together on sites over time and develop informal methods for establishing goals.
- In general, regions that conduct regular planning meetings for RPMs and other staff tend to perform better in meeting targets than regions with less frequent meetings and regions in which RPMs did not provide clear answers to the questions.
- The use of Odometer and e-facts as a method of auditing data corresponded with reasonable performance in two regions who also had relatively high numbers of targets.
- We could not identify any correlation between regional performance and the use of best practices documents or the incorporation of targets into Performance Agreements. However, regions with inconsistent responses regarding the incorporation of targets into Performance Agreements tended to have higher levels of total and negative variance, suggesting that increased emphasis on clear communication of targets could be useful.
- The majority of respondents indicated that training would be helpful. These results suggest that the planning process could be improved from some additional training opportunities, and that respondents in some of the regions that have had the most difficulty accurately planning would likely take advantage of these opportunities.

Question 2: Insights And Conclusions

Responses to questions about considerations for setting targets suggest that:

- Regions identify a range of methods for reducing uncertainty, but did not describe specific approaches. Oddly, however, regions that generally mention being conservative in developing targets report high negative variance (expressed as a percent). It appears, therefore, that more specific approaches than a general conservatism might be more useful in improving targeting.
- Regions that stress coordination with other agencies report lower levels of variance, suggesting that an emphasis on close coordination provides some improved information about the progress at the sites.
- We were unable to identify any correlation between regional planning performance and responses regarding site complexity (e.g., size, remediation type, presence of munitions). In addition, regions appear to have developed effective methods for coordinating with legal and technical resources.

Question 3: Insights and conclusions

Responses to questions about quantitative or formulaic methods for accounting for uncertainty in setting targets suggest that only three regions have employed quantitative methods. Methods differ by region, and at least one of these regions no longer uses the approach.

Performance among regions that do report quantitative approaches to adjusting targets varies widely. The region incorporating a 10 percent "rule of thumb" adjustment is high performing, but the region reporting that it previously used a 50 percent adjustment is surprisingly one of the regions with a relatively large number of "missed" goals (high negative variance). Finally, the region with a relatively sophisticated quarterly adjustment approach has a mixed performance, with a significant number of "missed" targets but also a large number of overall targets, resulting in a reasonable performance when variance is expressed as a percentage.

Question 4: Insights and conclusions

Responses to questions about the integration of CERCLIS into the planning process suggest that:

- There are varying opinions about how important CERCLIS is for measuring job performance as it pertains to accurate and achieved targets. It is notable that regions with high levels of variance and respondents that do not believe CERCLIS is used to set targets appear to be less likely to believe that CERCLIS is important for measuring job performance.
- We were not able to identify a correlation between variance and how CERCLIS is integrated, whether targets are based on data in CERCLIS, and the reported frequency of planning meetings. There is, however, a relationship between

inconsistent responses to the question about targets being based on data in CERCLIS and whether performance (i.e., high variance).

Question 5: Insights and conclusions

The table below summarizes each region for the following:

- Use of Management Review Function
- Regular CERCLIS Meetings
- Mention of Audit Frequency
- RPM Data Entry

Two of the regions with the best performance did answer "yes" to the most questions, and the two regions with the highest variance answered "yes" least often. It is important to note, however, that with the exception of one region at each end of the distribution, the answers did not vary much across regions.

EXHIBIT ES-2. SUMMARY OF CERCLIS CHARACTERISTICS

REGION	MANAGEMENT REVIEW FUNCTION?	REGULAR CERCLIS MEETINGS?	ARE THERE FREQUENT AUDITS?	DO RPMS ENTER DATA?	COUNT OF "YES'"
A	Yes	Yes	Yes	Yes	4
B	Yes	Yes	Unknown	Yes	3
C	Yes	No	Yes	Yes	3
D	Yes	Yes	Yes	Some	3
E	No	Yes	Yes	Some	2
F	No	Yes	Yes	Some	2
G	No	Yes	Yes	No	2
H	Yes	No	Unknown	Yes	2
I	Yes	No	Yes	No	2
J	No	No	Unknown	Yes	1

Based on responses to questions about CERCLIS data quality, we were unable to identify a correlation between variance and measures taken to ensure high data quality, the process for updating milestones, use of management review function, regular CERCLIS meetings, and data entry. However, it does appear that regions that clearly integrate multiple data quality activities into their processes general achieve better results in their planning process.

In addition, regions that provided consistent responses had lower variance than those regions that provided inconsistent responses between staff. Frequency of internal coordination and communication may therefore be a significant driver of performance, regardless of which specific data quality actions are implemented.

Question 6: Insights and conclusions

Responses to questions about common features among sites that present difficulty suggest that:

- In general, sites identified as "difficult" sites are now in late stages of the pipeline; 61 percent of all of the difficult sites identified by interviewees report completion of RA work at an OU as the most recent action.
- Identified difficult sites have Army and Air Force as the lead agency at a slightly higher frequency than other federal sites, but the relationship between difficult sites and agencies is not strong.
- "Difficult" sites are slightly less likely to have groundwater contamination than other sites, in contrast to responses from interviews about which site features most complicate planning. Conversely, difficult sites were more likely to have "other" contaminated media than other sites, suggesting that unique or uncommon contamination may be a contributor to creating a difficult site.
- Our assessment of the relationship between sites with "difficult site" features and overall regional performance in work planning did not identify any correlations.

In general, our investigation of the features of difficult sites did not isolate any specific site features that could be used to help regions predict what other sites will present planning difficulties.

Question 7: Insights and conclusions

Responses to questions about performance measures that present greater problems than others suggest that:

- A relationship exists between the total number of targets and total variance, but it does not appear to be linked to more "missed targets" (i.e., higher negative variance). Thus, having a greater number of targets does not necessarily predict more missed targets, but rather indicates a greater general uncertainty regarding the number of targets that can be accomplished in a year.
- Data from interviews and data from CERCLIS on the types of milestones that are most often missed appear slightly inconsistent, with interview respondents clearly noting RODs and Construction Completes as the most difficult milestones to plan for (third most frequent response was that all measures are equally difficult). In contrast, CERCLIS data documents that total Decision Documents (including RODs and other decision documents) frequently lag their schedules, but that Five-Year Reviews and RI/FS starts have historically had the most "missed schedules" and Five-Year Reviews have the longest average lag time, at over one year. Notably, however, in 2008 RI/FS starts and Five-Year Review completions were not missed in significant numbers.
- In general, the number of targets set for any specific milestone does not appear to be linked to overall regional performance. While we noted a strong correlation

between RI/FS Starts and total variance, this is difficult to interpret because negative variance (i.e., "missed" targets) is not strongly related.

Question 8: Insights and conclusions

Evaluation of the relationship between different regional variables and regional performance reveals the following:

- The "size" of a region's workload may be a factor in performance. The relationship between total OUs per FTE and the negative variance is positive, though the relationship is not very strong. More tellingly, however, EPA's workload model that adjusts for the level of effort required at sites in different stages of the Superfund pipeline, identifies several regions with estimated staffing "shortfalls" for 2008. These include regions that tend to report higher levels of total variance and in some cases negative variance.
- Regions where respondents mentioned good relationships with states tended to have lower levels of variance, suggesting that attention to relationships with states, and careful consideration of targets in states where relationships are difficult, may assist in planning.
- Correlation tests on recent pipeline actions at a site level and negative variance demonstrated a strong correlation between the number of sites with RA as the most recent pipeline action and negative variance (missed targets). This relationship seems inconsistent with respondent assertions that RODs are the most difficult pipeline stage to plan, but site-level actions may obscure difficulties associated with larger sites with multiple OUs. A more detailed analysis of actions at the OU level could clarify any relationships between planning and "pipeline stage."
- There is no national correlation between performance and lead agency, or performance and geographical area of regions.

Question 9: Insights and conclusions

Responses to questions about regional organization suggest that regional organization into multiple branches with federal sites, and distribution of workload among RPMs, do not have a significant impact on overall performance across regions.

CONCLUSIONS AND RECOMMENDATIONS

Relationship between Current Federal Facilities Response Program Work Planning Procedures and Regional Performance (Questions 1 through 3):

- Most regions have developed planning approaches over time, and have organically integrated the experience of staff and informal but well understood methods of coordination and data quality control.
- Clarity in communicating and implementing the planning process appears to be a key factor in ensuring a high level of regional performance.
- Current attempts to account for uncertainty in setting targets have limited success.

Role of CERCLIS Integration in Federal Facilities Response Program Work Planning Procedures and Regional Performance (Questions 4 and 5):

- Most regions describe CERCLIS as central to the planning process, and report using CERCLIS to both set targets and audit progress.
- Regions that report systematic efforts to ensure CERCLIS data quality tend to have better performance than other regions.

Impact of External Factors on Regional Work Planning Performance (Questions 6 trhough 9)

- The investigation of key features at difficult sites did not identify any features with strong links to overall regional performance in work planning.
- While data suggest that specific site milestones have different rates of success in achieving timely completion, we did not identify a simple relationship between number of targets for specific milestones and regional performance.
- Regional variables, including staffing level, relationship with states, and potentially the "stage" of a site in the pipeline, appear to have a limited relationship with regional planning performance. Other issues (geographic area, lead agency at each site, regional branch organization, and RPM assignments across federal and private sites) do not appear related to performance.

RECOMMENDATIONS

Based on the conclusions and insights from the data collection and analysis in this evaluation, we outline the following near-term and "stretch" recommendations. These recommendations are designed to respond to the second evaluation objective of identifying processes and methods that can be implemented to improve the effectiveness of the work planning process. Note that we incorporate suggestions for addressing external variables that may affect performance into the work planning procedure recommendations.

Work Planning Procedure Recommendations:

- **Headquarters/Regions: Develop and Deliver Training on Work Planning.** Additional training received interest among respondents in most regions, including all regions where planning has been a challenge. While training should reflect specific regional structure and approach, Headquarters could contribute to this process.
- **Headquarters/Regions: Develop a Planning "cheat sheet."** As part of training or as an alternative to training, develop a brief "trouble-shooting guide" for planning.
- **Regions: Improve communication about importance, structure of work planning process by:**
    - Clearly incorporating targets into all Performance Agreements (managers, RPMs, IMCs) where applicable. This clarifies expectations and raises priorities; and
    - Developing a formal schedule for communication during planning process. Based on the results of our analysis, the most successful regions have at least quarterly meetings for managers, RPMs, and IMCs that address planning, though in some cases the meetings are one-on-one meetings between managers and RPMs.
- **Headquarters/Regions: Account for Uncertainty.** Develop and implement a process for incorporating adjustments to targets to reflect significant uncertainty related to certain sites/actions. Requires additional analysis to determine appropriate adjustments, but consistent with regional reports, system could include probability adjustments for completing targets.
- **Stretch Recommendation - Headquarters: Rationalize Budget Allocation with Performance**. Develop a budget scoring system that differentially gives positive or negative credits for regional performance accounting for:
    - Total number of targets;
    - Total number of "missed" targets;
    - Total number of exceeded targets;
    - Differential credits for RODs, CCs, and Five-Year Reviews;
    - Differential credits for regions with low weighted FTE/active OU ratios.

CERCLIS Integration and Data Quality Recommendations:

- **Regions:** Develop document clarifying regional chain of responsibility for CERCLIS data quality and, to the extent possible, incorporate this policy into Performance Agreements. The policy should include one or more of the following:

  - Implementation of management review function;
  - Periodic audits of CERCLIS data;
  - Regular meetings between RPMs, managers and IMCs to rationalize/update CERCLIS data (informal audits; can be incorporated into planning meetings).

- **Regions:** Encourage monthly, rather than quarterly, updates from RPMs to facilitate expertise in using the system and increased awareness of site status and relevant events.
- **Stretch Recommendation – Headquarters:** Review and coordinate all HQ reporting requirements for CERCLIS to limit burden on regional staff; goal should be to reduce reporting requirements and time by a significant percentage, in order to improve focus on data quality for remaining reports.
- **Stretch Recommendation - Headquarters:** Respondents to interviews suggested several changes to CERCLIS, including:
  - Improve the transparency of the database locking feature as targets get locked in and RPMs have difficulty identifying and changing locked fields;
  - Allow for entry of schedules and planned dates in CERCLIS;
  - Create memo fields for notes as some sites may exist for many years;
  - Integrate project management or project tracking tools into CERCLIS.

INTRODUCTION AND PURPOSE OF EVALUATION

The U.S. Environmental Protection Agency's Federal Facilities Restoration and Reuse Office (FFRRO) works to improve the timeliness and cost-effectiveness of clean-up activities, and to promote the reuse of contaminated sites at federal facilities. This mission involves coordination with other agencies (e.g., Departments of Defense and Energy) and efforts to track and monitor site activities, primarily (but not exclusively) at sites on the National Priorities List (NPL or Superfund sites) and sites designated for Base Realignment and Closure (BRAC).

Consistent with input received from the Office of Management and Budget's (OMB's) Program Assessment Rating Tool (PART) exercise, FFRRO is strengthening its use of program evaluation methods with a targeted evaluation of the Federal Facilities Response Program (the "program").

One area of focus for the program has been the effectiveness of the planning process used to establish annual and long-term cleanup targets at federal sites. Clean-up targets for the program are linked to a series of internal cleanup milestones which are tracked in the Comprehensive Environmental Response, Compensation, and Liability Information System (CERCLIS).

EPA regions establish targets annually for the number of sites that will meet each milestone; the national program must then demonstrate that the targets have been met. In recent years, regions have achieved varying rates of success in setting and meeting targets.

Complicating this analysis is the inclusion of inaccurate data in CERCLIS which create unrealistic target dates for achieving site cleanup milestones for federal sites. The reliability of these data and targets is important to EPA and other agencies involved with the cleanup of these sites.

To address these issues, FFRRO has charged Industrial Economics, Incorporated (IEc) with completing an independent evaluation that focuses specifically on the effectiveness of the regional target-setting process. This evaluation is designed to support the development of an approach that will assist regions with optimizing their target-setting activities. This report presents the methodology used, results, and conclusions of the evaluation, along with a detailed set of recommendations.

OVERVIEW OF THE FEDERAL FACILITIES RESPONSE PROGRAM WORK PLANNING PROCESS

FFRRO works in conjunction with the Department of Defense, the Department of Energy and other federal agencies to provide efficient, creative, and cost-effective solutions to environmental contamination at federal facilities. As part of its program, FFRRO coordinates, tracks, and monitors site cleanup efforts among federal agencies at federal NPL sites. The remediation process at these sites involves significant communication

among agencies, and the development of formal agreements governing selected remedies, cleanup goals, and schedules. Regional federal facility staff use and update CERCLIS to track and monitor site remediation activities and progress toward the cleanup goals and schedules that are incorporated into the Federal Facility Agreement (FFA) for each site. While other agencies typically undertake the direct cleanup activities at federal sites, EPA is the regulatory agency responsible for cleanup oversight and environmental compliance, and ensuring that sites are effectively moving toward remediation.

To fulfill its oversight obligations, EPA Headquarters and regional staff participate in a work planning process in which each region commits to annual targets identifying the number of federal NPL sites that will achieve each of a range of clean-up milestones. The specific milestones that EPA tracks in this planning process are:

- RI/FS (Remedial Investigation and Feasibility Study)/RCRA Facility Investigation Starts
- Decision Documents
- Final Remedy Selected/Final Record of Decision (ROD)
- Remedial Action (RA)/Corrective Measure Implementation (CMI) Starts
- RA/CMI Completions
- Construction Completions
- Sitewide Ready for Anticipated Use (SWRAU)
- Five Year Review Completions

The national-level sum of the regional targets for each milestone then becomes part of EPA's performance for the Superfund program as reported under the Government Performance and Results Act (GPRA) and as part of the agency's Strategic Planning process.

EPA regional staff identify the number of sites, and for some metrics, the names of NPL sites that will achieve each of the milestones outlined above in each fiscal year and report back to HQ staff who request and collect the data for use in planning accomplishments and workload.

The planning process includes estimates with a two-year lead-time (i.e., regions first estimated fiscal year (FY) 2008 targets at the end of FY 2006). These initial estimates are adjusted and finalized at the end of the fiscal year prior to the target fiscal year – at the same time as regions are providing the final documentation for the current year targets. Therefore, as of the end of September 2008, regions have both reported on their FY 2008 achievements and finalized their FY 2009 targets.

The final number of federal NPL sites that meet the established targets is reported on a national basis along with Office of Superfund Remediation and Technology Innovation (OSRTI) achievements. EPA's Administrator and external reviewers such as Office of Management and Budget through its Program Assessment Rating Tool (PART) evaluate these results. In this context, "missed" targets can present a significant issue for FFRRO

and OSRTI. For example, in 2007 FFRRO missed its national target for the number of sites achieving Construction Completions.

PURPOSE OF THE EVALUATION

In recent years, results of the Federal Facilities Response Program's work planning process have revealed that EPA regions have had differing levels of success in establishing and meeting targets under this system. While some regions are reasonably successful at meeting the targets they set, other regions have significantly over- or under-estimated the number of sites that will achieve milestones.

In addition, efforts by the Federal Facilities Response Program to analyze the reporting process have identified a few data quality issues in CERCLIS that may be complicating the work planning process. For example, in some cases, site information is not updated, resulting in data discrepancies that either fail to document site progress or do not reflect changes in schedules that delay completion of milestones at the site.

Both regions and Headquarters rely on the targeting process and CERCLIS information to report progress and guide workload allocation. To help ensure the validity of the process, the Federal Facilities Response Program has identified two overarching objectives for this evaluation:

- Identify the reasons that certain Federal Facilities Response Program site cleanup milestone targets are not met under current planning and implementation processes.
- Identify processes that regions can put in place to more effectively establish and meet targets. If possible, this would result in a "preferred" work planning process, or "planning formula," that could be adopted to aid planning and implementation across all regions.

More specifically, this evaluation focuses on assessing the factors that affect planning (including planning processes, CERCLIS data quality procedures, and external factors), and identifying factors that may improve planning.

This section describes in detail the methodology that IEc has undertaken to perform this evaluation. We first provide a description of the evaluation scope and questions, and then outline a two-pronged assessment of the relationship between both regional practices and key site features with regional planning performance in setting and meeting targets for achieving Federal Facilities Response Program site milestones.

In addressing the evaluation objectives outlined above, it is important to first define the scope and focus of the evaluation by outlining key definitions and evaluation questions.

EVALUATION SCOPE

This evaluation is designed to focus specifically on the portion of the Federal Facilities Response Program that relates to the development, monitoring and tracking of site cleanup milestones and targets.

The evaluation also focuses specifically on the following:

- **NPL Sites:** While federal site cleanup includes both NPL and non-NPL sites, performance measures and milestones reported under GPRA and PART include only sites on the National Priorities List. While this evaluation considers the impact that non-NPL site management has on the regional site planning processes, the evaluation focuses on planning related to NPL sites.
- **The role of CERCLIS:** CERCLIS is the central tracking system for the Superfund program, including both federal and non-federal sites. Federal facilities staff wish to optimize the role of CERCLIS as both a planning and tracking tool, and the evaluation examines opportunities for implementing practices, staff, and systems that most effectively use and review this data set.
- **Regional practice:** Regional level systems and planning efforts are the core of the Federal Facilities Response Program work planning process. This evaluation focuses primarily on EPA regional staff, including regional Federal Facilities Program Managers, Information Management Coordinators (IMCs), Remedial Project Managers (RPMs), and the systems and approaches used by regions that appear to support the most effective work planning efforts.

EVALUTION METHODOLOGY

EVALUATION QUESTIONS AND GENERAL APPROACH

To inform the objectives of this evaluation, IEc and EPA collaborated to develop a set of nine specific questions. The questions form the framework for the information collection and analysis. One question addresses CERCLIS data quality efforts, and four questions address issues external to the planning process, including specific site features and regional organization. The questions are:

1. What existing processes do regions have in place for planning and tracking FFRRO site remediation milestones and goals?
2. What factors do regions consider when developing their targets?
3. What quantitative or formulaic methods (if any) are used to develop targets (e.g., does the region have a system where certain calculations are used to evaluate the probability that a site will achieve a milestone)?
4. How is CERCLIS integrated into the FFRRO site remediation goal and milestone planning process?
5. What measures do the regions take to ensure high data quality in CERCLIS?
6. Do sites that present difficulty when meeting targets share any common features (e.g., contaminant types, specific federal agencies)?
7. Do certain specific performance measures present greater problems than others when developing and meeting targets?
8. Do any features unique to a region (e.g., type of sites, number of sites) affect the ability to plan and implement goals?
9. How does regional organization (e.g., how RPMs are assigned, how CERCLIS is maintained) affect what planning processes are in place?

Finally, the evaluation builds on the relationships between regional performance and regional planning or site features. The goal is to identify potentially successful practices that may be valuable across regions and factors (such as site features) that may require consideration in planning and assessing regional performance. The remainder of this section describes in more detail the data collection and analytic approaches.

DATA COLLECTION APPROACH

This evaluation employs data from two key sources:

Responses to detailed questions in a series of telephone interviews with selected staff in all EPA regions; and

Data from CERCLIS addressing key issues raised by regional staff in the interviews as factors in the work planning process.

In addition, the evaluation used data from planning cycles for 2006, 2007, and 2008 to assess the accuracy of regional planning. See Appendix A for interview questions and Appendix B for planning cycle data for 2006 – 2008.

Regional Staff Interviews

The central data collection effort for this evaluation consisted of a series of phone and in-person interviews with regional staff. IEc interviewed at least one RPM, one manager, and one IMC from each of EPA's 10 regions. Regions have only one IMC, one or two Federal Facilities Program Managers, and multiple RPMs. Through 35 telephone interviews and two in-person group interviews, IEc interviewed 13 RPMs, 19 managers, and 10 IMCs for a total of 42 respondents.

EPA identified specific individuals to participate in the interviews, ensuring a range of experience and knowledge of the planning processes. This approach maximized the likelihood that a limited number of interviews will result in collection of high-quality data, but limited the use of a statistical analysis because the interviewees were not randomly selected.

Interview Format

The interviews were organized into three general topic areas that were mapped to the evaluation questions:

1. **Data/perceptions on region's planning and implementation processes:** This set of open-ended questions was designed to elicit a description of the current regional planning process. More detailed questions were designed to address evaluation questions 1-3 and included the following:
    - What has been your experience with the planning process (e.g., how well has the process worked?)?
    - Describe your role in the planning process.
    - How is your planning process implemented (e.g., what staff, meetings, processes are implemented)?
    - How much effort does the planning process require?
    - How does your region develop targets and cleanup dates?
    - Are the targets based on any quantitative or formulaic methods?

2. **Regional uses and interpretation of CERCLIS data:** This set of more standardized questions was designed to elicit information about current data management practices and inform evaluation questions 4 and 5:
    - How is CERCLIS integrated into the planning process?
    - Do you update all milestones on a site when targets/goals shift?
    - How can CERCLIS be used to make the planning process more effective and efficient?

- How do you ensure the quality of data in CERCLIS?
- Are regular CERCLIS meetings held or other formal processes used to update the data?

3. **Data/perceptions on planning issues:** This set of both standard and open-ended questions was designed to elicit information about both internal limitations in regional processes and external issues (e.g., site features) that can affect the planning process; the questions inform evaluation questions 6 through 9:

   - Are there common characteristics among sites that present difficulty when meeting targets?
   - Do certain specific performance measures (e.g., beginning or ending Remedial Actions, publishing RODs) present greater problems than others when developing and meeting targets?
   - How does regional organization limit/direct what planning processes are in place?
   - Do you have a list of "difficult" sites?
   - Do any features unique to your region affect the ability to plan and implement goals (e.g., office organization, staffing levels, geography, distribution of responsibility across NPL/non-NPL or federal/non-federal sites)?

Implementation Procedure

Before conducting phone interviews, IEc tested the interview questionnaire in an in-person meeting with several regional staff in Region 1, including two managers (one a former RPM) and an IMC. The test served to both collect regional information and to gauge the effectiveness of the interview format by identify potential issues and/or revise interview questions. Region 1 served as a test region both because of its proximity to IEc and because the region's record of effectiveness in setting and meeting cleanup milestone targets suggested that regional staff could potentially provide insights into the factors that determined planning success.

