



Users' Guide

Cost Estimation Toolkit (CET)

Version 2.4

September, 2008

Greg Hunolt, Bud Booth, Mel Banks, SGT

Funded by NASA Contract NNG05CA99C



Table of Contents

Section	Topic	Page
1.0	Introduction to the Cost Estimation Toolkit (CET)	6
1.1	Other CET Documentation	6
1.2	Notes on Version 2.4 of the CET	6
1.3	For More Information	6
2.0	Data Activity Cost Estimation Using the CET	7
2.1	Overview of Cost Estimation Using the CET	7
2.2	Preparing to Use the CET: Describing Your Data Activity	8
2.2.1	Data Activity Level Information	9
2.2.2	Functional Area Information – Operating Functions	9
2.2.3	Functional Area Information – Non-Operating Functions	12
2.3	Interpreting and Using the CET’s Life Cycle Cost Estimate	13
2.3.1	Life Cycle Cost Estimate, An Example	14
2.3.2	CET Quality Report	16
2.3.3	Interpreting the CET’s Life Cycle Cost Estimate - Overall Notes	18
2.3.4	Interpreting the CET’s Life Cycle Cost Estimate - Itemized Notes	20
3.0	Cost Estimation Toolkit Users’ Guide	24
3.1	Setting Up the CET	24
3.1.1	What the CET Package Consists Of	24
3.1.2	How to Install and Start the CET	25
3.2	Beginning: Selecting an Activity Dataset	26
3.3	Define New Activity Dataset	27
3.3.1	New Site / Activity Level Information	28
3.3.2	New Ingest Information	30
3.3.3	New Processing Information	31
3.3.4	New Documentation Information	33
3.3.5	New Archive Information	33
3.3.6	New Access and Distribution Information	35
3.3.7	New User Support Information	38
3.3.8	New Non-Operating Functional Areas Information	38
3.4	Work With an Existing Activity Dataset	41
3.4.1	Modify Activity / Site Level Information	41
3.4.2	Modify Ingest Information	42
3.4.3	Modify Processing Information	43
3.4.4	Modify Documentation Information	44
3.4.5	Modify Archive Information	45
3.4.6	Modify Access and Distribution Information	45
3.4.7	Modify User Support Information	46
3.4.8	Modify Non-Operating Functional Areas Information	46
3.5	Produce Life Cycle Cost Estimate and Quality Report	47
3.6	Generate Graphs	48
3.7	Perform Sensitivity Test	56
3.7.1	Background and Purpose of the Sensitivity Test	56
3.7.2	Sensitivity Test and the ‘What-If’ Function	56
3.7.3	Sensitivity Test Menu	57
3.7.4	Example of Sensitivity Test Results Example	58
3.8	Save Your Activity Dataset	59

Section	Topic	Page
3.9	Run a “What If” on Your Current Activity Dataset	59
3.10	Review the CET’s Life Cycle Cost Estimate	61
3.11	Delete No Longer Needed Activity Datasets	61
3.12	Exiting from the CET	62
3.13	Reviewing the CET Life Cycle Cost Estimate	62
3.13.1	CET Estimate Review Categories	63
3.13.2	Running the CET Estimate Reviewer	64
3.14	Accessing and Using the CET Worksheets	68
3.14.1	CET Internal Worksheets	68
3.14.2	Activity Dataset Worksheets	70
3.14.3	Saved Revised Life Cycle Cost Estimate Worksheets	70