Following the interview and feedback from Region 1, IEc administered a revised questionnaire to staff in remaining regions by telephone, with the exception of Region 5; IEc staff were on site and interviewed key staff in person for convenience. IEc distributed the evaluation questions to regional staff prior to calls. After interviews were complete, IEc administered a short follow-up survey by e-mail to assist in documenting specific "difficult" sites that have missed one or more specific targets in recent years.

CERCLIS and Headquarters Data Collection

To evaluate the impact of certain external variables (e.g., the number of active operable units (OUs) and sites in a region), IEc collected a range of data from the CERCLIS database and from the work planning records managed by EPA Headquarters.

The CERCLIS database contains background data on the FFRRO program as well as information regarding number of sites, key site features, and most recent milestones achieved at each OU. In addition, Headquarters maintains CERCLIS audit reports and tracking information about regional targets, achievements, and distribution of staff measured in full time equivalents (FTEs).

Data from CERCLIS and Headquarters used to frame the analysis included the following information:

- Regional work planning (site milestone) targets, mid-year updates, and achievements for fiscal years 2006 - 2008;
- Descriptive information from CERCLIS about federal NPL sites, including data on location, lead agency, media contaminated, number of OUs, and recent milestones recorded for each OUs. In addition, IEc identified sites with specific contaminants (e.g., Military Munitions Response Program (MMRP) sites) and regional features such as total NPL sites and non-NPL federal sites; and
- Data on the FTEs associated with federal facilities in each region, and internal EPA Outyear Cost Model staff allocation analyses estimating the appropriate allocation of staff across sites at different stages in the "pipeline."

DATA ANALYSIS

Using the data collected as outlined above, IEc:

- Evaluated regional performance in the FY 2006 – FY 2008 planning cycles;
- Investigated the relationship between regional planning and data quality processes and regional performance; and
- Investigated the relationship between external factors and regional performance.

Assessment of Regional Performance Using CERCLIS Data

To compare performance across regions, IEc examined four different measures of accuracy by examining the "variance," or differences between the number of targets that regions set for each milestone and the number that they achieved in FY 2006, 2007, and 2008. The types of variance we examine are:

1. **Total variance:** The absolute value of the difference between the total number of achievements and the total number of targets. This metric identifies how many targets were both missed or exceeded; a low total variance is a general indicator of accuracy.

2. **Negative variance**: The total number of targets "missed" across all milestones. This metric isolates the targets that regions failed to meet and does not consider cases where regions exceed targets.
3. **Percent total variance:** The total variance expressed as a percentage of total targets for each region. This metric considers the overall accuracy of a region in meeting targets, but also considers the number of targets; a region with a large number of targets could therefore have a higher total variance and still report a relatively low percent total variance.
4. **Percent negative variance:** Negative variance expressed as a percentage of total targets. Again, this metric addresses targets "missed" and accounts for the number of targets set.

Analysis of The Regional Planning Process

A central portion of this evaluation is a review of the different processes used by regions to develop estimates and targets for site cleanup milestones as determined through interviews. Work planning processes are defined by both regional personnel organization (e.g., assignment of responsibility for accurate planning and data tracking) and specific planning activities (e.g., meetings to review and evaluate targets, quantitative adjustments to targets to account for uncertainty).

To investigate these processes, we coded the relevant interview responses to provide results meaningful to the evaluation questions. We coded interview responses to be uniform to allow comparison and then categorized or grouped responses in ways that most effectively supported the objectives of the analysis. For example, in answering one question about meeting frequency, we defined all meetings that occurred quarterly or more often as "frequent." Specifically, we coded responses to identify:

- Differences in regional practices- grouped to reflect useful distinctions (e.g. regular meetings vs. infrequent meetings).
- Regions where significant internal disagreement/confusion is evident in responses (e.g., two respondents in the same region provided conflicting information on how often meetings occur).

In several cases, we referred to information in a range of questions to verify and refine our coding. For example, we examined whether respondents' specific positions, or regional structure (e.g., the number of sections within federal facilities), would clarify the answers that were given about the structure of the planning process.

We then compared the coded results to regional performance from CERCLIS to identify practices that are highly correlated with success (or lack of success). Note, however, that because most data for this analysis were collected through interviews in which interviewees are not selected randomly, the approach supports only limited, descriptive correlation analysis.

Analysis Of The Impact Of External Factors

In addition to the planning process, a range of external factors, such as specific site features or management practices of specific federal agencies, affects the ability of regions to establish accurate targets for achieving cleanup milestones. For example, other federal agencies sometimes have differing requirements for finalizing milestones such as remedial actions. Our analysis of data from interviews and CERCLIS sought to determine whether some or all of the variation in planning performance may be related to external factors outside the control of regional EPA staff.

Note On Limitation Of Statisical Analyses

In the following sections, we discuss the results of our analysis. The two most common forms of results are counts and discussions on correlation tests. Counts are straightforward and are used to provide EPA with an understanding of the frequency with which practices are employed or issues arise. A correlation is used to identify a relationship between two variables. It is important to note that a correlation shows whether a relationship exists between two variables, such as links between specific contaminants at a site and the length of time remedial actions last. It does not determine whether one variable causes change in the other (i.e., causality).

For variables with sufficient quantitative data (e.g., in CERCLIS) we sought through our correlation tests to determine where relationships exist between variables potentially contributing to a region's overall performance. After plotting the two data points for each region, we used a linear regression line to estimate the significance of the relationship between the two variables. A highly-sloped regression line suggests that one variable may have a large impact on the other, but only when the data points are located close to the regression line does the test suggest a high level or correlation between the two variables.

RESULTS

This section of the report summarizes the results of our analyses as outlined in the methodology. We first summarize our assessment of recent regional work planning performance in achieving the targets established during the planning process, and then describe the results of our analysis of each of the nine evaluation questions.

For each evaluation question, we identify the specific interview questions that we use to frame the assessment, and then describe the relationships between regional responses or data and regional performance. Finally, we provide a brief summary of the general insights that our analysis provides into each question.

RECENT PERFORMANCE IN THE PLANNING PROCESS

Our analysis considers information that might explain success or difficulty in setting targets and achieving targeted milestones. Exhibit 1 summarizes the total and negative variance for each region for the three annual planning cycles that began with identification of targets in September 2005 and completion of achievements by September 30, 2008. High performing regions are those with the lowest variance.

Note that variance can be calculated in several additional ways (e.g., positive variance only) that could potentially be useful to the Federal Facilities Response Program. However, it is most important to consider total variance (overall inaccuracy) and negative variance (missed targets only) in the context of GPRA and PART reporting.

In addition to considering both total and negative variance, Exhibit 1 documents regional performance both in terms of "total" variance (i.e., actual number of targets missed) and variance expressed as a percentage of total targets. This approach accounts for regions with higher number of targets and identifies cases where low absolute variance is potentially a result of setting limited targets. Note that total variance is not the same as total accomplishments less total targets. Variance is calculated using that formula; however it is done so on a target type by target type basis, not cumulatively across each region. For example, if Region X targeted 10 construction completions and 5 final RODs and accomplished 5 and 10, respectively, Region X's total variance would be the sum of the absolute values of (10-5) and (5-10) = 10, even though it had the same number of total accomplishments and targets.

EXHIBIT 1. TOTAL AND NEGATIVE REGIONAL VARIANCE FROM 2006-2008 PLANNING CYCLES

REGION	TOTAL TARGETS	TOTAL ACCOMPLISHMENTS	TOTAL VARIANCE[A]	PERCENTAGE TOTAL VARIANCE	NEGATIVE VARIANCE	PERCENTAGE NEGATIVE VARIANCE
A	106	126	30	28%	5	5%
B	45	60	15	33%	0	0%
C	132	205	85	64%	6	5%
D	19	23	8	42%	2	11%
E	208	265	81	39%	12	6%
F	23	21	10	43%	6	26%
G	65	54	17	26%	14	22%
H	22	15	15	68%	11	50%
I	52	66	28	54%	7	13%
J	74	50	44	59%	34	46%

Total variance includes all variance. Thus, negative variance is counted in both total variance and negative variance.

Regional performance varied significantly across all categories of variance. In addition, Appendix B provides separate tables sorted by each of the four types of variance, to provide easier examination of regional performance. Note that while regional performance varied across the three years examined, we did not notice any significant trends toward improved accuracy in any regions, and the "ranking" of regional performance did not vary significantly from year to year. We therefore consider overall regional performance across all three reporting years.

In general, two regions perform well in all categories, in that both their total and negative variance is very low (i.e., they achieve the targets they set with some accuracy), whether expressed in raw numbers or as a percentage. When looking only at total variance numbers, four regions have low variance, but they also tend to have fewer sites and/or OUs, and therefore set fewer targets. The number of targets is typically related to both the number and age of regional sites, but is an important factor in considering performance.

If the number of targets is taken into account, then regions with larger total variance can demonstrate relatively strong performance if variance is expressed as a percentage (see Region E, for example). Four regions (Regions A, B, C, and E) show relatively small numbers of targets "missed" therefore low negative variance, though two of these regions (E and C) have the highest total variance due to "overshooting" many targets. Nevertheless, the number of targets missed for these regions is less than 10, and the percentage of negative variance is also below 10 percent.

Two regions record high percentages "missed" for both total variance and negative variance. Region G shows the lowest total variance (26%) but the 4th highest level of negative variance, suggesting that when the region's planning is wrong, the result is

missed targets. Three regions are "in the middle" in terms of both total and negative variance when number of targets is considered.

No single metric is appropriate for identifying high (or low) performing regions in all cases. While it would be possible to develop a single scoring system that combines variance (and potentially other features of the regional programs), this would require careful attention from Headquarters and regional staff to ensure that each part of the score is properly identified and weighted. We therefore do not attempt to develop an overall ranking of regional performance. We consider total variance and total percent variance throughout our analysis, and consider negative variance and percentage of total targets except in cases where we are considering normalizing variables such as number of sites, number of operable units, or FTEs.

The remainder of this section examines the relationship between different responses to interview questions and regional performance. Although we refer to the results noted above in defining regional performance, we do not identify specific regions; our aim is to identify the practices and features of planning processes that are most strongly linked to success in planning across multiple regions.

EVALUATION QUESTION 1.

What Existing Processes Do Regions Have In Place For Planning And Tracking Federal Facilities Response Program Site Remediation Milestones And Goals?

To evaluate which existing processes appear to lead to an effective planning process, IEc considered the relationship between performance and answers to the following interview questions:

- Do regions employ formal planning processes described in written regional guidance or "best practices" documents?
- What systems do regions use for measuring and tracking progress toward goals?
- How do regions foster communication among staff in the planning process? To address this question, we examine the frequency and regularity of regional meetings with staff to address planning and target-setting.
- Do regions incorporate planning and targets into Performance Agreements?
- Do regions provide/utilize training for the planning process?

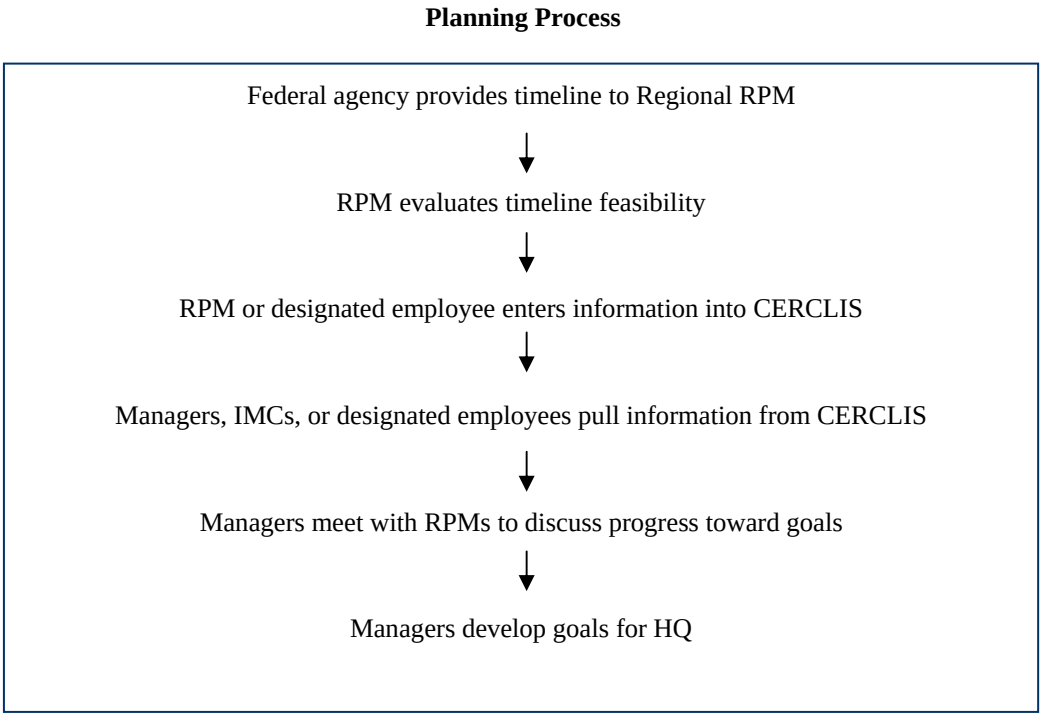
Below we outline responses to these questions and note areas where responses appear to be correlated with low variance (e.g., high accuracy) in the planning process.

Note On General Description of the Planning Process

Based on our interviews, the general description of the planning process in all regions is a variable system in which RPMs, managers, and IMCs coordinate in different ways depending on regional organization, site needs, and management style.

Still, regions did report certain consistent details. For example, all regions reported that CERCLIS is used as a foundation for planning and tracking progress toward goals. All of the regional representatives responded that their targets are based on data in CERCLIS, which is mapped to the milestones and schedule in the Federal Facility Agreements (FFA's) in place at the sites. Respondents indicated that their regions use audit reports to monitor progress and data quality – in some cases regions generate their own audit reports, but many use Headquarters' reports. Other specific reports mentioned include Superfund Comprehensive Accomplishment Plans (SCAP) reports, including SCAP 2, 14, and 15 as well as the "Odometer" and e-facts. Exhibit 2 presents the general planning process identified by all regions.

EXHIBIT 2. GENERAL OVERVIEW OF THE REGIONAL FEDERAL FACILITIES RESPONSE PROGRAM PROJECT PLANNING PROCESS



The implementation of this process differed by region, but most regions stressed that direct, informal communication between RPMs and federal agency representatives, and between RPMs and managers, is important to setting appropriate targets.

Use of Formal Documentation for Planning

In terms of documentation of the planning processes, it does not appear that regions employ written guidance or best practices documents, and it is therefore impossible to identify any link between these materials and regional planning performance. Every region responded that the current process does *not* include a manual or a formal best practices document. However, one region noted that a document had been previously developed. This document is not currently used, though the developer still has a central role in planning.

Measuring and Tracking Progress

The majority of regions responded that they rely on CERCLIS audit reports to track progress toward their goals and for quality control of their data. Frequently mentioned audit reports include:

- e-facts reports
- The "Odometer" in e-facts
- SCAP reports

The most frequently cited report is the SCAP 14. We could identify a correlation with the use of SCAP reports and variance, in part because all but one region mentioned the use of SCAP reports. Furthermore, since SCAP reports are based on CERCLIS data, their use in planning may be affected by inaccurate data.

Respondents in two regions mentioned that they use e-facts and the Odometer; both of these regions report relatively low variance when expressed as a percentage of targets missed. However, these regions also report relatively high numbers of targets. It is therefore unclear whether these additional measures contribute to improving accuracy, or merely indicate that regions with higher numbers of targets require more effort for the planning process and find that multiple tools can assist this process.

Frequency of Communication

While regions reported a wide range of communication types and frequency in the planning process, we focused on the role of regular meetings in supporting the planning process. In describing the planning process, respondents of one region noted meetings and check-ins as frequent as twice weekly, while respondents in another region noted formal planning discussions occurring only once a year, with limited preparation for the mid-year progress meeting as well.

To evaluate the relationship between the regularity of planning meetings and regional performance in meeting targets, we coded all responses and categorized regions into one of four groups: regions where all staff meet regularly (i.e., at least quarterly); regions where only managers meet regularly, with ad hoc meetings for RPMs, regions in which planning meetings are limited to once or twice a year, and finally, one region with conflicting information from different respondents. Exhibit 3 provides an example of the raw coded data that we used indicating the number of responses to the interview question.

When we compared responses to regional performance with meeting targets, it appears regions that meet more frequently have a higher level of accuracy in their targeting process. Regions with the most frequent communication (i.e., regions coded as "frequent") tend to be among those with lower levels of variance. In contrast, regions with less frequent planning meetings tend to report higher variance of all types.

While we were not able to code data for other aspects of communication, we examined qualitatively the responses in high performing regions – respondents in these regions tended to report that a significant amount of time is spent in planning, and that planning is a year-long process. These insights are consistent with the meeting schedules identified, and suggest that frequent and consistent communication increases the effectiveness of planning.

EXHIBIT 3. RAW CODED DATA FOR EVALUATION OF PLANNING FREQUENCY

REGION	WEEKLY OR MORE OFTEN	BIWEEKLY OR MONTHLY	QUARTERLY OR OTHER REGULAR	SEMI-ANNUAL	AS NEEDED	ANNUAL	NEVER/ NO RESPONSE	GRAND TOTAL	GROUPS
A		4					1	5	regular
B		3						3	regular
C			2	2	2*			4	regular
D			4					4	regular
E	2			2				4	mgrs frequent
F	3		1		1	1		6	mgrs frequent
G			2	1		1		4	not frequent
H	1**			2		1		3	not frequent
I				1		1	2	4	not frequent
J		1					3	4	conflicting answer
Total	5	8	9	8	1	4	6	41	

* denotes ad hoc meetings in addition to "semi-annual"
** IMC reported regular CERCLIS meetings but not strictly for planning

Incorporation of Targets in Employee Performance Agreements

In addition to the frequency of meetings, we examined whether inclusion of specific site milestones and targets in RPM and managerial (and also IMC) Performance Agreements is correlated with regional performance.

Most of the regions reported that targets are specifically incorporated into employee Performance Agreements. In only one region did all respondents indicate that milestones are *not* specifically incorporated, though general performance goals related to meeting targets are incorporated. Another region noted that emphasis on planning is increasing in these agreements. Responses varied in that some regions mentioned the agreements are site-specific, and others identified more "generic" wording that does not specify sites but notes targets. Responses also varied within regions, and even within positions (e.g., two RPMs or two managers in the same region would sometimes indicate different practices).[1]

We examined the different answers provided and did not find a strong link between specific answers and regional performance, in part because the answers in most cases were fairly similar. However, we do note one specific pattern: those regions in which respondents provided inconsistent responses to this question tended to have higher levels

[1] Note that some variance across Performance Agreements is to be expected depending on how the chain of command is designed at different regions. For example, in some regions managers may have specific targets, while RPMs and IMCs have more general requirements. In other regions this may be reversed, or may vary year to year if only GPRA-related targets are tracked closely.

of variance of all types. This suggests that consistent communication and incorporation of targets into Performance Agreements may be a useful tool in improving planning.

Training for the Planning Process

All regions reported that they do not offer regularly scheduled training for the planning process; several regions noted that staff are very experienced and do not require formal training. Generally, training for the planning process appears to be included in new employee training. Additional training is offered as needed, particularly for the introduction of new features or modules in CERCLIS, but these training sessions do not focus specifically on planning and target-setting.

Regions reported a range of methods and materials for assisting staff in becoming familiar with or improving the process of target-setting, development of relationships with other agencies, and management of data. Identified sources include "business plans," regional planning documents, various internal memoranda (unspecified) and more standardized sources such as SCAP 14 reports, audit spreadsheets provided by headquarters, and the Superfund Program Implementation Manual (SPIM). Unfortunately, in part because of the range of sources identified and in part because it is difficult to identify how often the regions use these materials, we were not able to identify a relationship between use of any specific material and regional performance.

In response to a separate question about whether additional training for the planning process could be helpful, respondents in three regions indicated that additional training would not be helpful; it is noteworthy that all of these regions reported low or moderate levels of variance. The majority of respondents in the remaining regions (including all of the regions reporting high levels of variance) indicated that training would be helpful. These results suggest that the planning process could be improved from some additional training opportunities, and regions that have had the most difficulty accurately planning would likely take advantage of these opportunities.

QUESTION 1: INSIGHTS AND CONCLUSIONS

Responses to questions about the planning process suggest that:

- The overall work planning process is similar across regions, and most regions emphasize that the process is in part developed "organically" as RPMs and managers work together on sites over time and develop informal methods for establishing goals.
- In general, regions that conduct regular planning meetings for RPMs and other staff tend to perform better in meeting targets than regions with less frequent meetings and regions in which RPMs did not provide clear answers to the questions.
- The use of Odometer and e-facts as a method of auditing data corresponded with reasonable performance in two regions who also had relatively high numbers of targets.

- We could not identify any correlation between regional performance and the use of best practices documents or the incorporation of targets into Performance Agreements. However, regions with inconsistent responses regarding the incorporation of targets into Performance Agreements tended to have higher levels of total and negative variance, suggesting that increased emphasis on clear communication of targets could be useful.
- The majority of respondents indicated that training would be helpful. These results suggest that the planning process could be improved from some additional training opportunities, and that respondents in some of the regions that have had the most difficulty accurately planning would likely take advantage of these opportunities.

EVALUATION QUESTION 2.

What Factors Do Regions Consider When Developing Their Targets?

To address this question, we examined the specific processes that regions use to determine when a site will achieve a milestone within a year. To evaluate the factors that regions consider when developing targets, we evaluate the following interview questions:

- What site factors are taken into consideration in setting targets?
- How does complexity of the site factor into developing targets?
- How do regions account for uncertainty in establishing targets?
- How do regions incorporate coordination with legal and other technical resources in developing targets?
- How do regions coordinate with other agencies in setting targets?

This section examines the responses to these questions and their links to regional performance in meeting targets.

Site Factors Considered in Setting Targets

All regions identified "site complexity" as the principal factor in developing targets. However, we were unable to identify a consistent definition of "complexity" – respondents typically noted that "it depends on the site" and other responses that could not be used to develop a coded response. Similarly, several regions made general mention of site type or location as driving factors, but could not define these clearly. However, respondents in several regions also mentioned the following factors as important considerations in establishing targets:

- **Budget/Funding:** Specifically, respondents identified difficulty in meeting targets when budget priorities in lead agencies shifted during the year.
- **Lead Agency:** Respondents in all regions noted that certain lead agencies were difficult to work with, though the specific "difficult" agencies varied by region. The following agencies and branches were mentioned most frequently:
    - Air Force (Six regions);
    - Army (Two regions);
    - DOD (Four regions);
    - DOE (Four regions);
    - DOI (One region); and
    - Navy (Two regions).
- **Community Groups:** Respondents in one region noted that extensive public hearings and comment responses are difficult to predict and can affect schedules.

- **Groundwater and contamination type:** Respondents noted that actions at sites requiring groundwater remediation generally are more difficult to predict (this represents a large majority of sites).
- **Munitions:** Sites with munitions appear to have presented a recent complication in the planning process, as different investigations and remedial actions are required when munitions are present; DOD's Military Munitions Response Program (MMRP) has delayed or complicated cleanup actions as many sites..

We discuss relationship between a number of these variables and regional performance below as part of questions 6 and 7. While we were unable to identify CERCLIS fields or other data that would recognize sites with budget issues and community relations requirements, we investigated the impact of lead agency, contaminant type, and the number of OUs (a proxy for site "complexity") under question 6. Under question 7, we considered the impact of planning for sites at different stages in the pipeline – this is also a potential indicator of site complexity as it relates to planning.

Accounting for Uncertainty in Setting Targets

Respondents noted that the most common way to account for uncertainty in the planning process is to "be conservative" and ensure that targets have sufficient time in their schedules to accommodate changes. In addition, regions reported employing the following methods to reduce uncertainty, but did not specifically outline approaches:

- Communication with federal counterparts
- Communication with other staff members
- Consideration of complexity and issues
- Drawing on experience/historical perspective
- Case-by-case scrutiny of targets

Notably, four of the six regions that emphasized a "conservative" approach in the planning process also reported the highest levels of *negative* variance among regions. This result is counter to what would be expected if a conservative approach is effective in eliminating overstatement of targets. Moreover, in some cases respondents in these regions indicated specific approaches in addition to general "conservation." This result suggests that there may be a role for more specific approaches to accounting for uncertainty.

Coordination with Legal and Other Technical Resources

Region 1 respondents noted that their planning process includes close coordination with regional technical staff to ensure that legal and other technical resources are adequate to perform the activities (e.g., reviewing RODs) that are necessary to meet targets. We therefore asked other regions how they ensured adequate technical support for completing the milestones identified in the target-setting process.

Respondents in general noted that managers are responsible for ensuring the availability of legal and technical resources, but a number of regions noted that technical resources are not a limiting factor in work planning because they utilize contracted staff, or they have in-house staff that are readily available.

Regions also noted a number of methods for coordinating technical resources. In some regions, RPMs coordinate with technical staff directly; in other regions, one staff member or manager is responsible for all coordination. Our review of the responses indicates that, while this is potentially an important issue, most regions have identified effective methods for considering technical support and there is no strong correlation between the different approaches identified and regional performance.

Coordination with Other Agencies (Agencies other than the Federal Facility Counterparts)

Five regions indicated that their sites require coordination with states and other agencies in addition to the lead agencies responsible for federal facility cleanups. Examples of other agencies include:

- Fish and Wildlife Service (FWS)
- National Oceanic and Atmospheric Administration (NOAA)
- U.S. Geological Survey (USGS)

The remaining five regions responded that they coordinate only with their federal facility counterparts. While respondents did not describe the coordination in detail, the regions that coordinate with other agencies included three regions with the lowest levels of variance, and regions that did not coordinate included those regions reporting some of the highest levels of variance. It is not clear, however, whether or not expanded coordination with additional federal partners represents an opportunity for improving planning.

QUESTION 2: INSIGHTS AND CONCLUSIONS

Responses to questions about considerations for setting targets suggest that:

- Regions identify a range of methods for reducing uncertainty, but did not describe specific approaches. Oddly, however, regions that generally mention being conservative in developing targets report high negative variance (expressed as a percent). It appears, therefore, that more specific approaches than a general conservatism might be more useful in improving targeting.
- Regions that stress coordination with other agencies report lower levels of variance, suggesting that an emphasis on close coordination provides some improved information about the progress at the sites.
- We were unable to identify any correlation between regional planning performance and responses regarding site complexity (e.g., size, remediation type, presence of munitions). In addition, regions appear to have developed effective methods for coordinating with legal and technical resources.

EVALUATION QUESTION 3.

What Quantitative Or Formulaic Methods (If Any) Are Used To Estimate Final Targets?

In some cases regions reported using a range of quantitative or formulaic methods to adjust targets in CERCLIS. We examined the relationship between quantitative methods identified by regions who use them and the performance of these regions.

In general, respondents in most regions report that they set and adjust targets on a case-by-case basis, considering various site features and past performance. Respondents in one region, however, reported that they employ a method of factoring a probability of 100% in first quarter, 75% in second quarter, 50% in third quarter, and 25% in fourth quarter to provide a probability-adjusted set of targets. One respondent in another region reported that regional staff would once divide targets by two, but the region no longer employs this approach. Finally, one region, noted a "10 percent rule" (presumably a downward adjustment of targets by 10 percent to account for uncertainty), but the respondents did not elaborate.

We examined the regions that reported current or past quantitative adjustment methods, and their performance ranged considerably. The region with the former "divide by two" approach reports high variances of all types, suggesting that this approach was either ineffective or inconsistently applied. In contrast, the region employing the 10 percent rule reports high performance (i.e., low variance). The region employing quarterly adjustments reports high total and negative variance, but also has a higher number of targets than many regions. It is difficult to determine whether the quarterly adjustment approach is effective; the larger number of sites may have complicated implementation of the approach.

Based on the limited experiences reported by the regions, it is difficult to determine whether a specific quantitative method of adjustment is an effective planning tool.

QUESTION 3: INSIGHTS AND CONCLUSIONS

Responses to questions about quantitative or formulaic methods for accounting for uncertainty in setting targets suggest that only three regions have employed quantitative methods. Methods differ by region, and at least one of these regions no longer uses the approach.

Performance among regions that do report quantitative approaches to adjusting targets varies widely. The region incorporating a 10 percent "rule of thumb" adjustment is high-performing, but the region reporting that it previously used a 50 percent adjustment is surprisingly one of the regions with a relatively large number of "missed" goals (high negative variance). Finally, the region with a relatively sophisticated quarterly adjustment approach has a mixed performance, with a significant number of "missed" targets but also a large number of overall targets, resulting in a reasonable performance when variance is expressed as a percentage.

EVALUATION QUESTION 4.

How Is CERCLIS Integrated Into The Federal Facilities Response Program Site Remediation Goal And Milestone Planning Process?

To evaluate how the integration of CERCLIS into the work planning process contributes to effective target-setting, we considered the responses to the following questions:

- How is CERCLIS integrated into the planning process?
- Are targets based on data in CERCLIS?
- Are regular CERCLIS meetings held to coordinate input or updates?
- How important is CERCLIS in relation to measuring job performance?

Below we outline responses to these questions and note areas where responses appear to be correlated with high performance (e.g., low variance) in the work planning process.

Integration of CERCLIS in the Work Planning Process

All regions noted that CERCLIS is an integral part of the planning process and is often a starting point for measuring progress, monitoring data quality, and setting targets. Many regions view CERCLIS as a tracking tool and many also report using features such as SCAP reports and the Odometer. Only one RPM responded that CERCLIS is not used to set targets. While this region reports a significant number of missed targets, other respondents in the region report using CERCLIS for setting targets, and it is not clear that the region's performance is related to confusion about the role of CERCLIS. However, effective communication about the use of CERCLIS in all regions is likely to be an important factor in an effective work planning effort.

Relationship Between Targets and CERCLIS

Respondents in all regions reported that their targets are based on data in CERCLIS, although in three regions one respondent answered "no" or "not really." The three regions with these inconsistent responses have high levels of total variance; two of the three also have high levels of negative variance. In contrast, respondents in two regions with high performance (i.e., low levels of variance) answered consistently that the targets are based on data in CERCLIS and went on to further explain that CERCLIS is a helpful tool for tracking. These results, though strictly illustrative, appear to underscore the effectiveness of a clear integration of CERCLIS into the planning process.

Role of Regular CERCLIS Meetings

There was some inconsistency in answers from regions about the role of regular CERCLIS meetings, but in general, most respondents indicated that they either have regularly scheduled planning meetings that involve CERCLIS or they discuss planning as needed with RPMs and other staff. We were not able to identify a correlation between responses, variation in responses, and performance.

Importance of CERCLIS in Job Performance

All regions that responded to this question had at least one respondent indicate that CERCLIS is important in measuring job performance; this is expected because one of the respondents in each region is an IMC with responsibility for managing CERCLIS data. However, it is notable that in four regions, one or more respondents indicated that CERCLIS is not important. Two of these regions also had respondents that indicated that the targets are not based on data in CERCLIS, and both of these regions had high rates of both total and percentage variance; the other two regions had moderate performance. It therefore appears that regions with staff members that do not highly value CERCLIS may have weaker performance. In contrast, management and staff commitment to integrating CERCLIS into the planning process is likely to be correlated with more effective work planning.

QUESTION 4: INSIGHTS AND CONCLUSIONS

Responses to questions about the integration of CERCLIS into the planning process suggest that:

- There are varying opinions about how important CERCLIS is for measuring job performance as it pertains to accurate and achieved targets. It is notable that regions with high levels of variance and respondents that do not believe CERCLIS is used to set targets appear to be less likely to believe that CERCLIS is important for measuring job performance.
- We were not able to identify a correlation between variance and how CERCLIS is integrated, whether targets are based on data in CERCLIS, and the reported frequency of planning meetings. There is, however, a relationship between inconsistent responses to the question about targets being based on data in CERCLIS and whether performance (i.e., high variance).

EVALUATION QUESTION 5.

What Measures Are Taken To Ensure High Data Quality In CERCLIS?

To evaluate what measures are taken to ensure high data quality and which measures may lead to an effective planning process, IEc evaluated responses to the following questions:

- What measures are taken to ensure high data quality?
- What is the process for updating milestones?
- Do regions use the management review function?
- Are regular CERCLIS meetings help to coordinate input or updates?
- Do RPMs enter data?

Below we outline responses to these questions and note areas where responses appear to be correlated with high performance (i.e., low variance) in the planning process.

Measures to Ensure High Data Quality

Regions reported that they conduct reviews of audit reports and actual data in CERCLIS to identify potential issues. There is at least one person in each region that pulls audit reports and/or data for quality control. Respondents in seven regions mentioned that audits are done at least monthly or quarterly while three regions did not specify. Of the three regions that did not specify, two have high levels of variance. There may be an inverse relationship between awareness of data audits and variance, in that variance decreases with increased knowledge. It is important to note that the interview questions did not directly ask how frequently audit reports are used. However, respondents from many regions provided this information in their responses to data quality questions.

Process for Updating Milestones

Respondents in all regions reported that milestones are updated as needed, and that management reviews all changes. In all regions, RPMs are expected to identify milestones that require revision. The process by which changes are made varies greatly among regions. Respondents from some regions indicated that an informal conversation with a manager is sufficient, while in other regions formal notification may be required. In many regions the IMC or other designated data quality reviewer occasionally reviews dates to verify that they are up-to-date. The process for updating milestones is varied and particular processes do not appear to correlate with variance.

Management Review Function

Seven regions reported that they use the management review function in CERCLIS. Of the three regions that do not use the function, only one is a region with high variance. It is not clear whether the use of the management function alone is important in supporting regional planning. Those regions that do not use the management function rely on RPMs to notify managers of the change in target date, and two of the three regions that do not use the management review function rely heavily on monthly audit reports, suggesting

that frequent audit reports may contribute more to performance than use of the management function.

CERCLIS Meetings

All regions had at least one respondent that indicated there are regular meetings, but the responses to this question varied. Two regions provided consistent responses, all of which were that regular meetings are held.

Data Entry by RPMs

Eight regions indicated that some or all RPMs enter data into CERCLIS. Three out of these eight regions indicated that some RPMs enter data, but not all. Two regions indicated that RPMs do not enter data into CERCLIS. There does not appear to be a correlation between who does data entry and variance.

QUESTION 5: INSIGHTS AND CONCLUSIONS

The table below summarizes each region for the following:

- Use of Management Review Function
- Regular CERCLIS Meetings
- Mention of Frequency Audits
- RPM Data Entry

Two of the regions with the best performance answered "yes" to the most questions, and the two regions with the highest variance answered "yes" least often. It is important to note, however, that with the exception of one region at each end of the distribution, the answers did not vary much across regions.

EXHIBIT 4. SUMMARY OF CERCLIS CHARACTERISTICS

REGION	MANAGEMENT REVIEW FUNCTION?	REGULAR CERCLIS MEETINGS?	ARE THERE FREQUENT AUDITS?	DO RPMS ENTER DATA?	COUNT OF "YES'"
A	Yes	Yes	Yes	Yes	4
B	Yes	Yes	Unknown	Yes	3
C	Yes	No	Yes	Yes	3
D	Yes	Yes	Yes	Some	3
E	No	Yes	Yes	Some	2
F	No	Yes	Yes	Some	2
G	No	Yes	Yes	No	2
H	Yes	No	Unknown	Yes	2
I	Yes	No	Yes	No	2
J	No	No	Unknown	Yes	1

Based on responses to questions about CERCLIS data quality, we were unable to identify a correlation between variance and measures taken to ensure high data quality, the process for updating milestones, use of management review function, regular CERCLIS meetings, and data entry. However, it does appear that regions that clearly integrate multiple data quality activities into their processes general achieve better results in their planning process.

In addition, regions that provided consistent responses had lower variance than those regions that provided inconsistent responses. Frequency of internal coordination and communication may therefore be a significant driver of performance, regardless of which specific data quality actions are implemented.

EVALUATION QUESTION 6.

Do Sites That Present Difficulty When Meeting Targets Share Any Common Features?

This question considers factors external to the planning process that may complicate the ability of regions to meet established targets. To evaluate which site features may cause difficulty for the planning process, IEc considered two aspects of "difficult sites." First, we examined whether the sites named by regional staff as problems had any significant similarities. Based on regional responses to Question 2 – we examined the following site features:

- Average number of OUs at difficult sites versus non-difficult sites (this is a general proxy for site complexity)
- Frequency of lead agencies for difficult sites and non-difficult sites
- Frequency of specific site characteristics, including groundwater contamination or other contamination type, site stage, and MMRP status.

Second, we examined the extent to which the number of sites with features identified as difficult in Question 2 is correlated with regional performance.

We also examined the current stage of difficult sites, though we discuss the relative difficulty of planning different milestones in more detail under Question 7.

Through our interviews and follow-up correspondence, respondents identified 49 sites as difficult sites, due either to difficulties in planning with the lead agency, or to some other site feature. The following analyses of the difficult sites include data from three sources: interviews, follow-up correspondence, and NPL site data provided by EPA HQ. For 11 sites mentioned in interviews we were unable to match the site names with CERCLIS records; we therefore based our analysis on the 38 sites we could clearly identify. Please refer to Appendix C for a list of all "difficult" sites showing the varying degrees of specificity of each site.

Current Site Stage

Of the 38 identified difficult sites for which we were able to link to NPL site data, 23, or 61 percent, had remedial action (RA) as the last completed pipeline action. However, it also may reflect the fact that sites with a long and difficult history are most likely to be identified as "difficult" – and many of the missed targets may pre-date the data used for this study.

Lead Agency

Federal facilities NPL site information identified 16 lead agencies across all sites. Of these 16 agencies, eight were the lead at only one site, and three were the lead at only two sites. Exhibit 5 below shows the breakdown of sites by lead agency divided into six categories aimed at better showing the number of sites for each main lead agency.

EXHIBIT 5. NUMBER OF SITES IN EACH REGION BY AGENCY

	DOE	USAF	USAR	USNV	DOD-OTHER[1]	NON-DOD OTHER[2]	TOTAL
Region 1	0	3	4	7	0	1	15
Region 2	3	3	3	4	0	1	14
Region 3	0	3	8	15	1	3	30
Region 4	3	3	4	8	1	0	19
Region 5	2	2	4	1	0	1	10
Region 6	1	2	3	0	0	2	8
Region 7	1	0	5	0	0	0	6
Region 8	2	4	2	0	1	0	9
Region 9	3	11	4	10	2	1	31
Region 10	6	6	5	9	0	4	30
Total							

Notes:
1. DoD-Other agencies includes: Defense Logistics Agency and DODF.
2. Non-DoD Other agencies includes: COEN, DOAG, DOIN, DOTR, EPAA, FAAA, NAGU, NASA, SBAA, and USCG.

For "difficult" sites, the two most common lead agencies are the Army and the Air Force. The Army accounts for 13, or 34 percent of all "difficult" sites, and the Air Force is the lead agency at 11 (29 percent) of the identified "difficult" sites. Both of these values are slightly higher than for all federal NPL sites which are Air Force-led 22 percent of the time, and Army-led 24 percent of the time. This comparison indicates that sites led by these two agencies may contribute to a greater likelihood of a site being classified as a "difficult" site, but the relationship is not strong enough to be conclusive.[2]

Contaminated Media

Of the 38 difficult sites we could positively identify and match to NPL site data, 33 (87 percent) had groundwater contamination, suggesting that in general, groundwater contamination is too common among federal NPL sites to explain variation in region-wide performance. However, it is interesting that the percentage of difficult sites with groundwater contamination is slightly *smaller* than the percentage of total sites, given the emphasis on groundwater reported by interview respondents.

While not specifically identifying a contaminated media, the prevalence of difficult sites with contamination identified as "other" suggests that unique or unusual contamination incidents are more difficult to address. CERCLIS data indicate that overall, 31 percent of all NPL sites noted "other" for a type of media contamination.[3] For the difficult sites, 39

[2] For an analysis of overall performance by lead agency, please refer to Appendix D.

[3] Explicit types of media affected include: Air; Buildings/Structures; Debris; Free Phase NAPL; Fish Tissue; Groundwater; Landfill Gas; Leachate; Liquid Waste; Residuals; Sediment; Sludge; Soil; Soil Gas; Solid Waste; Subsurface Soil; Surface Soil; and Surface Water.

percent of the 38 we could identify reported having "other" media contamination. Again, this difference is small, but it suggests that unusual contamination types are at least as important as groundwater in presenting difficulty when planning.[4]

Military Munitions Response Program (MMRP) Sites

From interviews and follow-up correspondence, we were able to identify 14 of the 49 difficult sites (29 percent) as having munitions. This information was obtained solely from non-CERCLIS data sources as most CERCLIS data does not include MMRP information. In comparison, as of July 2008, 86 of 131 (66 percent) DoD NPL sites were on the MMRP inventory, suggesting a weak or even inverse relationship between MMRP status and difficulties with site planning.[5] However, incomplete data on the number of MMRP sites by region limits our ability to correlate the number of MMRP sites with overall performance. In addition, it is important to note that interviewees were generally strong in their assertion that presence of munitions affected schedules and remedies at sites, suggesting that MMRP status should not be discounted as a reason for difficulty meeting regional targets.

QUESTION 6: INSIGHTS AND CONCLUSIONS

Responses to questions about common features among "difficult" sites suggest that:

- In general, sites identified as "difficult" sites are now in late stages of the pipeline; 61 percent of all of the difficult sites identified by interviewees report completion of RA work at an OU as the most recent action.
- Identified difficult sites have Army and Air Force as the lead agency at a slightly higher frequency than other federal sites, but the relationship between difficult sites and agencies is not strong.
- "Difficult" sites are slightly less likely to have groundwater contamination than other sites, in contrast to responses from interviews about which site features most complicate planning. Conversely, difficult sites were more likely to have "other" contaminated media than other sites, suggesting that unique or uncommon contamination may be a contributor to creating a difficult site.
- Our assessment of the relationship between sites with "difficult site" features and overall regional performance in work planning did not identify any correlations.

In general, our investigation of the features of difficult sites did not isolate any specific site features that could be used to help regions predict what other sites will present planning difficulties.

[4] For an analysis of overall performance in relation to frequency of groundwater and other contaminated sites, please refer to Appendix E.

[5] "Military Munitions Response Program Update July 2008," presentation, U.S. EPA Office of Solid Waste and Emergency Response.

EVALUATION QUESTION 7.

Do Certain Specific Performance Measures Present Greater Problems Than Others When Developing Or Meeting Targets?

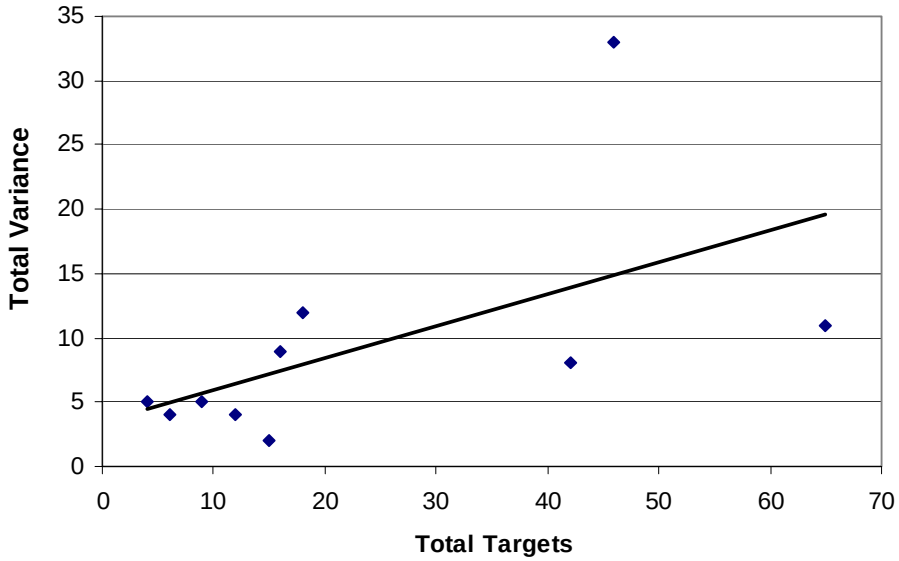
To evaluate which performance measures may cause difficulty for the planning process, we asked respondents to note the specific milestones that are most difficult to meet, and we also examined data on actual start and completion dates for different activities. Respondents reported with great consistency across regions that remedial investigations (RIs) are the most difficult to predict, and as a result, RODs (a subset of decision docs) are the most complex milestone to meet. This appears due to the uncertainty related to RI findings and to the need to coordinate with states, communities, and the lead agency on the final document. Respondents also indicated that CCs present challenges.

In addition, we examined information on the overall regional target workload (i.e., total number of targets and number of targets for each milestone) in comparison with regional staffing (full-time equivalents or FTEs).

Relationship of Total Number of Regional Targets with Performance

Using 2008 data, we compared the number of total targets set by each region and their total variance and negative variance, to determine whether "larger" regions or regions with more sites in active stages of the pipeline might experience more difficulty meeting targets. We identified a correlation between the total number of targets set and total variance, but surprisingly, we did not identify a correlation between *negative* variance (or "missed" targets) in regions. Exhibit 6 below shows the correlation data for total variance.

EXHIBIT 6. TOTAL TARGETS SET VERSUS TOTAL VARIANCE IN 2008



That a correlation exists between the total number of targets and total variance, but not negative variance, suggests that having a greater number of targets does not necessarily

lead to more missed targets, but rather indicates a greater uncertainty regarding the number of targets that could be met that were not originally planned for. It is also reasonable to assume that regional planning approaches focus more on assuring that targets are met (i.e., avoiding negative variance), and this is another reason that a relationship between targets and total variance is stronger than a relationship specifically between the number of targets and negative variance.

Specific Performance Measures

During our interviews, we asked each participant what performance measures presented the most problems in terms of meeting targets. The two most common responses were RODs and Construction Completes. Exhibit 7 below presents these results in full.

Exhibit 8 shows historical data on the percent of targets missed by target and, on average, how late they were in being met. Additionally, the exhibit shows the likelihood of missing each of type of target in 2008.

Some disconnect appears to exist between the information provided directly by respondents and the data on actual missed targets. Respondents indicated that RODs and Construction Completes were the most difficult measures to plan; the third most frequent response was that all measures are equally difficult. The historical data confirm that RODs and other decision documents in general are subject to delays, but Five-Year Reviews are also subject to high "miss rates" – and the delays associated with Five-Year Review targets exceed a year on average.