Appendices

Section	Topic	Page
App A	Appendix A - CET DSP Reference Model - Functional Areas	71
A.1	Data Service Provider Reference Model	71
A.1.1	Functional Areas - Areas of Cost	71
A.1.2	Reference Model Parameters	72
A.1.3	Reference Model Requirements / Levels of Service	72
A.2	Functional Areas	72
A.2.1	Ingest	73
A.2.2	Processing	73
A.2.3	Documentation	74
A.2.4	Archive	74
A.2.5	Search and Order	75
A.2.6	Access and Distribution	75
A.2.7	User Support	76
A.2.8	Instrument / Mission Operations	76
A.2.9	Sustaining Engineering	76
A.2.10	Engineering Support	76
A.2.11	Technical Coordination	77
A.2.12	Implementation	77
A.2.13	Management	77
A.2.14	Facility / Infrastructure	78
App B	Appendix B - Activity Dataset Checklist	79
B.1	Data Activity Level ADS Information	79
B.2	ADS Information for Operating Functional Areas	80
B.2.1	Ingest ADS Information	81
B.2.2	Processing ADS Information	82
B.2.3	Documentation ADS Information	83
B.2.4	Archive ADS Information	84
B.2.5	Access and Distribution ADS Information	85
B.2.6	User Support ADS Information	88
B.3	ADS Information for Non-Operating Functional Areas	88
B.3.1	Implementation ADS Information	88
B.3.2	Sustaining Engineering ADS Information	89

Section	Topic	Page
B.3.3	Engineering Support ADS Information	89
B.3.4	Technical Coordination ADS Information	89
B.3.5	Miscellaneous Non-Staff Cost Items	90
App C	“Moxie”, an Example of an Activity Dataset	91
C.1	Activity Level Information for Moxie	91
C.2	Ingest Information for Moxie	91
C.3	Processing Information for Moxie	92
C.4	Documentation Information for Moxie	92
C.5	Archive Information for Moxie	92
C.6	Access and Distribution Information for Moxie	93
C.7	User Support Information for Moxie	93
C.8	Sustaining Engineering, Engineering Support & Implementation Information for Moxie	93
C.9	Technical Coordination Information for Moxie	94
C.10	Miscellaneous Non-Staff Cost Information for Moxie	94
C.11	“What-If” Changes to the Moxie Activity Dataset	94

List of Tables

Table	Title	Page
Table 1	Life Cycle Cost Estimate for the “Moxie” Activity Dataset	14
Table 2	Technical and Operational Staffing Breakdown for “Moxie” Activity Dataset	15
Table 3	CET Version 2.4 Performance Measurement	19
Table 4	What-If Differences for the “Moxie” Activity Dataset	60
Table 5	CET Estimate Review Effort Categories and Lines Each Contains	63
Table 6	CET Estimate Review Cost Categories and Lines Each Contains	64
Table 7	CET Visible Internal Worksheets	69

List of Figures

Figure	Title	Page
Figure 1	CET Concepts	8
Figure 2	CET Quality Report, Example	17
Figure 3	CET Title Worksheet	26
Figure 4	Select Activity Dataset Box	27
Figure 5	Select Functional Areas Box	28
Figure 6	Data Activity Information Box	29
Figure 7	Ingest Stream Information Box	30
Figure 8	Data Activity Processing Information Box	31
Figure 9	Processing Stream Information Box	32
Figure 10	Documentation Information Box	33
Figure 11	Archive Information Box	34
Figure 12	Distribution Information Box	35
Figure 13	Operational Distribution Stream Information Box	36
Figure 14	By Request Distribution Information Box	37
Figure 15	User Support Information Box	38
Figure 16	Implementation and Engineering Information Box	39
Figure 17	Technical Coordination Information Box	40

Figure	Title	Page
Figure 18	Miscellaneous Non-Staff Cost Information Box	40
Figure 19	Select Functional Areas to Modify Box	41
Figure 20	Delete or Modify Ingest Stream Box	42
Figure 21	Delete or Modify Processing Stream Box	44
Figure 22	Delete or Modify an Operational Distribution Stream Box	45
Figure 23	Life Cycle Cost Estimate Box	47
Figure 24	Graph Menu	48
Figure 25	Graph 1, Estimated Total FTE with Error Range	49
Figure 26	Graph 2, Total Mission Life Staffing Breakdown by Functional Area	49
Figure 27	Graph 3, Total Mission Life Staffing Breakdown by Labor Cost Category	50
Figure 28	Graph 4, Average Annual Staffing by Functional Area – Implementation	51
Figure 29	Graph 5, Average Annual Staffing by Functional Area – Operations	51
Figure 30	Graph 6, Total Estimated Costs, Staff vs Non-Staff	52
Figure 31	Graph 7, Total Staff Costs by Labor Category	52
Figure 32	Graph 8, Total Non-Staff Costs by Cost Category	53
Figure 33	Graph 9, Annual Non-Staff Costs by Category	53
Figure 34	Graph 10, What-If vs Base Estimate: Estimated Total FTE	54
Figure 35	Graph 11, What-If vs Base Estimate: FTE Changes by Category	55
Figure 36	Graph 12, What-If vs Base Estimate: Costs	55
Figure 37	Menu of Workload Parameters for Sensitivity Test	57
Figure 38	Sample Sensitivity Test Results	58
Figure 39	Choose Name for Activity Dataset to be Saved Box	59
Figure 40	Delete Activity Dataset Box	62
Figure 41	Estimate Review Category Menu Box	65
Figure 42	Review Estimate Screen Including Cost Estimate Line Review Box	66

1.0 Introduction to the Cost Estimation Toolkit (CET)

This document presents information you need to use the Cost Estimation Toolkit (CET) developed by SGT, Inc., for NASA. This Users' Guide is for the September 2008 version of the CET, Version 2.4. The CET employs the cost estimation by analogy approach, using information about existing data activities as the basis for estimating life cycle costs (i.e. from initial implementation through a period of operations) for your new data activity. The CET is a package of tools including the CET Estimator which you can use to prepare a life cycle cost estimate for a new data activity, and the CET Reviewer which you can use to review and tailor the CET estimate. The CET tools embody the SGT general data services provider reference model.

The remainder of Section 1 describes other documentation of the CET and the work on which it is based, and includes notes on the current version of the CET.

Section 2 provides an overview of the cost estimation process used by the CET Estimator (which for brevity will be referred to as simply the "CET") and guidance concerning preparing information needed by the CET and interpreting and using the estimates produced by the CET.

Section 3 is the Users' Guide proper, with detailed step-by-step information on installing, running, and using the CET.

Appendix A provides background information on the general Data Services Provider reference model upon which the CET is based, and describes the functional areas that are included in the model and how these are used in the CET. Appendix B contains the Activity Dataset Checklist, a more detailed description of the information required by the CET to describe a new data activity. Appendix C contains a sample Activity Dataset that is also included in the CET workbook.

1.1 Other CET Documentation

The CET Technical Description Document contains a detailed description of how the CET works; i.e. how the effort estimation by analogy is accomplished.

1.2 Notes on Version 2.4 of the CET

CET Version 2.4 is the September 2008 version of the toolkit. Version 2.4 provides the same capabilities as Version 2.3 and uses updated 2008 comparables information.

The CET includes an estimate of error for its effort (i.e., staffing level) estimates (discussed below in Section 2.3.3).

The CET and its documentation incorporate much of the feedback received from users of the earlier versions. Feedback from users has been of great value in pursuit of a primary goal of the CET effort: to develop a tool that is user friendly and that produces a genuinely useful product.

1.3 For More Information

For more information or help on using the CET, please contact any of the persons listed below:

Bud Booth, abooth@sgt-inc.com, 1-800-310-8161 ext 3147

Greg Hunolt, greghunolt@excel.net, 1-920-893-0422

Mel Banks, mbanks@sgt-inc.com, 1-800-310-8161 ext 3142

Kathy Fontaine, NASA POC, GSFC, Kathy.Fontaine@nasa.gov, 1-301-614-5582

2.0 Data Activity Cost Estimation Using the CET

This section will give you an overview of data activity cost estimation using the CET, including guidance on preparing the information the CET needs in order to produce an estimate, and in interpreting and using an estimate produced by the CET. Section 3, the detailed Users' Guide, will lead you, step-by-step, through the process of using the CET to generate a life cycle cost estimate for your new data activity.