EXHIBIT 7. PERFORMANCE MEASURES THAT PRESENT GREATER PROBLEMS, AS REPORTED BY INTERVIEW RESPONDENTS

PERFORMANCE MEASURE	RESPONSES
RODs	11
Construction Completes	8
All Targets are Equally Difficult	5
GPRA Targets[1]	4
Sitewide RAUs	4
Five-Year Reviews	2
Decision Documents	2
Other[2]	14

1. Note that while some respondents listed specific measures others grouped together the GPRA measures.

2. Note many of these answers included responses that were not measures including groundwater sites and proposed plans.

EXHIBIT 8. HISTORICAL AVERAGE LATENESS BY PERFORMANCE MEASURE

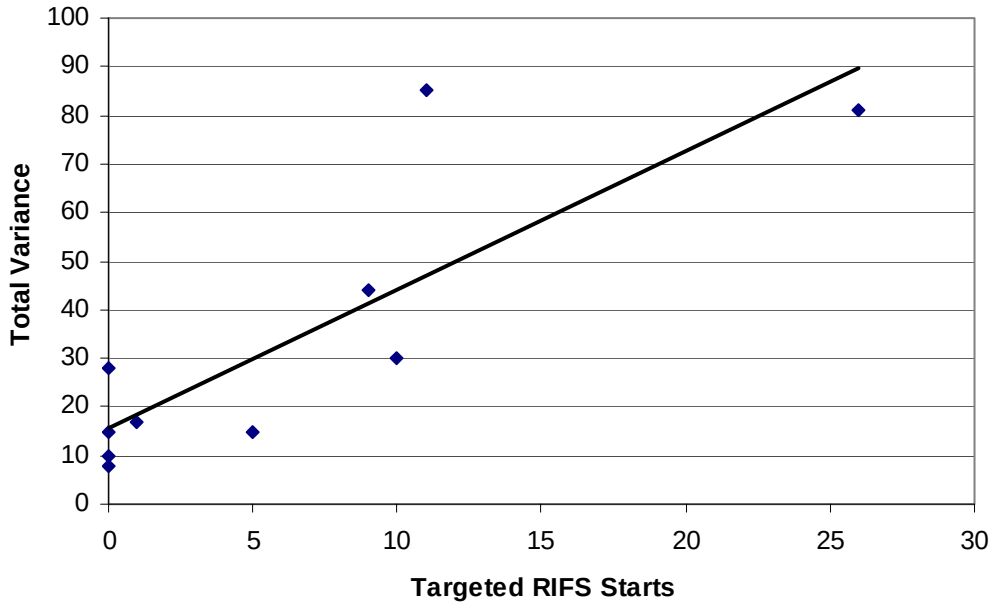
PERFORMANCE MEASURE[1]	HISTORICAL PERCENT OF TARGETS MISSED	HISTORICAL AVERAGE QUARTERS LATE	NEGATIVE VARIANCE IN 2008 (PERCENT MISSED)
Decision Documents	11.0	1.5	6 (8%)[2]
Final RODs	1.5	0.5	
Five-Year Reviews	12.4	4.2	1 (5%)
RA Completions	8.2	2.2	3 (7%)
RA Starts	7.9	2.3	7 (13%)
RI/FS Starts	14.4	2.0	2 (9%)

Notes:
1. The data do not allow for the calculation of historical averages for construction completions.
2. Final RODs are included in the Decision Documents target and therefore cannot be broken out as the historical data can be.

In addition, historically RI/FS Starts are the most frequently missed targets. In 2008, however, RA Starts were the most frequently missed target and Five-Year Reviews were the least missed target. It is noteworthy that RA Starts may be affected by difficulties completing Decision Documents (including RODs).

An analysis of the number of RI/FS Starts and Five-Year Reviews targeted from 2006 through 2008 found no correlation between total variance and Five-Year Reviews, but we identified a strong correlation between total variance and the number of targeted RI/FS Starts. Exhibit 9 shows this result of this significant correlation.

EXHIBIT 9. RELATIONSHIP OF TARGETED RI/FS STARTS TO TOTAL VARIANCE FROM 2006-2008



An important note, however, is that neither RI/FS Starts nor Five-Year Reviews were correlated with *negative* variance. If in fact these two targets were more likely to be missed, one would expect a greater correlation between the number of targets set and negative variance, rather than total variance. This suggests there may be an additional factor linking RI/FS Starts to variance creating the strong correlation shown above.[6]

QUESTION 7: INSIGHTS AND CONCLUSIONS

Responses to questions about performance measures that present greater problems than others suggest that:

- A relationship exists between the total number of targets and total variance, but it does not appear to be linked to more "missed targets" (i.e., higher negative variance). Thus, having a greater number of targets does not necessarily predict more missed targets, but rather indicates a greater general uncertainty regarding the number of targets that can be accomplished in a year.
- Data from interviews and data from CERCLIS on the types of milestones that are most often missed appear slightly inconsistent, with interview respondents clearly noting RODs and Construction Completes as the most difficult milestones to plan for (third most frequent response was that all measures are equally difficult). In contrast, CERCLIS data documents that total Decision Documents (including RODs and other decision documents) frequently lag their schedules, but that Five-Year Reviews and RI/FS starts have historically had the most "missed schedules" and Five-Year Reviews have the longest average lag time, at over one year. Notably, however, in 2008 RI/FS starts and Five-Year Review completions were not missed in significant numbers.
- In general, the number of targets set for any specific milestone does not appear to be linked to overall regional performance. While we noted a strong correlation between RI/FS Starts and total variance, this is difficult to interpret because negative variance (i.e., "missed" targets) is not strongly related.

[6] For details on the relationships of each target and each performance metric, please refer to Appendix F.

EVALUATION QUESTION 8.

Do Any Features Unique To Specific Regions Affect The Ability To Plan And Implement Goals?

To evaluate which features may cause difficulty for the planning process IEc considered the following characteristics that vary among regions and could affect overall performance in work planning:

- Non-NPL sites,
- Staffing levels,
- State government relationships,
- Geographical issues,
- Lead agencies, and
- Maturity of sites.

Number of Non-NPL Sites

We were not able to examine the relationship between regional performance and the prevalence of non-NPL sites quantitatively due to limited information about non-NPL sites. However, interview respondents in general reported that non-NPL sites do not affect work planning for NPL sites, though many regions stressed that their non-NPL activities are limited or complete. When respondents did mention impacts, they noted the following issues that are not directly related to work planning:

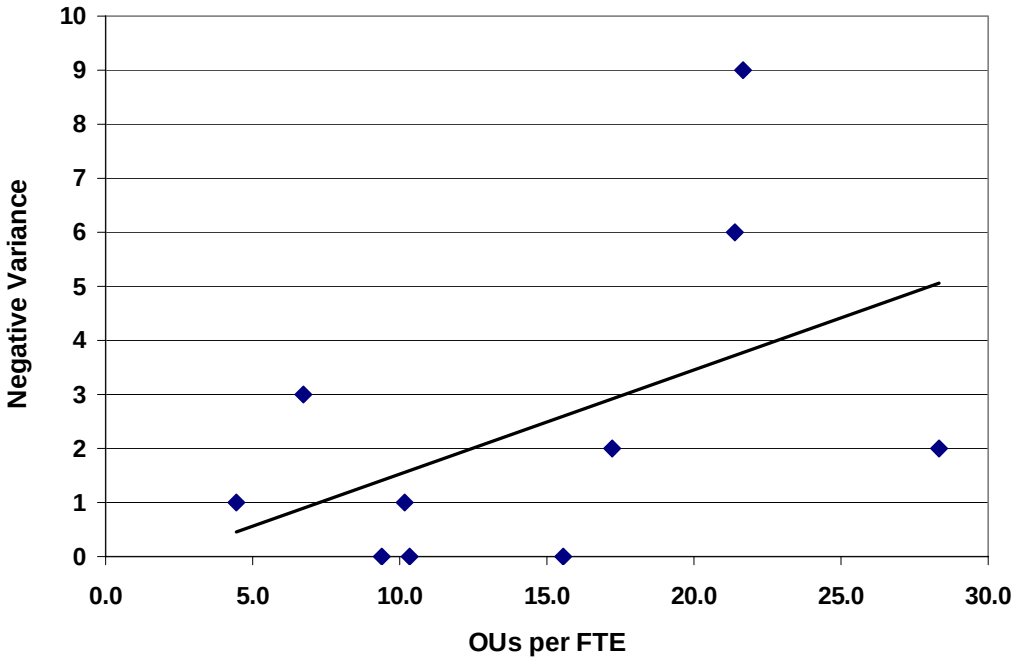
- Lack of credit for work is frustrating as non-NPL sites use resources but do not count toward their FFRRO targets.
- Working on non-NPL sites is frustrating due to a feeling of responsibility for sites with little authority.
- Non-NPL sites compete for resources and make it difficult to prioritize activities.

Staffing Level Within Each Region

The following graph shows the relationship between workload per full-time equivalent (FTE) and total net variance. The regions with the three highest levels of workload have three of the five highest total variances. We define workload as the number of FTEs in a region per total OUs.[7]

[7] Note that we were not able to examine FTEs per "active" OU – defined as OUs that had not yet achieved RA completions – because our data only identifies latest action at the site level. An analysis of this factor could potentially reveal a different relationship between FTEs and performance, but it is very similar to the type of analysis already undertaken in EPA's Workload Planning model, which we discuss below.

EXHIBIT 10. OUS PER FULL-TIME EQUIVALENT VERSUS NEGATIVE VARIANCE



As evidenced by Exhibit 10 above, the relationship between OUs per FTE and the negative variance is positive, but the strength of the correlation is limited. Therefore, while the number of OUs per FTE appears to be a factor, other factors, such as the stage of the site in the pipeline, may be more important.

Note that EPA has also developed a workload planning tool that considers a weighted approach to account for the different levels of effort required by RPMs at different stages of the site remediation process. An examination of the results of this analysis reveal that, in fact, the regions with the largest FTE requirements (and in some cases projected FTE shortfalls) are the regions reporting some of the higher levels of total variance.[8]

State Government Relationships

The majority of respondents indicated that relationships with state governments and other agencies are generally going well. Three out of the four regions that did not mention good relationships have the highest variance levels while the regions with the lowest variance report that their relationships are good. In general, however, the feedback for this question was positive with minor comments about individual state issues. Therefore, this correlation is likely to have only a limited impact on overall planning. Planning processes that focus on or adjust for actions at sites where state relationships are not strong, might improve planning accuracy in some cases.

[8] Please refer to Appendix G for the EPA workload planning tool analysis.

Geography

Some respondents identified distance or communication between offices, regional "sub-offices" or sites and regional offices as negatively affecting communication and planning. While we are unable to develop a meaningful evaluation of impact of this issue due to the limited number of regions with multiple offices and the range of site locations, it seems reasonable to consider steps to improve communication within regions, particularly when regions are spread out geographically.

A general analysis of overall geographic area and performance yielded no significant correlation between land area and variance suggesting that larger distances between sites do not systematically affect performance, but specific barriers between sites (e.g., mountains, water) within regions may have an impact.

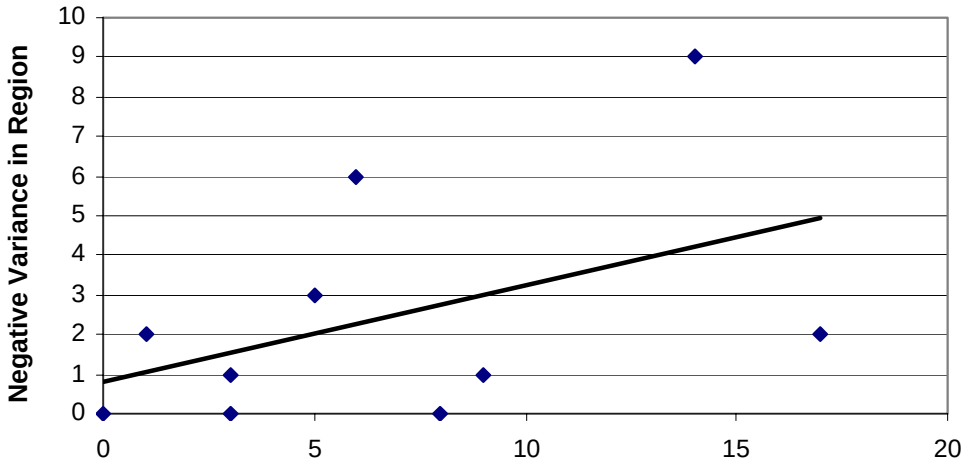
Lead Federal Agencies

A correlation test on each of the six groupings of lead agencies (Department of Energy, U.S. Air Force, U.S. Army, U.S. Navy, Department of Defense- Other, and Non-Department of Defense- Other) against negative variance yielded no significant correlations of note. These results suggest that despite interviewees noting difficulties at sites with specific lead agencies, no national-level relationship exists between lead agency and the likelihood of missing work planning targets.

Latest Pipeline Action

Data on each site includes a field called "Latest Pipeline Action." The actions that were included in the data are: CC, deleted, O&M, RA, RD, RI/FS, ROD, and SWRAU. Correlation tests on each of these actions and negative variance found a strong correlation only with the number of sites with RA as the most recent pipeline action. Exhibit 11 shows the results of this correlation and the trend line.

EXHIBIT 11. COMPARISON OF THE NUMBER OF SITES WITH "RA" AS THE LATEST PIPELINE ACTION VERSUS NEGATIVE REGIONAL VARIANCE



A limitation of this analysis is that we performed it at the site level; this may obscure activities at large sites with many OUs. For example, a site with an "RA completion" as a recent action could still have a number of OUs undergoing RI/FS stages. To fully assess the impact of "pipeline stage" on planning accuracy, we would ideally expand this analysis to the OU level.

QUESTION 8: INSIGHTS AND CONCLUSIONS

Evaluation of the relationship between different regional variables and regional performance reveals the following:

- The "size" of a region's workload may be a factor in performance. The relationship between total OUs per FTE and the negative variance is positive, though the relationship is not very strong. More tellingly, however, EPA's workload model that adjusts for the level of effort required at sites in different stages of the Superfund pipeline, identifies several regions with estimated staffing "shortfalls" for 2008. These include regions that tend to report higher levels of total variance and in some cases negative variance.
- Regions where respondents mentioned good relationships with states tended to have lower levels of variance, suggesting that attention to relationships with states, and careful consideration of targets in states where relationships are difficult, may assist in planning.

- Correlation tests on recent pipeline actions at a site level and negative variance demonstrated a strong correlation between the number of sites with RA as the most recent pipeline action and negative variance (missed targets). This relationship seems inconsistent with respondent assertions that RODs are the most difficult pipeline stage to plan, but site-level actions may obscure difficulties associated with larger sites with multiple OUs. A more detailed analysis of actions at the OU level could clarify any relationships between planning and "pipeline stage."
- There is no national correlation between performance and lead agency, or performance and geographical area of regions.

EVALUATION QUESTION 9.

How Does Regional Organization Affect What Planning Processes Are In Place?

This question is designed to evaluate whether current regional organization is related to the specific planning processes that are in place in different regions, and may therefore affect performance. IEc considered the following questions:

- How is the branch organized? Specifically, is there more than one branch addressing federal facilities NPL Sites?
- How many sites do RPMs typically cover?
- Do RPMs work exclusively on federal sites?

Below we outline responses to these questions and note areas where responses appear to be correlated with low variance (e.g., high accuracy) in the planning process.

Branch organization

Three regions indicated that the Federal Facilities branch is not located in one single branch, but instead is divided into two or more branches, most often divided by state boundaries. Of these regions, overall variance is low or moderate, and it is therefore not clear that organization has a significant impact on performance. Note, however, that while those regions separated into two or more branches did not have high levels of variance; respondents did indicate that coordination across branches can present a planning challenge.

Distribution of Workload: Number of Federal Sites Covered per RPM

The number of sites covered by RPMs seems to vary from one to eight with some RPMs covering all or part of one large site while other RPMs may have numerous small sites. The range of sites covered appeared to vary equally amongst regions with low and high variance. It does not appear that the distribution of site responsibilities within the Federal Facilities program has a significant impact on overall performance.

Distribution of Workload: Federal and Private Site Management

In five regions, RPMs appear to work exclusively or almost exclusively on federal sites; the remaining regions consistently reported a mix of both federal and private sites. The regions with the highest performance and those with the highest number of missed targets reflect both types of work assignment. We therefore conclude that the federal/private workload split does not have a significant impact on overall performance.

QUESTION 9: INSIGHTS AND CONCLUSIONS

Responses to questions about regional organization suggest that regional organization into multiple branches with federal sites, and distribution of workload among RPMs, do not have a significant impact on overall performance across regions.

GENERAL CONCLUSIONS AND RECOMMENDATIONS

Using the nine evaluation questions outlined above, the broad objectives of this program evaluation are to:

- Identify the reasons that certain FFRRO site cleanup milestone targets are not met under current planning and implementation processes.
- Identify processes that regions can put in place to more effectively establish and meet targets. If possible, this would result in a "preferred" work planning process, or "planning formula," that could be adapted to aid planning and implementation across all regions.

We first summarize the main outcomes of our analysis in three sections that map to the evaluation questions and address the first of the two objectives. We then provide a recommendations section that draws on the insights from the evaluation to suggest immediate options for improving Federal Facilities Response Program work planning, "stretch" options for Headquarters activities that could further the work planning process. This section also notes areas of uncertainty and priorities for additional analysis.

CONCLUSIONS

Our analysis of data collected through interviews with regional staff and data from CERCLIS and EPA Headquarters provides the following insights into the work planning process.

Relationship between Current Federal Facilities Response Program Work Planning Procedures and Regional Performance (Questions 1 through 3):

- **Most regions have developed planning approaches over time, and have organically integrated the experience of staff and informal but well understood methods of coordination and data quality control.**
    - Regions make only limited use of formal documentation or training for the work planning process.
    - Limited current use of formal training and documentation methods precludes an assessment of the effectiveness of these actions, but it is notable that many respondents indicate that additional training might be useful, and these respondents were more prevalent in regions that have difficulty accurately setting targets (i.e., regions showing high variance between targets and achievements).
- **Clarity in communicating and implementing the planning process appears to be a key factor in ensuring a high level of regional performance.**
    - Regions that consistently reported that the process required frequent meetings (at least quarterly) performed better (i.e., revealed lower variance) than regions where answers were conflicting or indicated only infrequent meetings.

- Regions reporting significant coordination with other federal agencies, and providing consistent and clear responses about the integration of planning targets into Performance Agreements also tended to have higher performance than other regions.
- In general, regions that provided differing or contradictory answers to questions about process reported lower performance in meeting their targets (i.e., had higher variance between targets and achievements).

- **Current attempts to account for uncertainty in setting targets have limited success.**
  - Use of quantitative methods by regions to adjust targets for uncertainty, other than one mention of a general "10 percent" adjustment to targets by one region, does not currently correspond with strong performance in avoiding "missed" targets (i.e., eliminating negative variance).
  - General assertions by regions that they are "conservative" in planning, when not combined with more specific methods, reveal a correlation with weaker performance and higher rates of missed targets.
  - Mention of more specific qualitative planning actions such as focused coordination with lead federal agencies does correlate with better performance in some cases, suggesting more detailed investigation of this aspect of regional planning practice may be useful.

Role of CERCLIS Integration in Federal Facilities Response Program Work Planning Procedures and Regional Performance (Questions 4 and 5):

- **Most regions describe CERCLIS as central to the planning process, and report using CERCLIS to both set targets and audit progress.**
  - While performance among these regions varies, performance in regions where one or more respondents said that CERCLIS is not important in setting targets is weaker than performance in other regions.
  - We were unable to detect a clear relationship between the different methods regions use to integrate CERCLIS into the planning process, in part because regional practice is relatively consistent with regard to the data used: observed variation in performance among regions may be better explained by other differences in the planning process, or by general quality of communication about planning.
- **Regions that report systematic efforts to ensure CERCLIS data quality tend to have better performance than other regions.** While no single data entry or reconciliation approach is strongly correlated with performance, performance does correlate generally with the implementation of multiple data quality efforts. We investigated use of the Management Review Function in CERCLIS, the implementation of regular CERCLIS meetings to address data quality, the use of regular CERCLIS audits to reconcile data with planning efforts, and entry of data

by RPMs. Regions that reported employing all or most of these actions, particularly regular data quality meetings and audits, also tended to be the regions with stronger performance.

Impact of External Factors on Regional Work Planning Performance (Questions 6 trhough 9)

- **The investigation of key features at "difficult" sites did not identify any features with strong links to overall regional performance in work planning.** Difficult sites do not appear systematically different from other sites in terms of type of media contaminated, lead federal agency, or MMRP status, and prevalence of sites with the features investigated (groundwater contamination, MMRP status, and Army/Air Force lead) did not correlate with general regional performance.
- **While data suggest that specific site milestones have different rates of success in achieving timely completion, we did not identify a simple relationship between number of targets for specific milestones and regional performance.**
    - Regions with higher total number of targets tend to report higher total variance, suggesting that it is harder to be accurate when planning for large numbers of targets. However, a larger number of total targets does not appear to be correlated with more "missed targets" (i.e., higher negative variance).
    - While interview respondents and data from CERCLIS on the types of milestones that are most often missed both reveal that RODs and other Decision Documents (when combined) are missed with some regularity, the CERCLIS data also reveal that Five-Year Reviews and RI/FS starts have historically been the targets missed at high rates, though these milestones were not missed in large numbers in 2008.
    - In general, the number of targets set for any specific milestone does not appear to be linked to overall regional performance. While we noted a strong correlation between RI/FS Starts and total variance, this is difficult to interpret because negative variance (i.e., "missed" targets) is not strongly related.
- **Regional variables, including staffing level, relationship with states, and potentially the "stage" of a site in the pipeline, appear to have a limited relationship with regional planning performance. Other issues (geographic area, lead agency at each site, regional branch organization, and RPM assignments across federal and private sites) do not appear related to performance.**
    - The relationship between total OUs per FTE and the negative variance is positive, though not strong. More tellingly, however, regions with staffing "shortfalls," as identified by EPA's workload model, tend to be

regions reporting higher levels of total variance and in some cases negative variance.

- Correlation tests on recent pipeline actions at a site level suggested that sites with at least one OU in the RA stage present a planning challenge, but a more detailed analysis of actions at the OU level would be necessary to fully assess the relationships between planning and "pipeline stage."

RECOMMENDATIONS

Based on the conclusions and insights from the data collection and analysis in this evaluation, we outline the following near-term and "stretch" recommendations. These recommendations are designed to respond to the second evaluation objective of identifying processes and methods that can be implemented to improve the effectiveness of the work planning process. Note that we incorporate suggestions for addressing external variables that may affect performance into the work planning procedure recommendations.

Work Planning Procedure Recommendations:

- **Headquarters/Regions: Develop and Deliver Training on Work Planning.** Additional training received interest among respondents in most regions, including all regions where planning has been a challenge. While training should reflect specific regional structure and approach, Headquarters could contribute to this process by:
  - Providing detail about HQ reporting requirements and how these are integrated with OSRTI for GRPA, PART, etc. Also, why year-to-year fluctuations matter;
  - Present and get feedback on use of audit reports, Odometer, other HQ tools; and support;
  - Attending training in order to get feedback on process as well as provide instruction.
- **Headquarters/Regions: Develop a Planning "cheat sheet."** As part of training or as an alternative to training, develop a brief "trouble-shooting guide" for planning that:
  - Identifies key Headquarters needs, concerns;
  - "Cross-walks" key differences between federal agency (particularly DOD and DOE) and EPA milestones and procedures;
  - Identifies "red flags" in site planning in considering whether targets are realistic;
  - Outlines the need and general process for keeping CERCLIS updated.

- **Regions: Improve communication about importance, structure of work planning process by:**
    - Clearly incorporating targets into all Performance Agreements (managers, RPMs, IMCs) where applicable. This clarifies expectations and raises priorities;
    - Developing a formal schedule for communication during planning process. Based on the results of our analysis, the most successful regions have at least quarterly meetings for managers, RPMs, and IMCs that address planning, though in some cases the meetings are one-on-one meetings between managers and RPMs;
    - Documenting approaches for coordinating with lead federal agencies at each site. Approaches will vary by EPA region, lead agency, and even by site, but should include at least a formal annual discussion between RPM and agency lead to reconcile planning for the fiscal year, and a mid-year progress check-in.
- **Headquarters/Regions: Account for Uncertainty.** Develop and implement a process for incorporating adjustments to targets to reflect significant uncertainty related to certain sites/actions. Requires additional analysis to determine appropriate adjustments, but consistent with regional reports, system could include probability adjustments for completing targets that:
    - Involve significant coordination with other agencies and entities (e.g., Construction Completes, RODs);
    - Involve a site with a history of delays;
    - Occur in a year with a large number of similar targets (e.g., multiple RODs) that will compete for agency resources;
    - Involve a lead agency with a recent personnel shift;
    - Involve a target that is unique to EPA or defined differently by the lead agency.