2.1 Overview of Cost Estimation Using the CET

At the outset, it is assumed that you (or a group you are part of) are planning a new data activity for which you need to assemble a technical and cost proposal.

A 'data activity' (or, synonymously, a 'data service provider') is an activity that will provide a set of data and information management and user services functions, where these can include data ingest, product generation, working storage or longer term archive and preservation, preparing documentation, providing users with a catalog and access to data and products, etc., (see Appendix A below for the description of functional areas used by the CET). Your data activity can be a discrete, stand-alone organization such as a DAAC, (Distributed Active Archive Center), ESIP (Earth Science Information Partner), or SIPS (Science Investigator-led Information Processing System), or it can be part of a larger organization, such as a flight project's data system.

The CET asks you to describe your new data activity in terms of its mission and workload (i.e. describe in at least approximate terms the data it will ingest, products it will produce, services it will provide to users, when implementation will begin, when its operations will begin, and how long it will operate). From your description the CET will produce a life cycle estimate of the staff effort that will be required to implement and operate your new data activity and of the life cycle costs for staff, hardware, software, facilities, etc. Section 2.2 below discusses this in more detail.

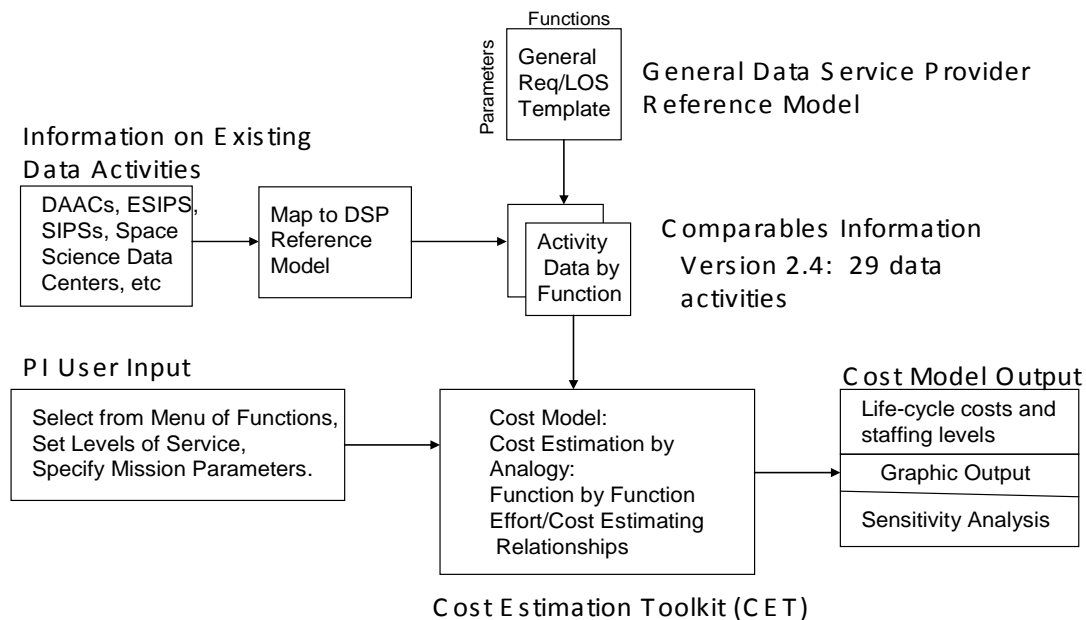
The effort and cost estimate for your data activity will be based on information about comparable existing data activities. Information about these has been compiled and assembled, structured according to the set of functional areas that constitute a general Data Service Provider (DSP) reference model as described in Appendix A below (where 'Data Service Provider' is synonymous with 'data activity'). The cost estimating approach, referred to as cost estimation by analogy, is to examine your data activity function by function, see which of the data activities are most similar to yours for each function separately, and compute a year by year effort estimate for each function over your projected data activity's life cycle. The estimate is then summed over the functional areas for the final year by year estimate. (A full description of the cost estimation techniques employed by the CET is included in the Cost Estimation Toolkit: Technical Description Document.). Figure 1 below shows how the CET uses the comparables information.