  The development process for this method could include coordinated effort between Headquarters and one or more regions, and could involve testing. Ideally, it would incorporate the "institutional knowledge" of senior regional personnel who already implement a similar process informally.
- **Stretch Recommendation - Headquarters: Rationalize Budget Allocation with Performance.** Develop a budget scoring system that differentially gives positive or negative credits for regional performance accounting for:
    - Total targets – encourage regions to avoid "aiming low;"
    - Total number of "missed" targets – encourage meeting targets;

- Total number of exceeded targets – (less than total number of targets and less than value for missed targets) – give credit for additional performance;
- Differential credits for RODs, CCs, and Five-Year Reviews (subject to verification that these are most difficult targets);
- Differential credits for regions with low weighted FTE/active OU ratios.

CERCLIS Integration and Data Quality Recommendations:

- **Regions:** Develop document clarifying regional chain of responsibility for CERCLIS data quality and, to the extent possible, incorporate this policy into Performance Agreements. The policy should include one or more of the following:
  - Implementation of management review function;
  - Periodic audits of CERCLIS data;
  - Regular meetings between RPMs, managers and IMCs to rationalize/update CERCLIS data (informal audits; can be incorporated into planning meetings).
- **Regions:** Encourage monthly, rather than quarterly, updates from RPMs to facilitate expertise in using the system and increased awareness of site status and relevant events.
- **Stretch Recommendation – Headquarters:** Review and coordinate all HQ reporting requirements for CERCLIS to limit burden on regional staff; goal should be to reduce reporting requirements and time by a significant percentage, in order to improve focus on data quality for remaining reports.
- **Stretch Recommendation - Headquarters:** Respondents to interviews suggested several changes to CERCLIS, including:
  - Improve the transparency of the database locking feature as targets get locked in and RPMs have difficulty identifying and changing locked fields;
  - Allow for entry of schedules and planned dates in CERCLIS;
  - Create memo fields for notes as some sites may exist for many years; and
  - Integrate project management or project tracking tools into CERCLIS.

APPENDIX A

FFRRO Evaluation Interview Questionnaire

Evaluation Objectives

- Why are certain annual FFRRO site cleanup milestone targets not met under current planning and implementation processes?
- Identify processes that regions can put in place to more effectively establish and meet targets. If possible, this would result in a "preferred" formula that would aid planning and implementation across all regions.

The interview questionnaire has three general categories of questions. The first section is a set of open-ended questions about the current FFRRO milestone target planning process in place. The second section is a more targeted discussion of data tracking approaches for documenting the planning process and also progress toward site milestones. The third section is a discussion focused on the issues and difficulties that affect the planning process. The questionnaire concludes with some open-ended questions about available data related to regional planning or results.

Current Planning and Implementation Process

1. General Interviewee Information
   a) Name?
   b) Title?
   c) What is the name of your position and what is your role?
   d) How long have you been in your job?

2. How is the planning process in your region implemented?
   a) Do you have a best practices document describing the planning process and how to set targets?
   b) What is your role in the planning process?
   c) Are regular planning meetings held within the region?
   d) How do you prepare for planning meetings with HQ?
   e) How much effort does the planning process require (for you and total region)?
      ei. Meetings?
      eii. Time?
   f) How much communication goes on among regional staff during the planning and implementation process?
   g) How does the availability of legal and other technical resources factor into the planning process?
   h) Who is responsible for designing and implementing the planning process?
   i) Does your region provide/utilize training for the planning process, if necessary?
   j) Are site targets or milestones incorporated into your individual Performance Agreement?
   k) In your experience, do other regions appear to have similar experiences or planning processes?

3. How does your region develop targets and cleanup dates?
   a) What site factors are taken into consideration (e.g., remediation type, complexity, community involvement, contaminants, etc.)?
   b) How does the complexity of the site factor into developing targets?
   c) Does your region use any quantitative methods to set or adjust targets?

d) How do you account for uncertainty (e.g., the need to depend on other agencies) in setting targets?
e) How does your region incorporate coordination with legal and other technical resources into developing targets?
f) Do you coordinate with other agencies in setting targets? If so, how?

Organization & Interpretation of Data

4. How is CERCLIS integrated into the planning process for setting FFRRO milestone targets?
   a) Are the targets based on data in CERCLIS?
   b) How does the regional planning process use CERCLIS data?
   c) What measures are taken to ensure high quality data in CERCLIS?
   d) When a milestone is changed in CERCLIS, what is the process for updating all corresponding dates and targets?
      di. How does the changed target date get updated?
      dii. Who updates it?
      diii. Are corresponding dates updated?
   e) Does your region implement the management review function in CERCLIS?
      ei. If so, who has management authority?
      eii. If not, do you use another review process?
   f) Are regular CERCLIS meetings held in the region to coordinate input or updates?
   g) In your experience, how can CERCLIS be used to make the planning process more effective and efficient?
   h) In your experience, would an increase in CERCLIS training opportunities make for a better planning process?
   i) How important is CERCLIS in relation to measuring job performance?

Planning Issues and Solutions

5. Do sites that present difficulty when meeting targets share any common features with each other?
   a) Reuse
   b) BRAC? Non-BRAC?
   c) Contaminants?
   d) Geology?
   e) Emerging contaminants?
   f) Community involvement?
   g) Specific agencies?
      gi. Why are these problems present?
      gii. How does coordination with other RPMs affect your ability to deal with difficult sites?
      giii. Do you have a list of "difficult" sites that are tracked (or generally known) in your region?

6. Do certain specific performance measures present greater problems than others when developing and meeting targets?

7. How does regional organization limit/direct what planning processes are in place?
   a) Is there a large number of new staff?
   b) Is frequent staff turnover/reorganization a problem?
   c) What are the major staffing issues you face?
   d) What is the chain of command for large sites?
   e) How many sites do RPM's typically cover?

ei. Are RPM's exclusively responsible for federal facilities or do they also work on non-federal facilities?

f) Are there multiple RPMs for one site?

g) How does the distribution of responsibilities affect coordination, QA/QC?

8. Do any features unique to your region affect the ability to plan and implement goals (e.g., need to communicate with specific state governments, role of regional organizations, unique climate issues, etc)?
   a) How do non-NPL sites affect your planning?
   b) Please describe your relationship with state governments and other agencies. Are there any conflicts or issues?

9. How does your planning and implementation strategy change from year to year? Does it change during the year?
   a) What drives this change?
   b) How important is a formal planning process to you and your region?
   c) What additional steps or resources do you think would benefit your planning process?

10. Do you see potential for closer communication between EPA and regions on the planning process?
    a) What kind of support/assistance would you like to see from HQ?

Information Requests and Additional Material

11. Is there any other information that you could provide that would assist us in our analysis?

12. Is there anyone else in your region that you suggest we speak with?

APPENDIX B

EXHIBIT B-1. PLANNING RESULTS FROM 2006-2008

2006 - 2008 Planning Accuracy - Sorted by Absolute Variance				
Region	Total Variance	Total Variance (Percent)	Negative Variance	Negative Variance (Percent)
D	8	42%	2	11%
F	10	43%	6	26%
B	15	33%	0	0%
H	15	68%	11	50%
G	17	26%	14	22%
I	28	54%	7	13%
A	30	28%	5	5%
J	44	59%	34	46%
E	81	39%	12	6%
C	85	64%	6	5%

2006 - 2008 Planning Accuracy - Sorted by Negative Variance				
Region	Total Variance	Total Variance (Percent)	Negative Variance	Negative Variance (Percent)
B	15	33%	0	0%
D	8	42%	2	11%
A	30	28%	5	5%
F	10	43%	6	26%
C	85	64%	6	5%
I	28	54%	7	13%
H	15	68%	11	50%
E	81	39%	12	6%
G	17	26%	14	22%
J	44	59%	34	46%

2006 - 2008 Planning Accuracy - Sorted by Percent Absolute Variance				
Region	Total Variance	Total Variance (Percent)	Negative Variance	Negative Variance (Percent)
G	17	26%	14	22%
A	30	28%	5	5%
B	15	33%	0	0%
E	81	39%	12	6%
D	8	42%	2	11%
F	10	43%	6	26%
I	28	54%	7	13%
J	44	59%	34	46%
C	85	64%	6	5%
H	15	68%	11	50%

2006 - 2008 Planning Accuracy - Sorted by Percent Negative Variance				
Region	Total Variance	Total Variance (Percent)	Negative Variance	Negative Variance (Percent)
B	15	33%	0	0%
C	85	64%	6	5%
A	30	28%	5	5%
E	81	39%	12	6%
D	8	42%	2	11%
I	28	54%	7	13%
G	17	26%	14	22%
F	10	43%	6	26%
J	44	59%	34	46%
H	15	68%	11	50%

APPENDIX C

EXHIBIT C-1. LIST OF DIFFICULT SITES NOTED DURING INTERVIEWS AND FROM FOLLOW-UP CORRESPONDENCE

REGION	SITE ID	EPA ID	SITE NAME	PROBLEMS/COMMENTS
2	0202438	NY4571924451	GRIFFISS AIR FORCE BASE (11 AREAS)	Yes - vapor intrusion & institutional controls
2	0201168	NJ3210020704	PICATINNY ARSENAL (USARMY)	Yes - risk of reuse; Yes - unspecified
2	0202439	NY4571924774	PLATTSBURGH AIR FORCE BASE	Yes - munitions; Yes - unspecified
2	0204694	PRN000204694	ATLANTIC FLEET WEAPONS TRAINING AREA	Yes - munitions, Yes - unspecified
2	0201164	NJ2210020275	FORT DIX (LANDFILL SITE)	Yes - munitions
2	0202425	NY0213820830	SENECA ARMY DEPOT	Yes - munitions
2	0201162	NJ0570024018	MCGUIRE AIR FORCE BASE #1	Yes - inability to sign FF agreement; Yes - unspecified
2	0202755	NJ0890090012	MIDDLESEX SAMPLING PLANT (USDOE)	Yes- RA status, Army kept changing
3	0300420	MD0570024000	ANDREWS AIR FORCE BASE	Yes - not specified
3	0300435	MD9210020567	FORT GEORGE G. MEADE	Yes - mentioned munitions and difficulty negotiating FFA. Issued order in August; Yes - not specified
3	0303768	VA2800005033	LANGLEY AIR FORCE BASE/NASA LANGLEY RESEARCH CENTER	Yes - groundwater and mentioned munitions, Yes - unspecified
3	0302862	VA7170024684	NAVAL SURFACE WARFARE CENTER - DAHLGREN	Yes- munitions
3	0302458	PA2210090054	LETTERKENNY ARMY DEPOT (PDO AREA)	Yes - Army keeps changing it's mind
3	0302467	PA6213820503	LETTERKENNY ARMY DEPOT (SE AREA)	Yes - Army keeps changing it's mind
3	0300421	MD2210020036	ABERDEEN PROVING GROUND (EDGEWOOD AREA)	Yes - Aberdeen Ground - munitions - some munitions along shoreline. Some munitions are still in sediment - potential human health risk from trespassers.
3				Spring Valley - high visibility non-NPL site
3				Fort Detrick - state remediation issues
3				3-4 sites have issues with federal facility agreements with DOD.
4	0404147	TN0210020582	MILAN ARMY AMMUNITION PLANT	Yes- dispute about allowing plumes to migrate, natural attenuation. Land use control portion causing trouble too.
4	0404152	TN1890090003	OAK RIDGE RESERVATION (USDOE)	Yes - unilaterally (DOE) decided to change schedules; Yes - complexity

REGION	SITE ID	EPA ID	SITE NAME	PROBLEMS/COMMENTS
4	0404794	KY8890008982	PADUCAH GASEOUS DIFFUSION PLANT (USDOE)	Yes - complex site
4	0401205	FL1570024124	TYNDALL AIR FORCE BASE	Yes
5				DOI site represents number of challenges and headaches. Chanute Air Force Base is unmanageable. Region tried to get it on NPL in 90's because 1) wanted a fed facility schedule to set parameters to guide cleanup 2) it would get higher funding priority. Problem with air force (HQ is aware of issue) and it is unique.
5	0501181	IL8143609487	SANGAMO ELECTRIC DUMP/CRAB ORCHARD NATIONAL WILDLIFE REFUGE (USDOI)	DOI site represents number of challenges and headaches.
6				Sandia national Labs in NM is a difficult DOE site.
6				Los Alamos national Labs in NM is a difficult DOE site.
6				Kelly Air Force Base in TX is very contentious. Community issues there.
7	0702020	NE2213820234	CORNHUSKER ARMY AMMUNITION PLANT	Yes - Nebraska ordinance plant - public involvement, PRPs, and army corps of engineers. Large city water supply nearby and they wrestled with the remedy interpretation as Army feels that they just need to contain contaminants.
7	0701773	MO5210021288	WELDON SPRING FORMER ARMY ORDNANCE WORKS	Yes - Huge site with lot of public interest.
7	0700743	KS6214020756	FORT RILEY	Yes - Fort Riley has 101,000 acres active base. Did ton of work and continue to do work. Looked at construction completion for 2009 or 2010 and MMRP issue arose now Construction Complete is pushed back to 2016. No funding to address MMRP contamination; Yes - NE ordinance plant they've wrestled on the remedy with the Army. Containment vs. source control that would impact # of wells they'd put in. The other site has an aquifer set for drinking water that should be cleaned up, others say monitoring natural attenuation is fine because nobody is using it now. munitions

REGION	SITE ID	EPA ID	SITE NAME	PROBLEMS/COMMENTS
7				Hastings - former Naval Ammunition Depot - groundwater has future risk as aquifer as classified as usable therefore it should be cleaned to drinking water standards.
8	0800359	CO7570090038	AIR FORCE PLANT PJKS	Yes
8	0800755	UT3213820894	TOOELE ARMY DEPOT (NORTH AREA)	Yes
8	0800017	WY5571924179	F.E. WARREN AIR FORCE BASE	Yes- munitions
8	0800753	UT0571724350	HILL AIR FORCE BASE	Yes- munitions
8	0800357	CO5210020769	ROCKY MOUNTAIN ARSENAL (USARMY)	Rocky flats and arsenal were bad because they were big and complex.
9	0902759	CA4570024337	MCCLELLAN AIR FORCE BASE (GROUND WATER CONTAMINATION)	Yes - Big sites with early transfers
9	0902762	CA4971520834	TRACY DEFENSE DEPOT (USARMY)	Yes - Construction complete snafu with other agency.
10			unspecified	Yes, there are two sites specifically. One is munitions and one is perchlorate and TCE vapor intrusion.
10	1000128	AK4170024323	ADAK NAVAL AIR STATION	Yes - munitions
10	1001139	WA9571924647	FAIRCHILD AIR FORCE BASE (4 WASTE AREAS)	Yes - TCE and perchlorate
10	1000302	ID3572124557	MOUNTAIN HOME AIR FORCE BASE	Yes - problem not specified
10	1001113	WA3170090044	JACKSON PARK HOUSING COMPLEX (USNAVY)	Yes - munitions site
10	1001114	WA3890090076	HANFORD 100-AREA (USDOE)	Yes - huge site
10	1001118	WA4890090075	HANFORD 1100-AREA (USDOE)	Yes - huge site
10	1001105	WA1890090078	HANFORD 200-AREA (USDOE)	Yes - huge site
10	100111	WA2890090077	HANFORD 300-AREA (USDOE)	Yes - huge site

APPENDIX D

AN ANALYSIS OF OVERALL REGIONAL PERFORMANCE AND LEAD AGENCY

Our analysis of difficult sites found that sites led by either the Army or Air Force may minimally increase the chance of a site being called a difficult site. Exhibits D-1 and D-2 show the correlation of the number of sites lead by each of these agencies and total variance.

Exhibit D-1 shows a strong, positive correlation between the number of Army-led sites and total variance. This suggests that a greater number of such sites may lead to greater uncertainty in the planning process. It is also important to note, however, that the strong correlation does not hold when compared to negative, total percent, or negative percent variance.

On the contrary, as evidenced by Exhibit D-2, no strong correlation exists between the number of Air Force-led sites and overall performance. This also hold true when correlated with negative, total percent, or negative percent variance.

EXHIBIT D-1. CORRELATION BETWEEN THE NUMBER OF ARMY-LED SITES AND OVERALL PERFORMANCE

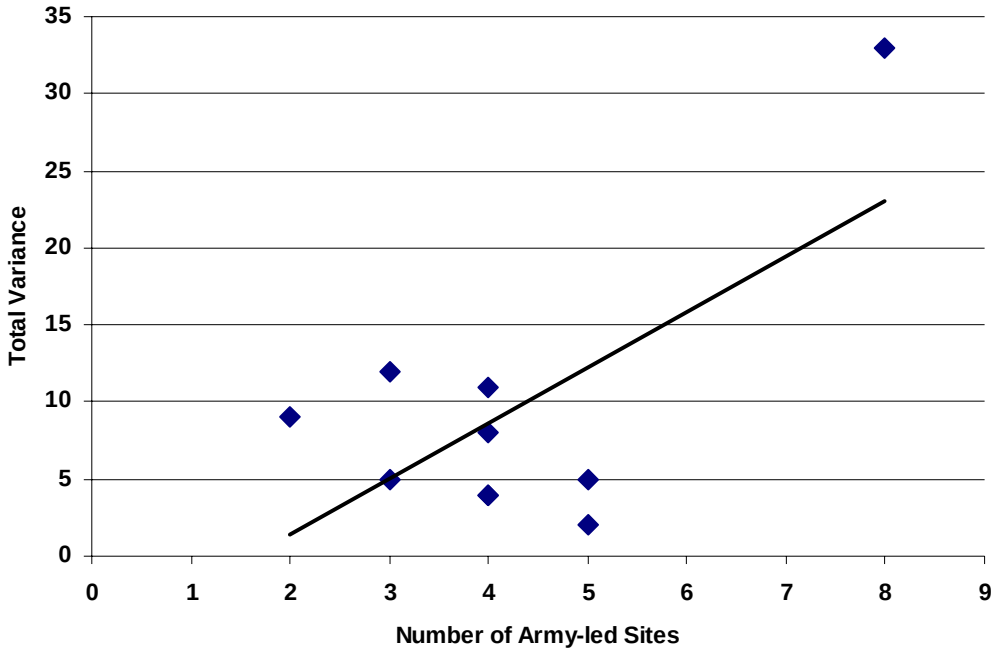
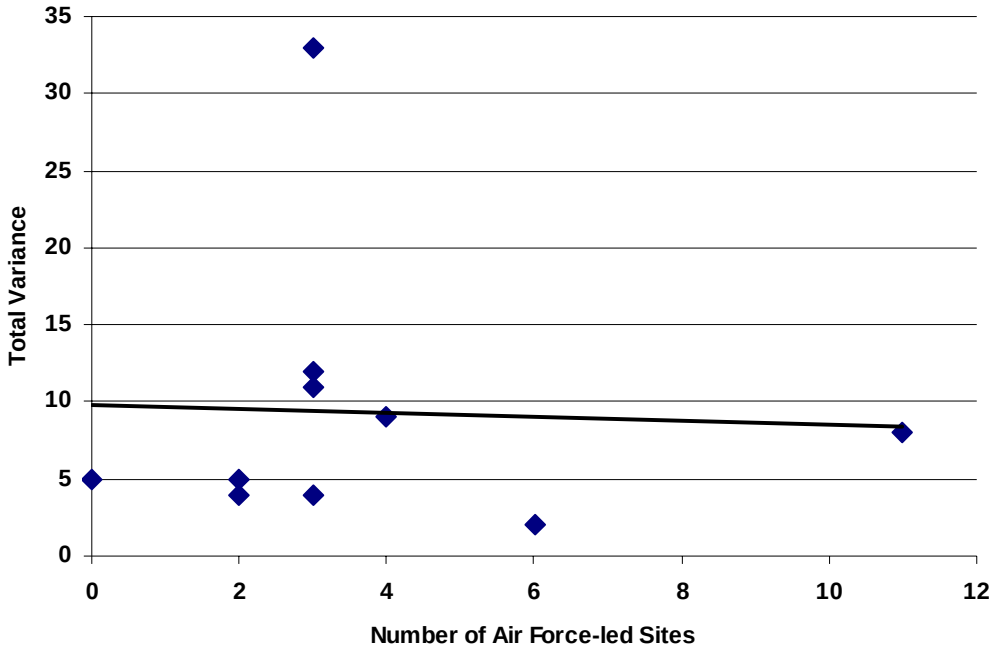


EXHIBIT D-2. CORRELATION BETWEEN THE NUMBER OF AIR FORCE-LED SITES AND OVERALL PERFORMANCE



APPENDIX E

AN ANALYSIS OF OVERALL REGIONAL PERFORMANCE AND CONTAMINATION TYPE

This section focuses on groundwater contamination and "other" media affected as proxy variables for site complexity. Each of type of contamination was tested against the four measures of regional performance: total variance, total percent variance, negative variance, and negative percent variance. Of these eight correlation tests, three resulted in significant correlations. Exhibits E-1, E-2, and E-3 below show these results.

EXHIBIT E-1. CORRELATION BETWEEN THE NUMBER OF GROUNDWATER-CONTAMINATED SITES AND TOTAL PERCENT VARIANCE

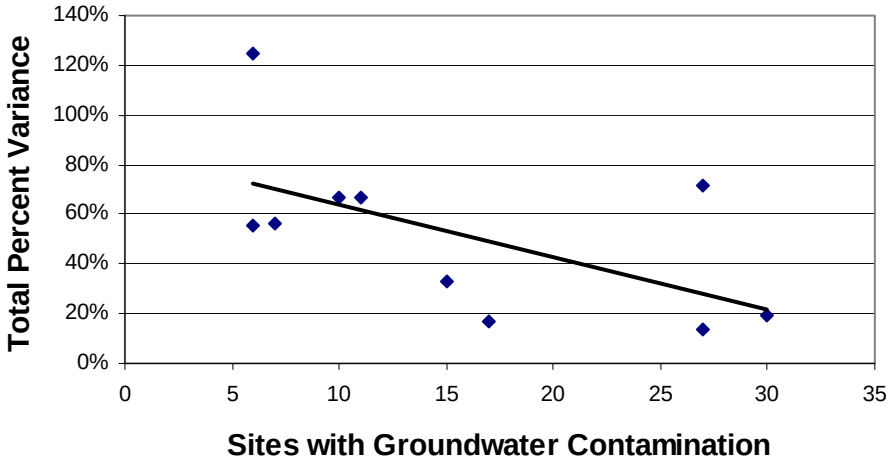


EXHIBIT E-2. CORRELATION BETWEEN THE NUMBER OF "OTHER"-CONTAMINATED SITES AND TOTAL PERCENT VARIANCE

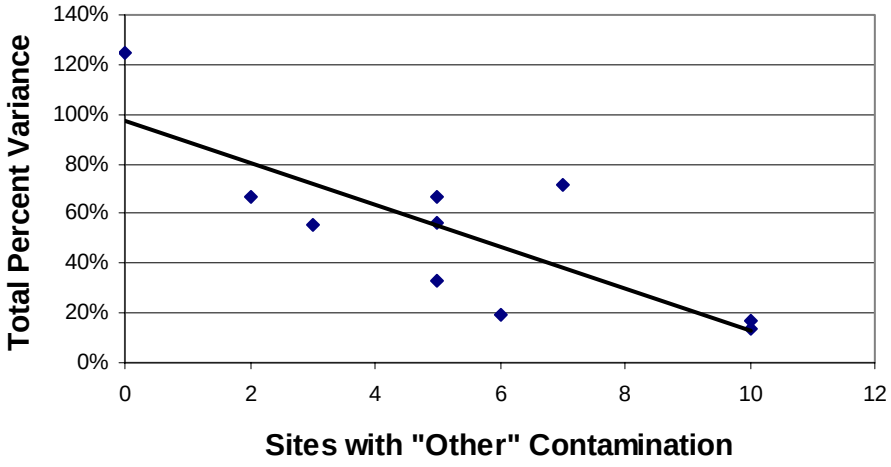
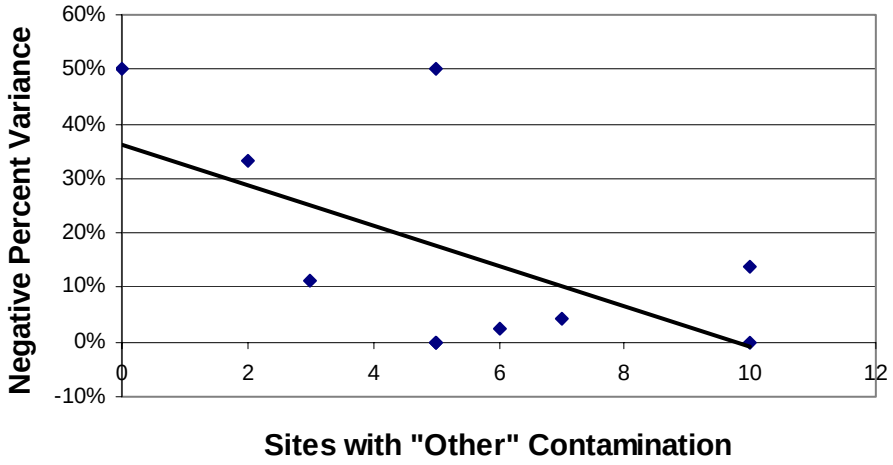


EXHIBIT E-3. CORRELATION BETWEEN THE NUMBER OF "OTHER"-CONTAMINATED SITES AND NEGATIVE PERCENT VARIANCE



The main point of significance from these three correlations is that they all suggest an inverse relationship between the frequency of sites of each type of contamination and overall performance relative to the number of targets each region sets. Each of these correlations, when analyzed using the total variance or negative variance, fails to show a significance relationship between the variables. This may suggest that if a region has a high number of groundwater or "other" contaminated sites, that region may be more likely to meet its targets because they could have a relatively smaller number of non-groundwater or "other" targets that could potentially cause planning issues. An alternate explanation could be that assuming these two types of contamination are the most complex to plan for, regions with relatively higher numbers of these sites take more time, and put more effort into correctly planning for them because they have to consider many more factors. A region with a relatively less complex set of sites may be more likely to gloss over the planning of those sites because they think they know how long it will take to meet each target and that can lead to an inferior planning schedule.

APPENDIX F

RELATIONSHIPS BETWEEN EACH TARGET AND EACH OVERALL REGIONAL PERFORMANCE METRIC

The following series of graphs depict the relationships between each of the eight targets (RI/FS/RFI Starts, Decision Documents, Final Remedy Selected, RA/CMI Starts, RA/CMI Completions, NPL Construction Completions, Sitewide RAU, and Five-Year Review Completions) and each of the four methods of measures regional performance (total variance, negative variance, total percent variance, and negative percent variance).

EXHIBIT F-1. RELATIONSHIP BETWEEN RI/FS STARTS AND TOTAL VARIANCE

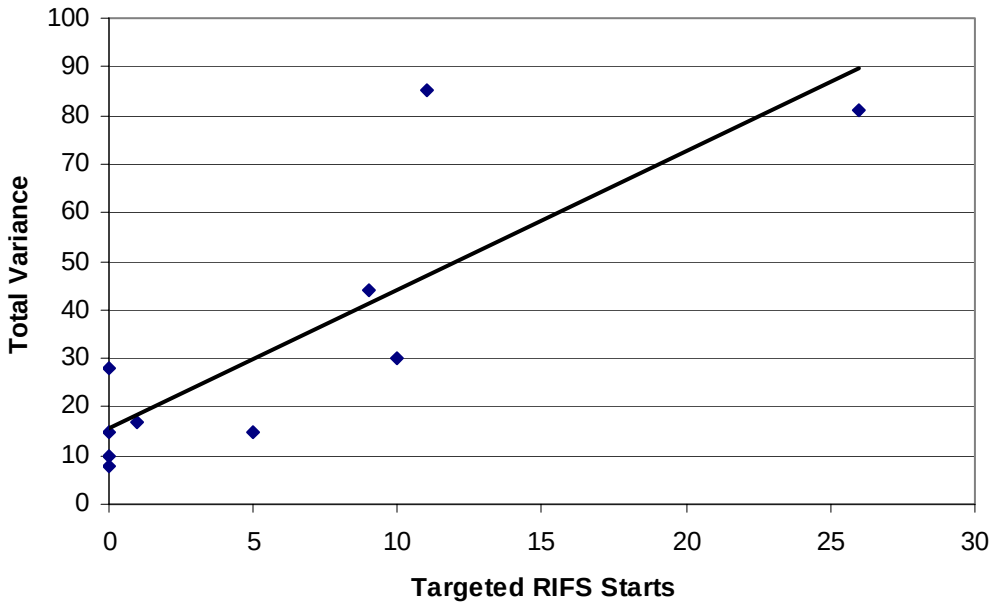


EXHIBIT F-2. RELATIONSHIP BETWEEN RI/FS STARTS AND NEGATIVE VARIANCE

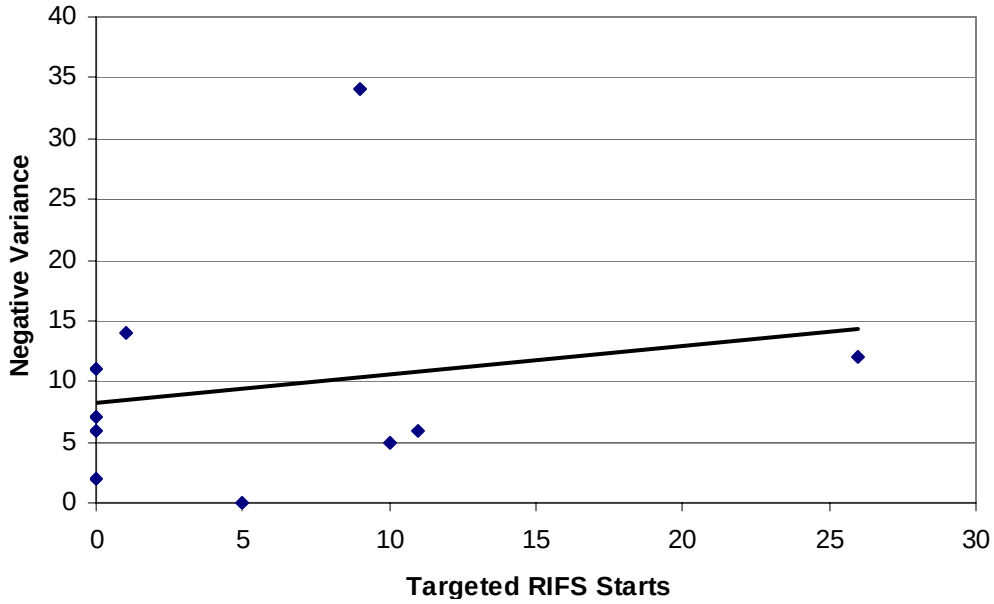


EXHIBIT F-3. RELATIONSHIP BETWEEN RI/FS STARTS AND TOTAL PERCENT VARIANCE

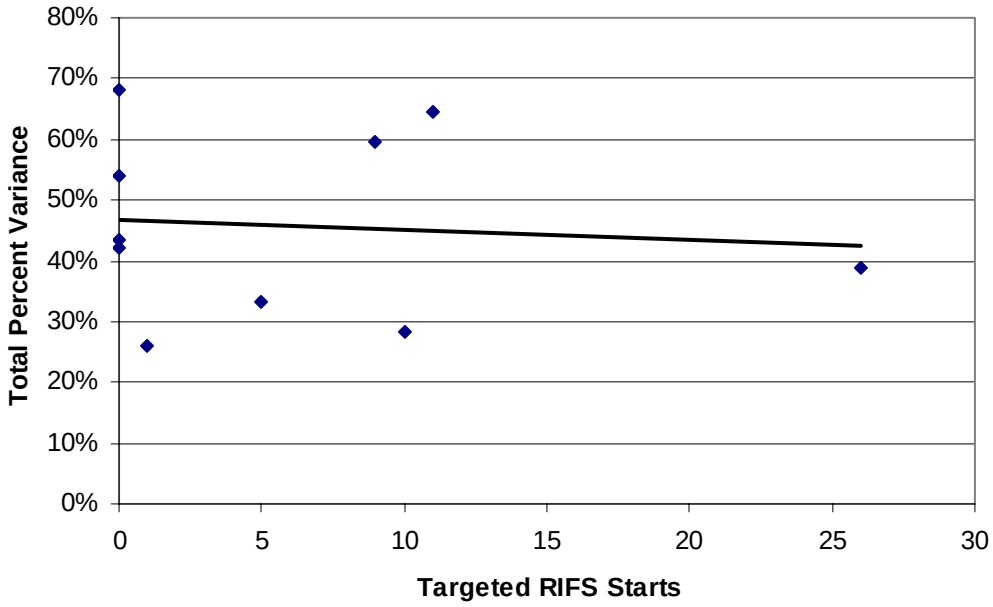


EXHIBIT F-4. RELATIONSHIP BETWEEN RI/FS STARTS AND NEGATIVE PERCENT VARIANCE

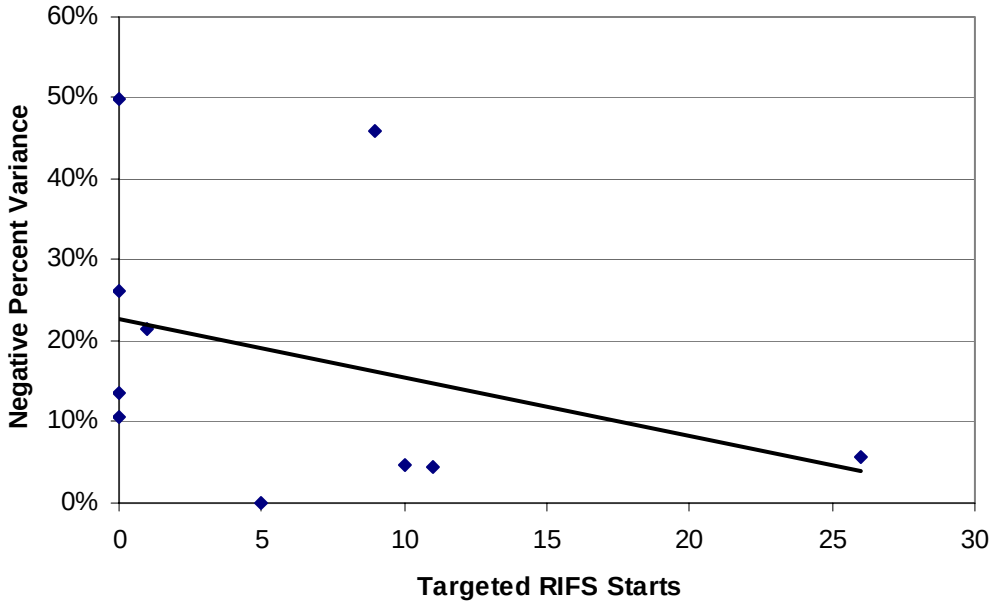


EXHIBIT F-5. RELATIONSHIP BETWEEN DECISION DOCUMENTS AND TOTAL VARIANCE

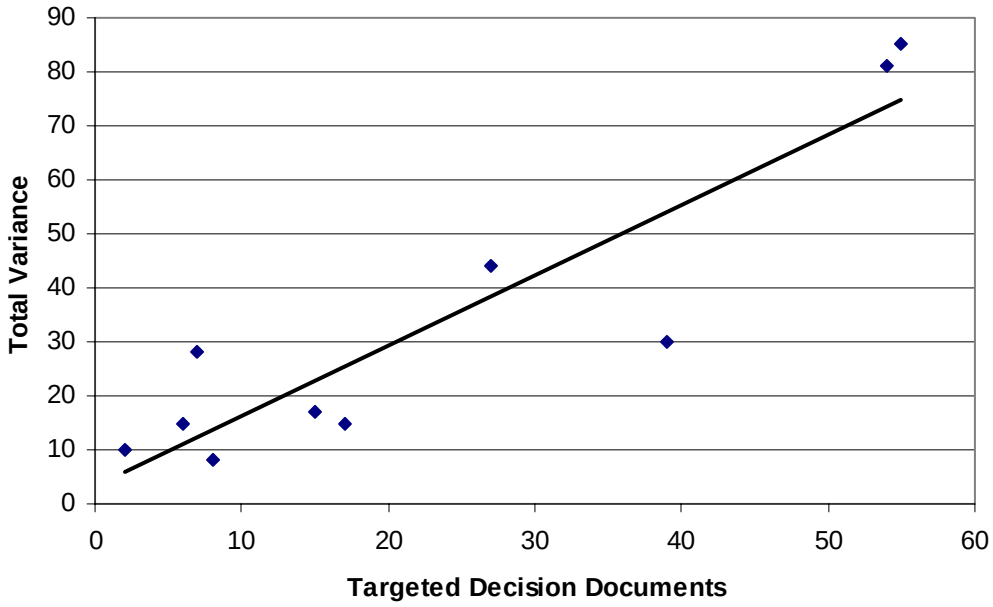


EXHIBIT F-6. RELATIONSHIP BETWEEN DECISION DOCUMENTS AND NEGATIVE VARIANCE

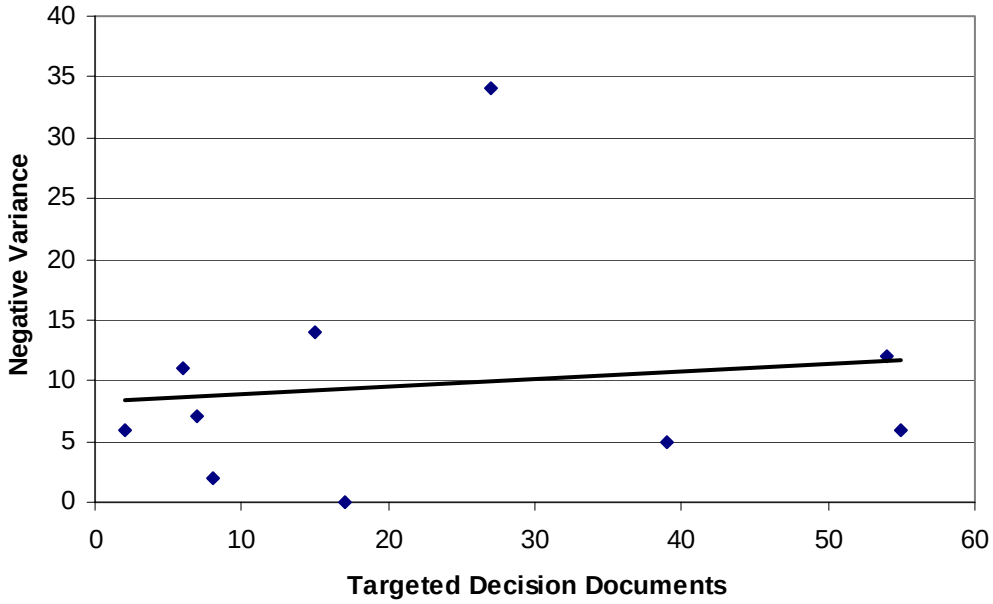


EXHIBIT F-7. RELATIONSHIP BETWEEN DECISION DOCUMENTS AND TOTAL PERCENT VARIANCE

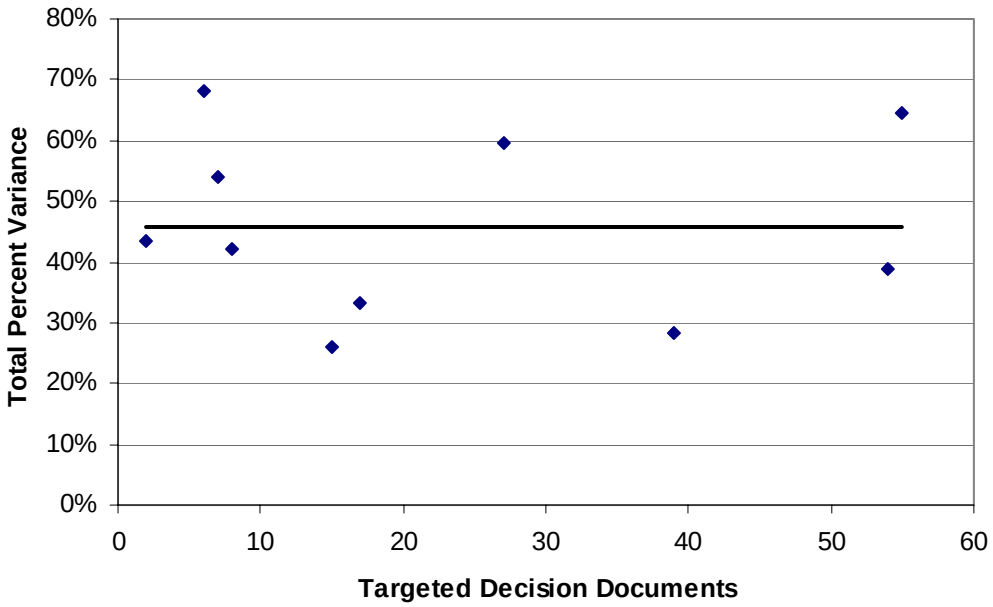


EXHIBIT F-8. RELATIONSHIP BETWEEN DECISION DOCUMENTS AND NEGATIVE PERCENT VARIANCE

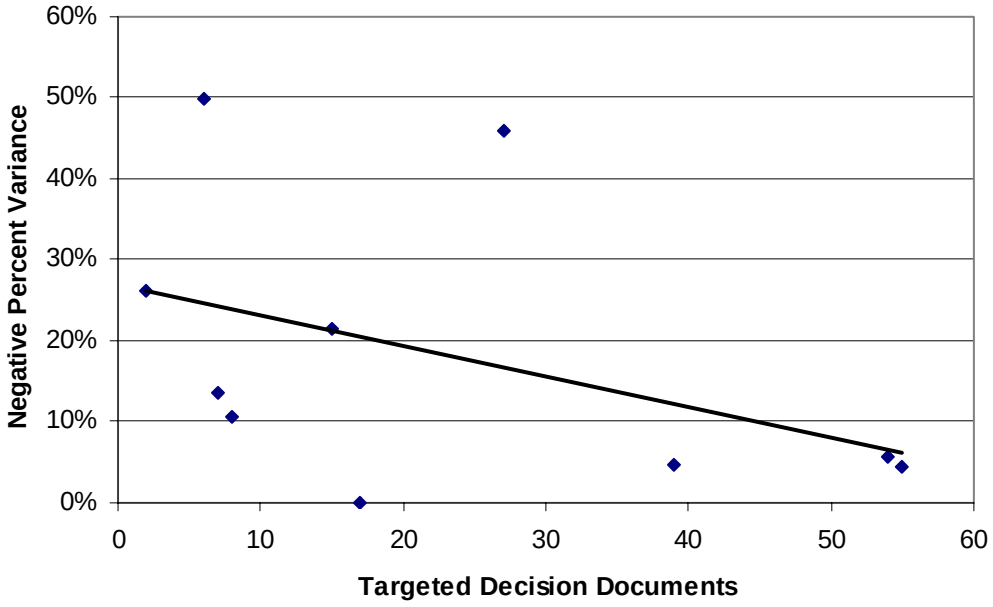


EXHIBIT F-9. RELATIONSHIP BETWEEN FINAL REMEDIES AND TOTAL VARIANCE

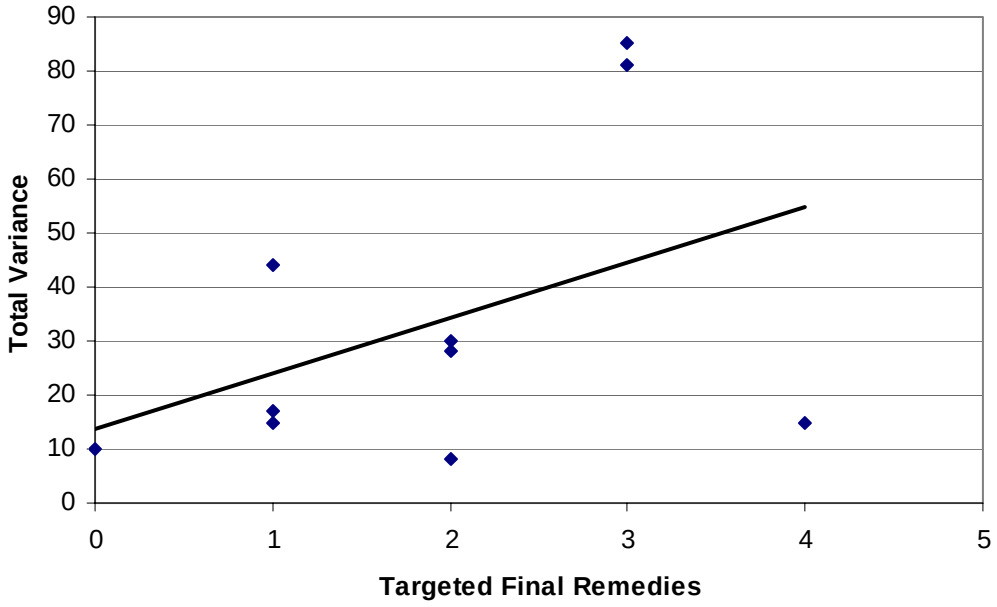


EXHIBIT F-10. RELATIONSHIP BETWEEN FINAL REMEDIES AND NEGATIVE VARIANCE

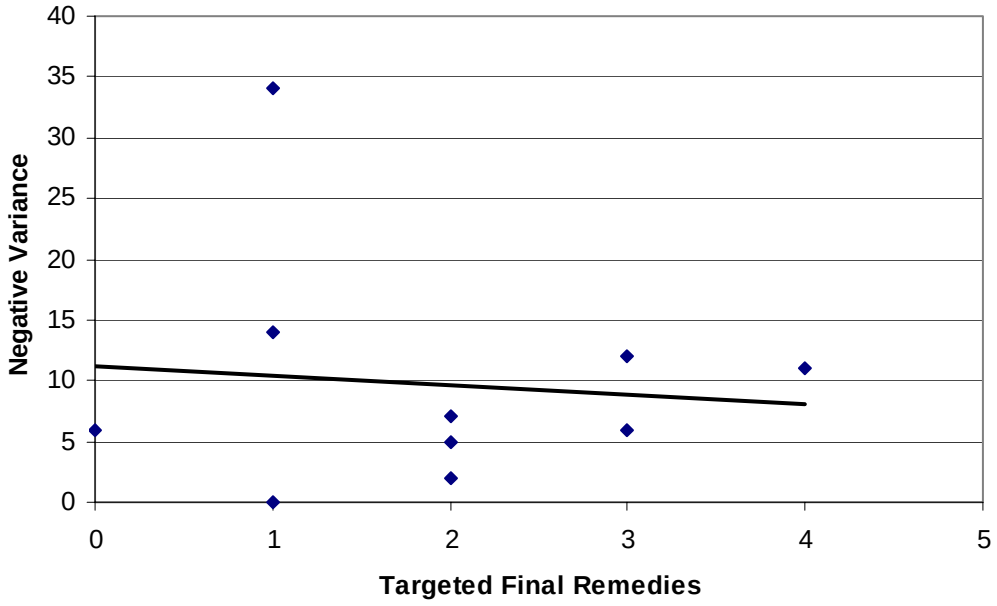


EXHIBIT F-11. RELATIONSHIP BETWEEN FINAL REMEDIES AND TOTAL PERCENT VARIANCE

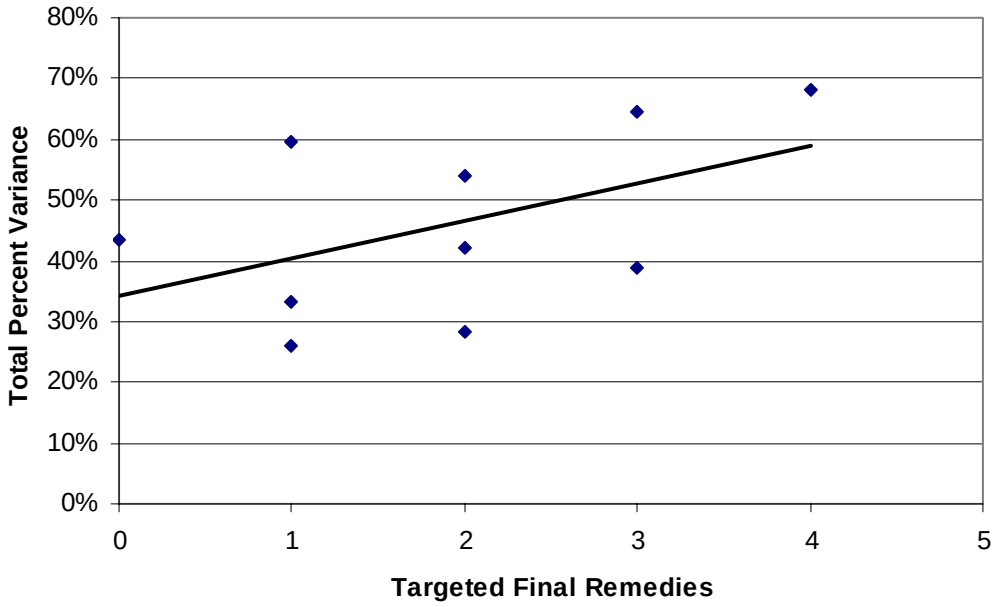