As noted above, the Data Service Provider Reference Model describes a general data service provider, or data activity, in terms of a set of functional areas (e.g. ingest, processing, archive, access and distribution, etc.). Within each functional area, a set of parameters is defined that describes it. (The parameters are described in this Users' Guide and the Technical Description Document.) The parameters are the project variables that directly influence the cost of performing one or more tasks within a functional area.

As illustrated in Figure 1, information from the existing data activities is collected from those activities, and then analyzed and mapped to the common set of functional areas and parameters according to the framework provided by the general model. The result is an internally consistent set of information describing existing ESE and other data activities.

As a CET user, you will enter an Activity Dataset, i.e. the information describing the mission, workload, etc., projected for your new data activity, and the CET will produce an estimated life cycle effort and cost estimate for it using a cost estimation by analogy approach that bases the estimate for the new data activity on information describing existing data activities that are similar, or comparable, to it. See the CET Technical Description Document for a detailed discussion of the CET's cost estimation by analogy process.

Figure 1 - CET Concepts



Since CET builds its life cycle cost estimate for your new data activity functional area by functional area, it accesses the comparables information functional area by functional area, rather than by data activity as a whole. Thus the CET will not necessarily use the same set of data activities for its estimate for all functional areas; for each functional area the CET uses information for those data activities which are the best 'comparables' for your new data activity.

The CET toolkit also includes an estimate review tool that you can use as a final step to examine the CET's life cycle cost estimate and make adjustments to it to take into account specific circumstances about your data activity that are beyond the scope of the CET (such as the extent to which you may re-use existing hardware, draw upon institutional resources, etc.) or adjust the estimated profiles (e.g. to smooth out estimated year-to-year fluctuations in staffing that track fluctuations in projected workload but would be impractical to implement).

2.2 Preparing to Use the CET: Describing Your Data Activity

Preparing to use the CET to produce a life-cycle cost estimate for your new data activity consists primarily of assembling the information needed by the CET to produce its estimate, and putting it into the form needed by the CET. This may involve some effort; and it is recognized that every

P.I. or group will have a preferred approach to describing its data activities, but of necessity the CET uses one approach to organizing information that is based on the general DSP reference model, and asks you to map your information to its model.

The term “Activity Dataset” is used as the general name for the set of information that describes the mission and planned work of your new data activity, expressed as needed by the CET. In order to be able to use the CET to produce a life-cycle cost estimate, you will assemble your information about your data activity and enter into the CET an Activity Dataset based on that information.

You can assemble more than one Activity Dataset, which allows you to use the CET to evaluate and compare effort and cost estimates for alternatives for workload or levels of service that you are considering. You can assign a name to each one and save it for future modification and use. In Tables 1 and 2, the effort and cost estimates were produced for an Activity Dataset named “Moxie” (which is included an example as Appendix C below and is provided with the CET).

The next sections describe the categories of information included in the Activity Dataset. They will refer to Appendix B, the Activity Dataset Checklist. The Activity Dataset Checklist is a description of all of the items of information that you will need in order to assemble an Activity Dataset describing your planned data activity. You will enter the Activity Dataset information by filling in the blanks on a series of forms. You can expedite this process by having the information at hand and ready to be entered; the Activity Dataset Checklist is intended to help you prepare the information you need to use the CET.

2.2.1 Data Activity Level Information

Data Activity level information needed by the CET is information that applies to your new data activity as a whole. It includes basic mission parameters such as the year implementation of your activity begins, how long it is planned that initial implementation will take, when your data activity will begin its operations, and how long it will operate (or over what period of operation you wish to produce a cost estimate). Note that operation of your activity can begin during the implementation period.