EXHIBIT F-12. RELATIONSHIP BETWEEN FINAL REMEDIES AND NEGATIVE PERCENT VARIANCE

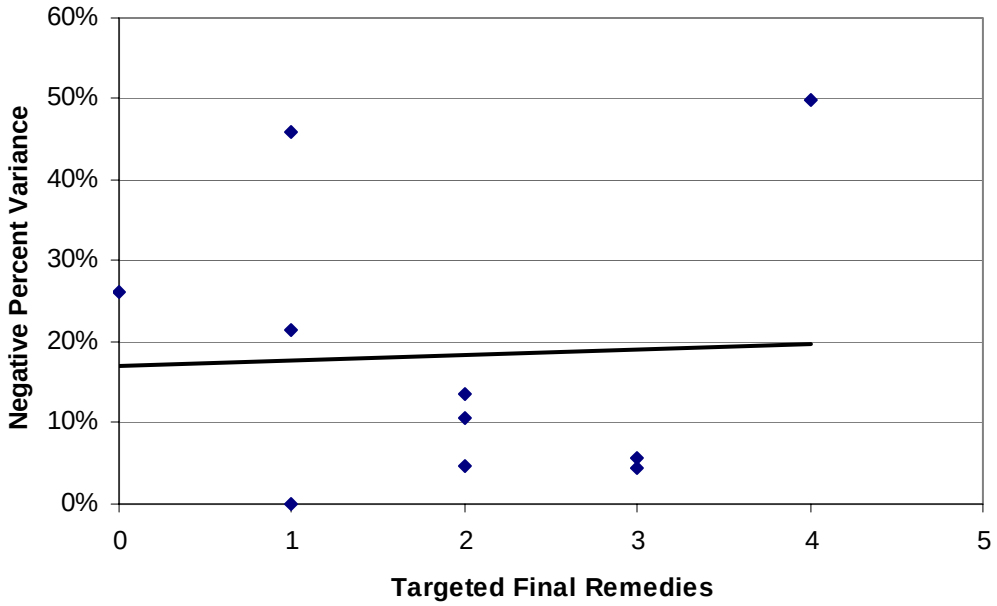


EXHIBIT F-13. RELATIONSHIP BETWEEN RA STARTS AND TOTAL VARIANCE

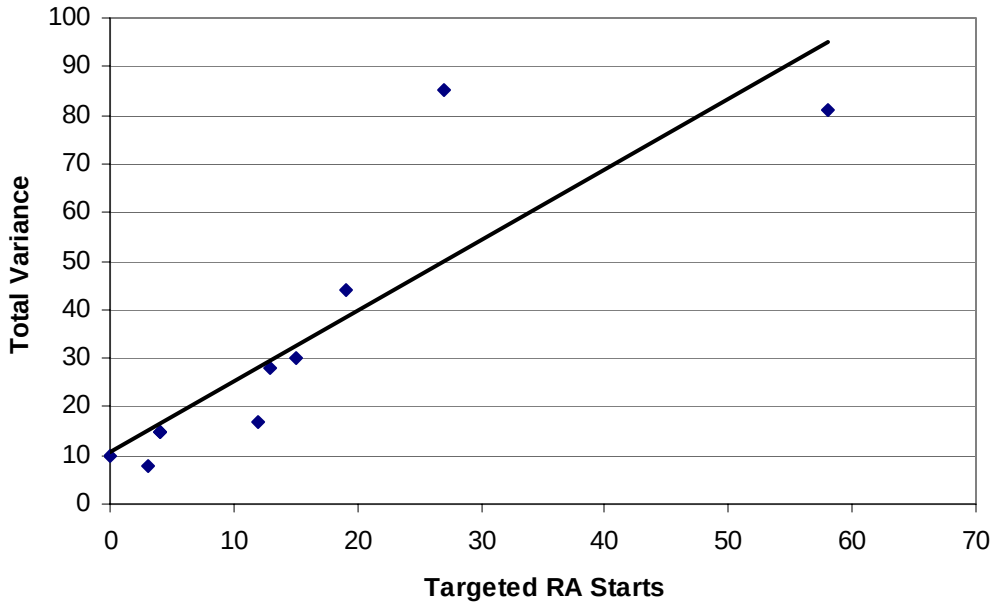


EXHIBIT F-14. RELATIONSHIP BETWEEN RA STARTS AND NEGATIVE VARIANCE

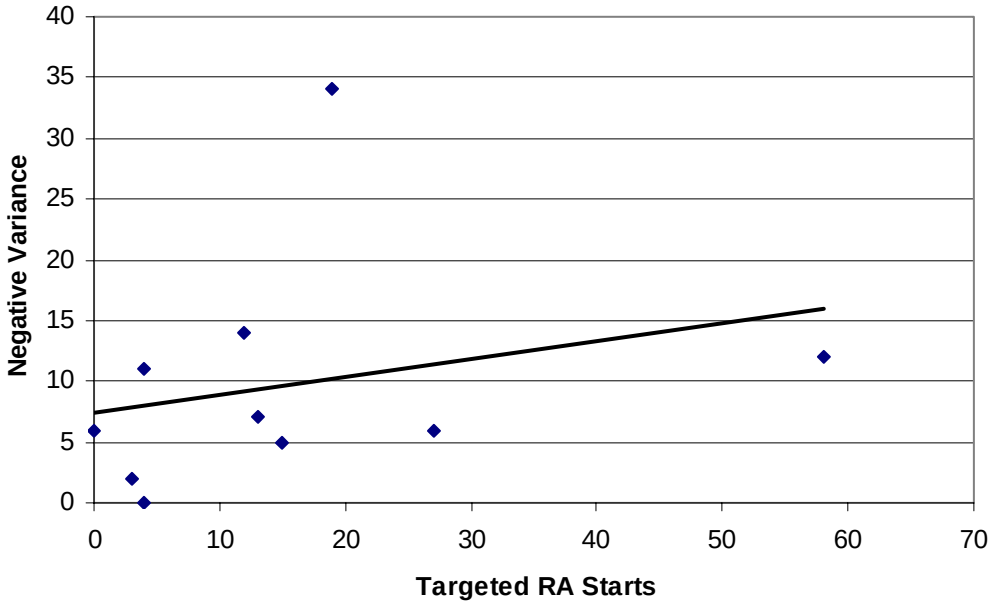


EXHIBIT F-15. RELATIONSHIP BETWEEN RA STARTS AND TOTAL PERCENT VARIANCE

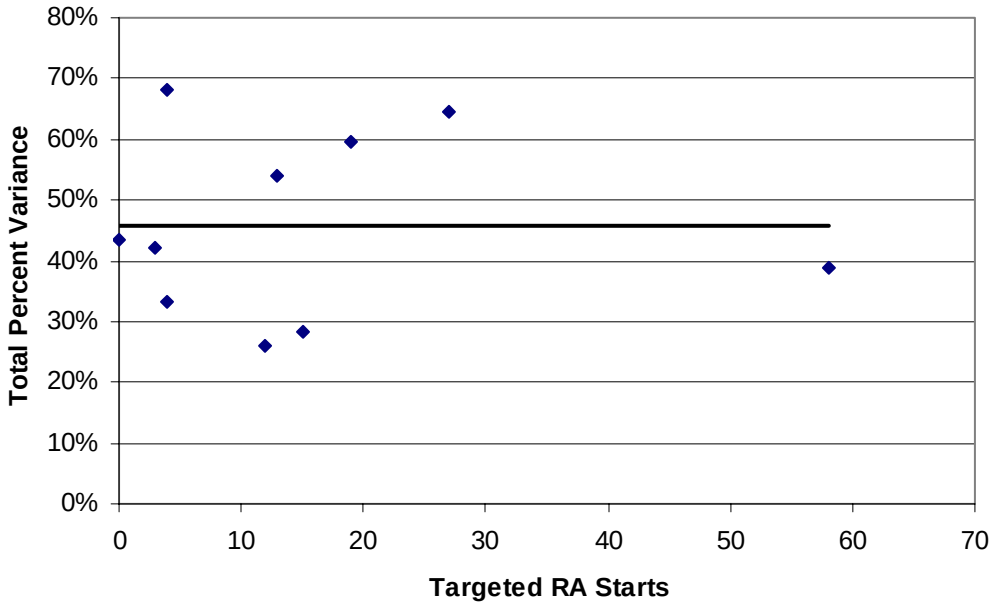


EXHIBIT F-16. RELATIONSHIP BETWEEN RA STARTS AND NEGATIVE PERCENT VARIANCE

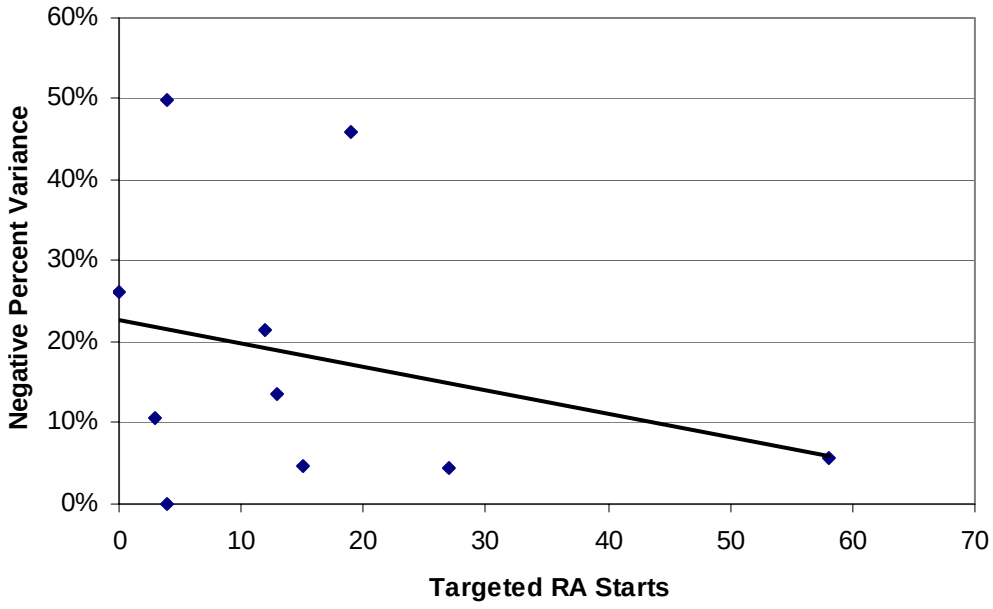


EXHIBIT F-17. RELATIONSHIP BETWEEN RA COMPLETIONS AND TOTAL VARIANCE

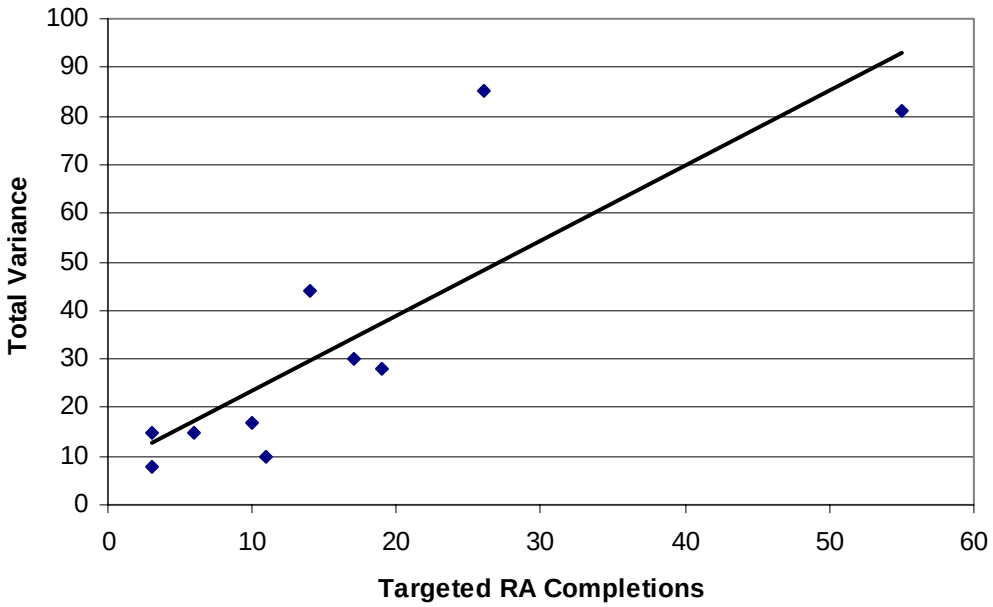


EXHIBIT F-18. RELATIONSHIP BETWEEN RA COMPLETIONS AND NEGATIVE VARIANCE

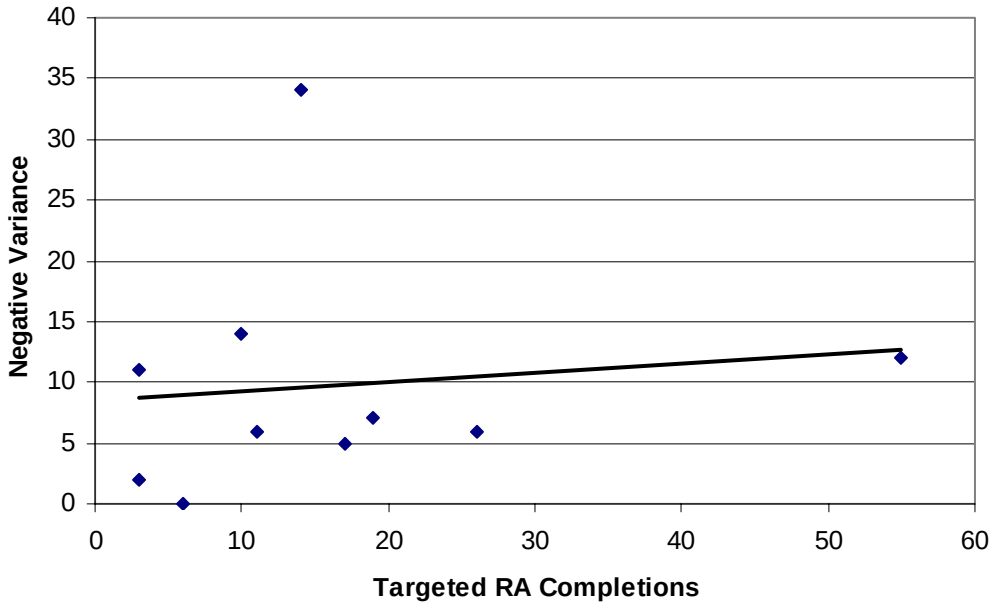


EXHIBIT F-19. RELATIONSHIP BETWEEN RA COMPLETIONS AND TOTAL PERCENT VARIANCE

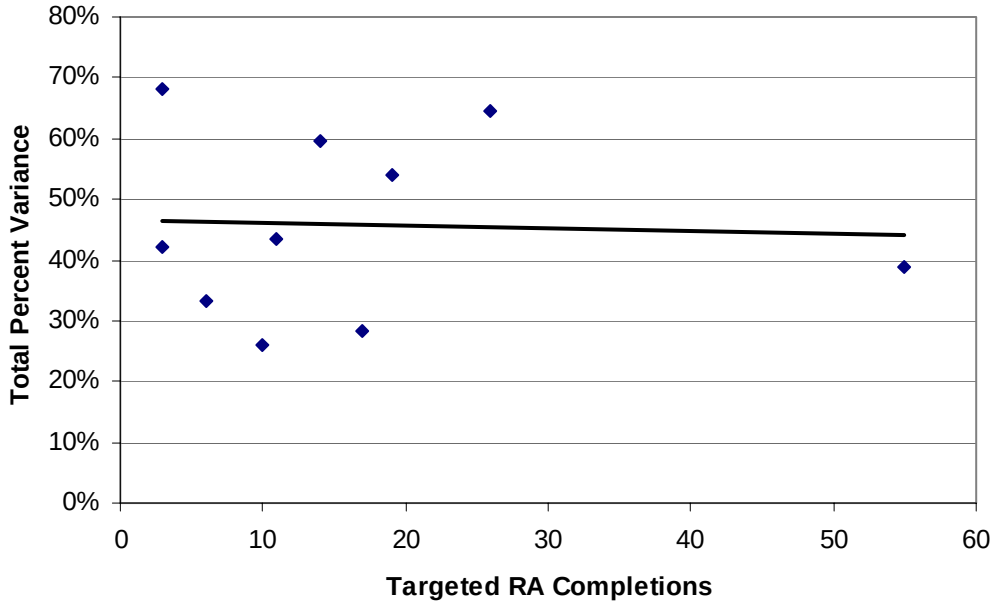


EXHIBIT F-20. RELATIONSHIP BETWEEN RA COMPLETIONS AND NEGATIVE PERCENT VARIANCE

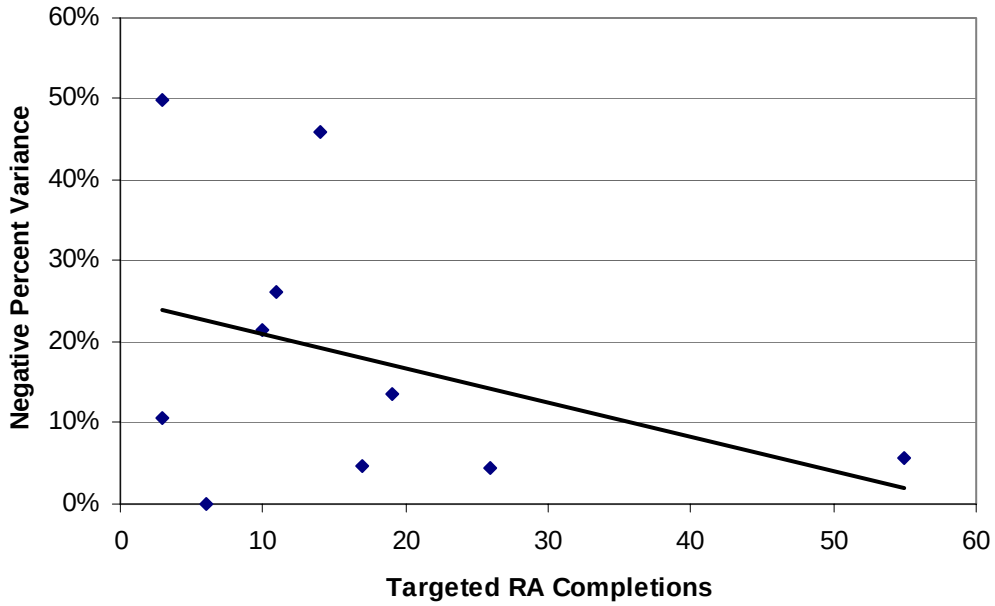


EXHIBIT F-21. RELATIONSHIP BETWEEN CONSTRUCTION COMPLETIONS AND TOTAL VARIANCE

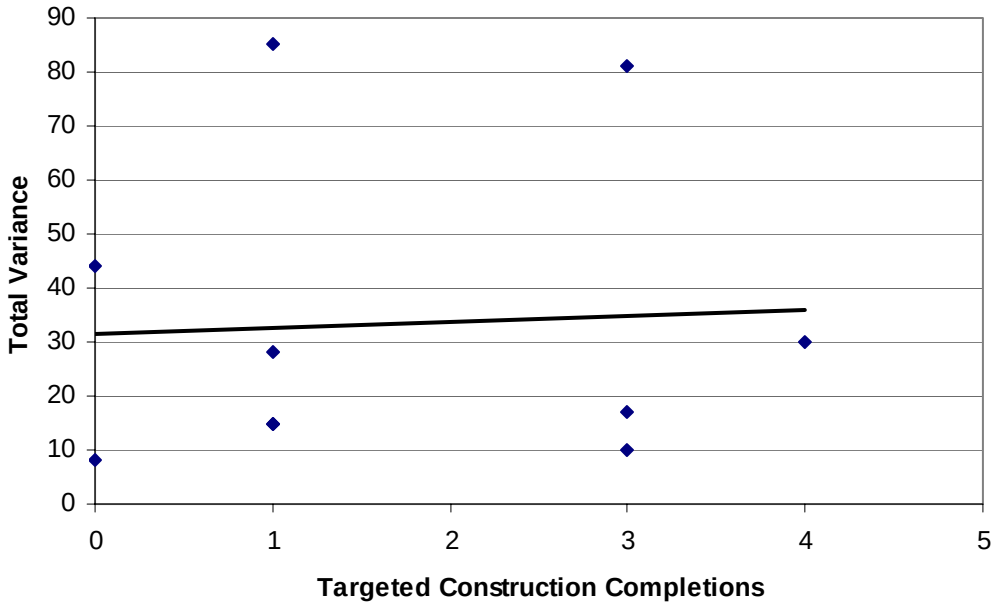


EXHIBIT F-22. RELATIONSHIP BETWEEN CONSTRUCTION COMPLETIONS AND NEGATIVE VARIANCE

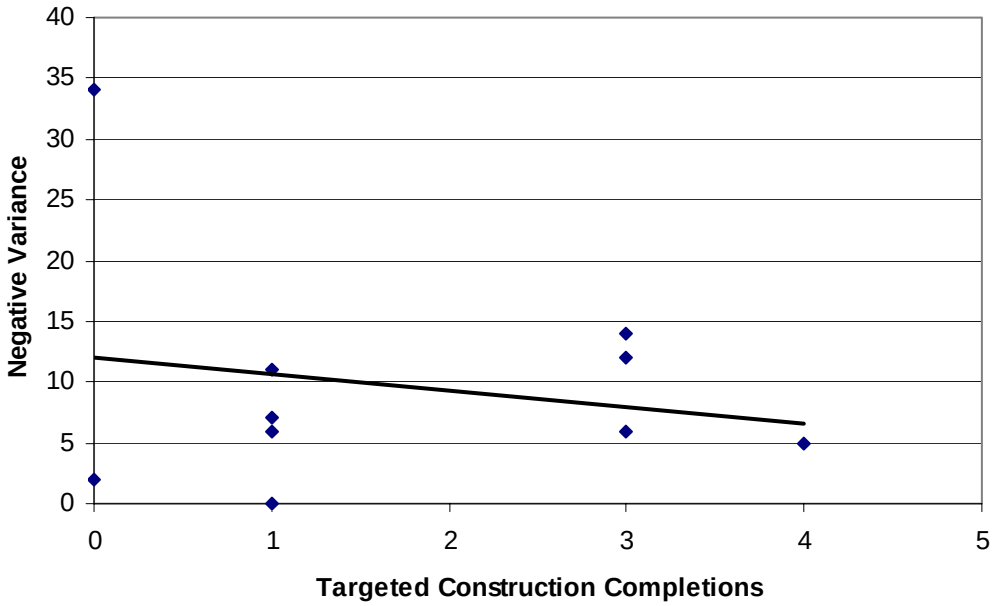


EXHIBIT F-23. RELATIONSHIP BETWEEN CONSTRUCTION COMPLETIONS AND TOTAL PERCENT VARIANCE

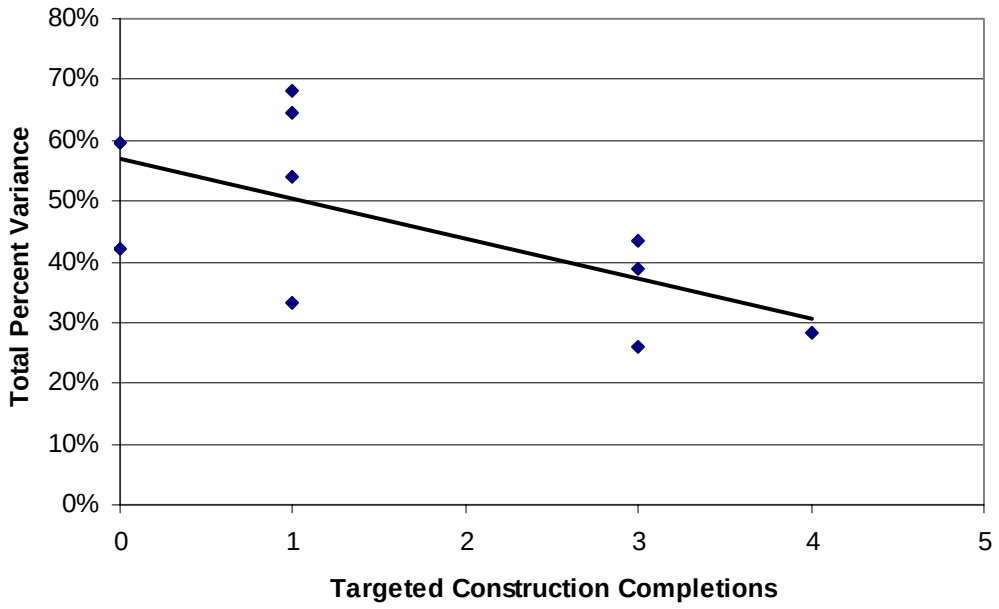


EXHIBIT F-24. RELATIONSHIP BETWEEN CONSTRUCTION COMPLETIONS AND NEGATIVE PERCENT VARIANCE

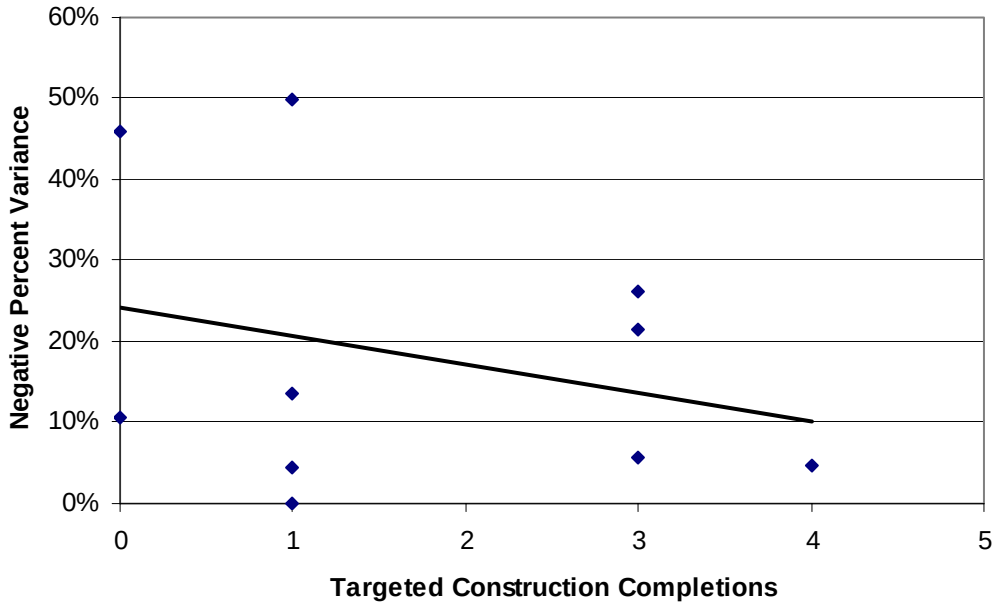


EXHIBIT F-25. RELATIONSHIP BETWEEN SITEWIDE RAUS AND TOTAL VARIANCE

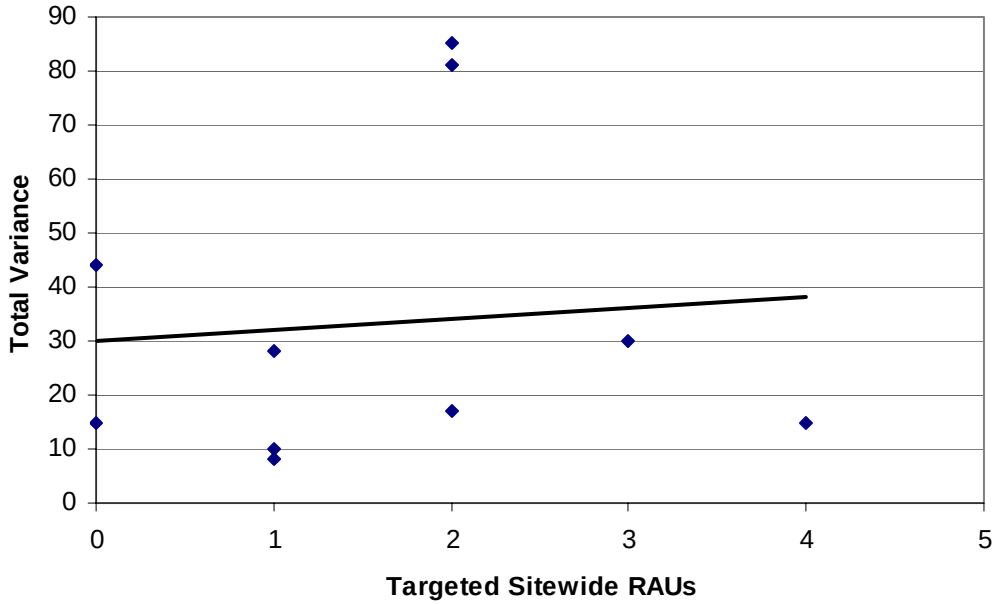


EXHIBIT F-26. RELATIONSHIP BETWEEN SITEWIDE RAUS AND NEGATIVE VARIANCE

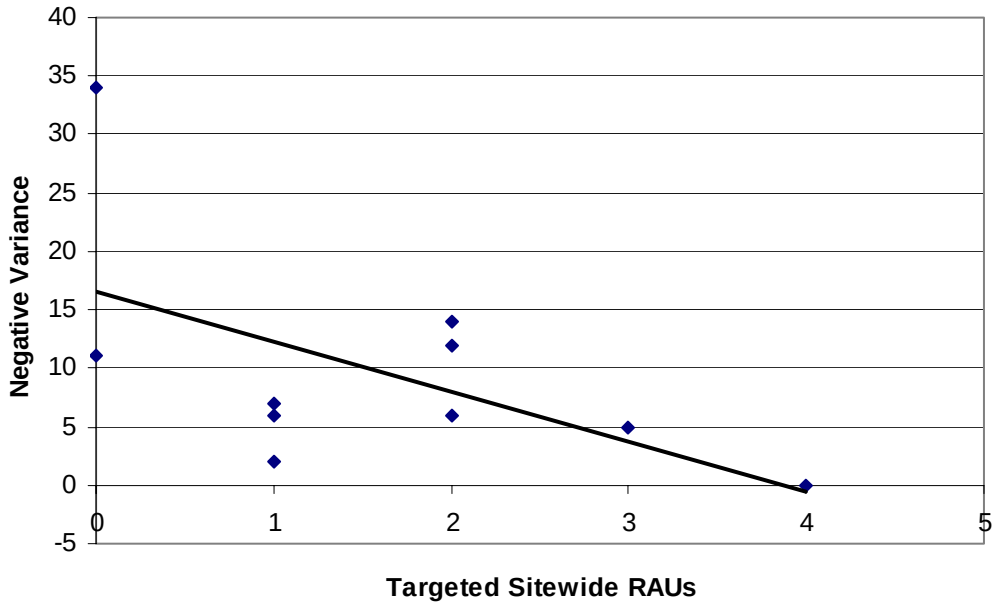


EXHIBIT F-27. RELATIONSHIP BETWEEN SITEWIDE RAUS AND TOTAL PERCENT VARIANCE

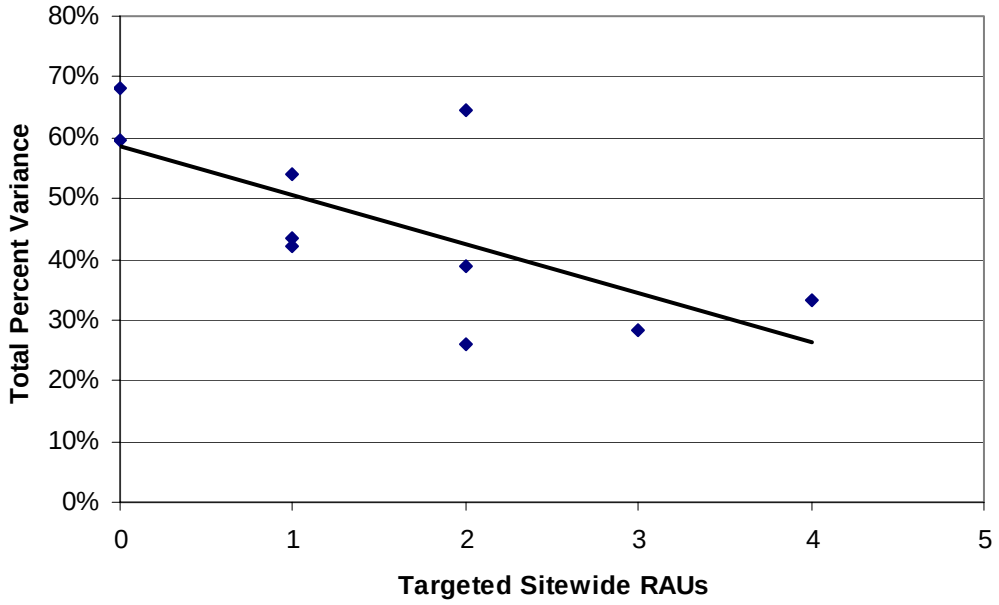


EXHIBIT F-28. RELATIONSHIP BETWEEN SITEWIDE RAUS AND NEGATIVE PERCENT VARIANCE

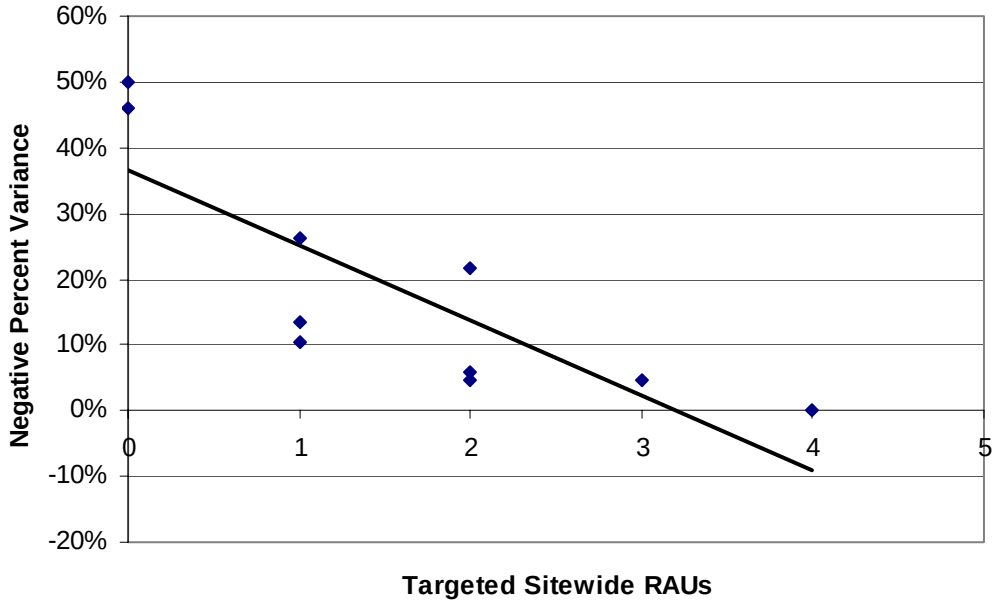


EXHIBIT F-29. RELATIONSHIP BETWEEN FIVE-YEAR REVIEWS AND TOTAL VARIANCE

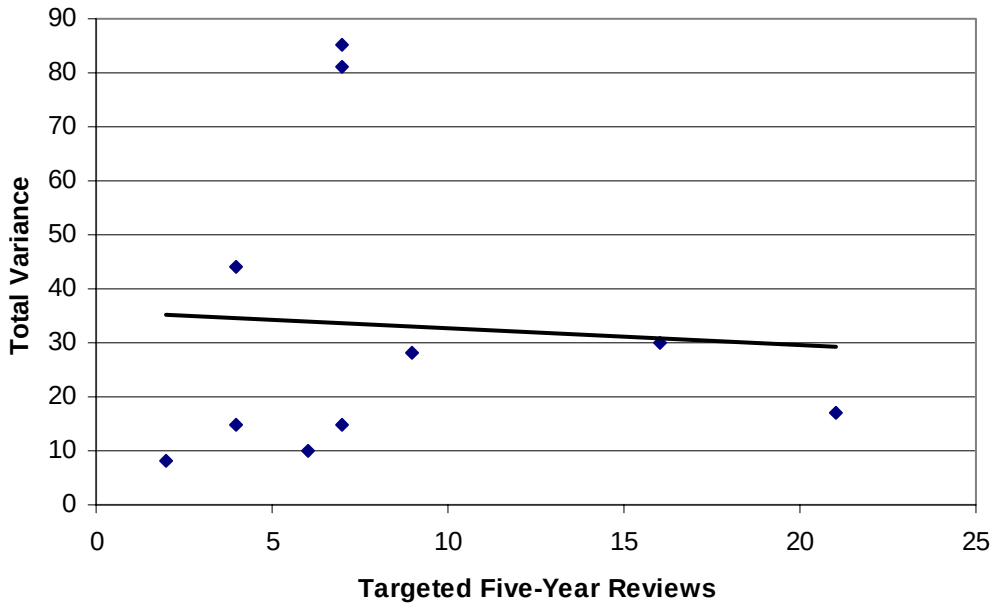


EXHIBIT F-30. RELATIONSHIP BETWEEN FIVE-YEAR REVIEWS AND NEGATIVE VARIANCE

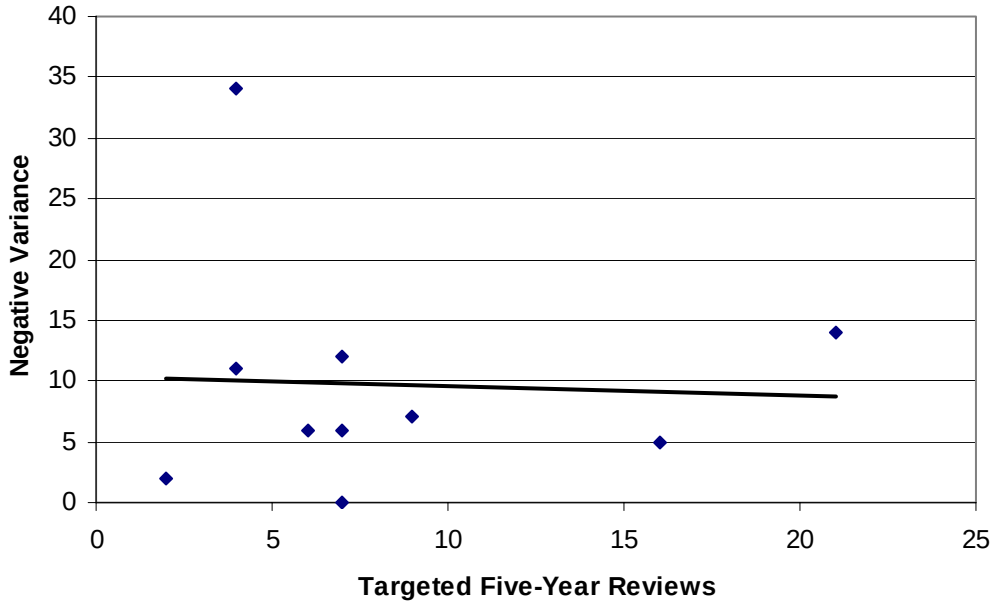


EXHIBIT F-31. RELATIONSHIP BETWEEN FIVE-YEAR REVIEWS AND TOTAL PERCENT VARIANCE

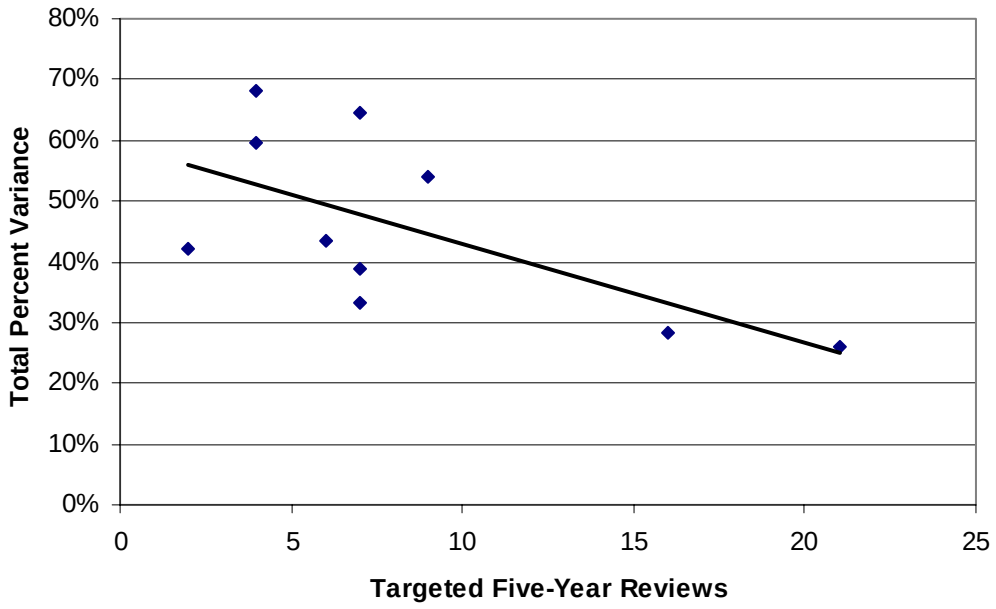
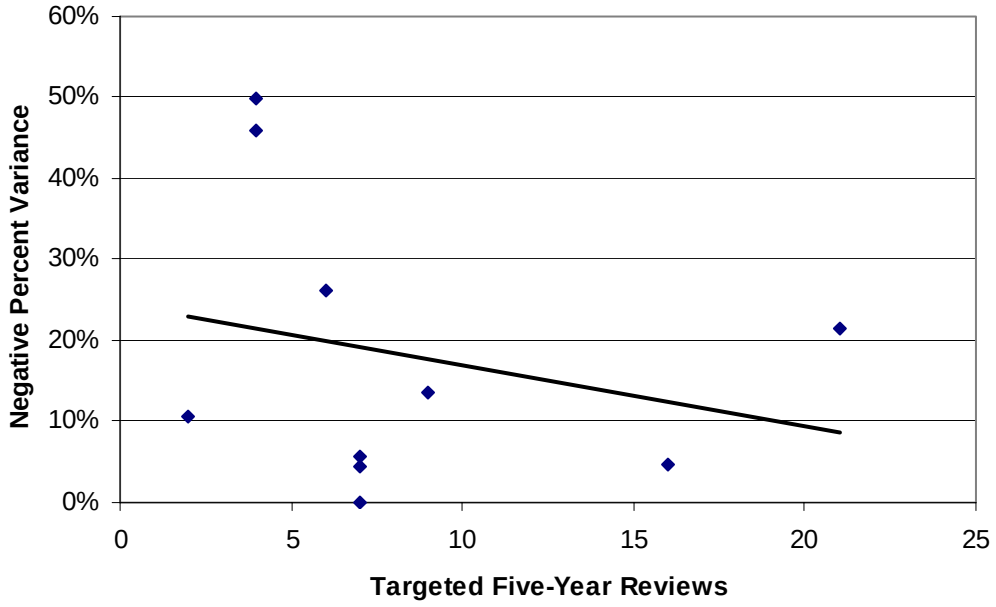


EXHIBIT F-32. RELATIONSHIP BETWEEN FIVE-YEAR REVIEWS AND NEGATIVE PERCENT VARIANCE



APPENDIX G

ANALYSIS OF EPA'S WORKLOAD PLANNING MODEL

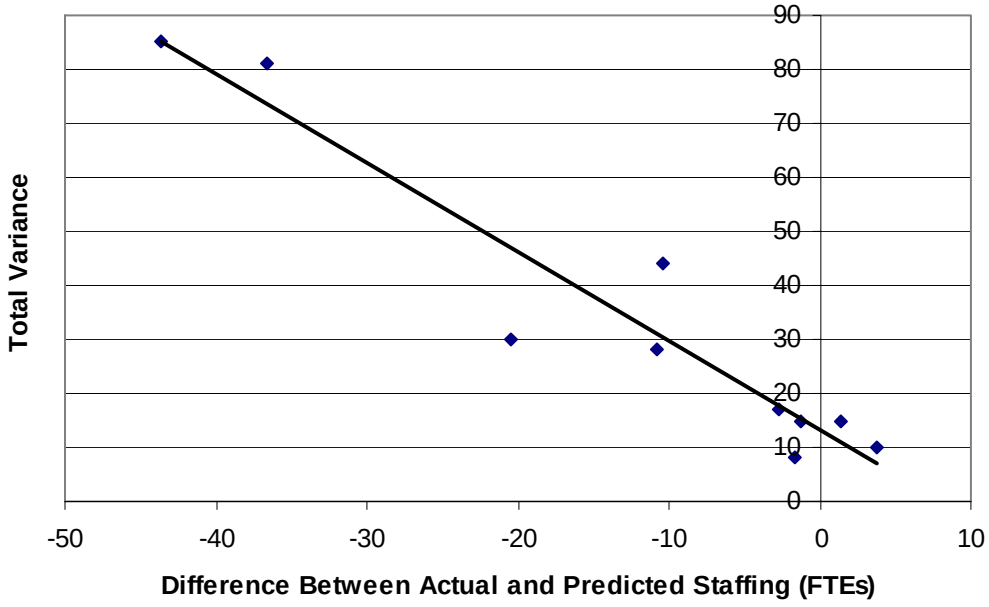
Exhibit G-1 below shows the model's suggested staffing level for FY 2008, the actual staffing level for FY 2007, and the difference between the two.

EXHIBIT G-1. PREDICTED AND ACTUAL STAFFING LEVELS

REGION	PREDICTED FY 2008 FTE LEVEL	ACTUAL FY 2008 FTE LEVEL	DIFFERENCE FROM PREDICTED TO ACTUAL
1	10.7	9.5	-1.2
2	18.3	7.9	-10.4
3	58.8	15.2	-43.6
4	53.9	17.3	-36.6
5	5.6	9.4	3.8
6	2.9	4.3	1.4
7	7.2	5.6	-1.6
8	22.3	11.5	-10.8
9	41.3	20.9	-20.4
10	20.2	17.5	-2.7

The following exhibit, G-2 shows the correlation between total variance and the difference in staffing level as shown above.

EXHIBIT G-2. CORRELATION BETWEEN THE DIFFERENCE IN STAFFING LEVEL AND TOTAL VARIANCE



The correlation shown above is highly significant and suggests that the regions that experienced actual staffing levels below what EPA's workload model suggests, were more likely to have higher total variances. The slope of the regression line also suggests that for each FTE short the region was, as compared to the model's estimate, a region could expect its total variance to increase by approximately 1.65.