Note that the CET operates with a time resolution or temporal granularity of one year. The time resolution of the comparables information is also one year. The comparables information smoothes out variations that occur in the course of individual years, and the CET cannot estimate changes occurring within a year. The mission dates are expressed in whole years; you may have much more specific dates in your plan, and the time periods involved may not be integral numbers of years, but for the CET you need to make a reasonable approximation in whole year terms. Remember that the CET is intended to produce a reasonable estimate of life cycle costs, and not display a degree of precision that is beyond its capabilities.

The data activity level information needed by the CET also includes a set of labor rates. The CET does not estimate labor costs directly; it produces an estimate of the effort required to carry out the work of your data activity, and then applies labor rates that you provide, and a standard inflation factor, to produce its labor cost estimates. This allows you to provide the labor rates appropriate for your area and organization. You may have a more complex wage structure than the five rates shown, but only reasonable approximations are needed by the CET.

The CET asks for fully loaded labor rates. The definition of “fully loaded” will vary by organization, but the intent is to include leave, benefits, and any other costs your organization bundles into overhead on salary. In some cases this might include items that the CET estimates separately, e.g. general supplies costs and recurring facility costs, in which case you can ignore those components of the estimate.

2.2.2 Functional Area Information - Operating Functions

Functional area information can be thought of as falling into two groups, “operating” and “non-operating”, where the distinction is drawn between those areas such as ingest and processing that involve data handling and services directly, and those that support the data handling functions, such as implementation and sustaining engineering. This section will discuss operating functions, and following sections will discuss non-operating functions. Specific details on entry of information are contained in Section 3, the detailed Users’ Guide, and Appendix B, the Activity Dataset Checklist, that follow below.

The operating functions include ingest, processing, archive, and distribution, which involve data handling directly, and documentation and user support, which involve additional services to users. Not every data activity will perform the full set of operating functions (as does the hypothetical “Moxie” sample from Appendix C shown in Tables 1 and 2).

This breakdown of functions may seem strange to you if you use a different model. When you describe your own data activity - i.e., when you build your own Activity Dataset(s), you will need to map your functional model to the model used by the CET, and in so doing you can select and provide information for those operating functions that your data activity will perform and ignore the others. The allocation of workload between or across functional areas can be approximate; more important is capturing on your Activity Dataset a good approximation of the overall total. The CET product is an overall estimate, and within the overall total an over-estimation of one functional area that is compensated for by an under-estimate of another area is fine (perhaps even likely) as long as the overall result is reasonable.

A. Operating Functions - Overall Information

There are two types of information needed for the operating functions. The first type is information that describes the function as a whole, such as overall levels of service. You will be asked to choose from among various levels of service or service options those that best characterize your planned data activity. In many cases your activity might be more complex, e.g. operating at different levels of service for different products. You should make a choice that roughly characterizes your data activity as a whole. Biasing your input toward a higher level of service might produce a higher cost estimate that might represent a conservative approach, protecting the most demanding requirement.

B. Operating Functions - Workload Information

The second type is information that describes mission workload for the function, i.e. details concerning the data handled, or other activity performed by the function. Workload is the primary driver of estimated effort - it is important that you describe your new data activity’s workload as accurately as you can in your particular situation. As you will see, it is possible to use fairly general ‘ball park’ information (which you can always refine as you are able to) to produce estimates with the CET, and you can use the CET’s ability to run sensitivity tests and ‘what-if’ cases to test the sensitivity of the CET’s estimates to workload parameters with the greater degree of uncertainty.

The basic data handling workload parameters used by the CET are product types, products, and volume.

The first parameter is product type (whether referring to a type of products ingested, generated, archived, or distributed) a collective term for a specified set of parameters, ranging from raw instrument data to derived geophysical information (and supporting information such as ephemeris, calibration, or quality control parameters) that is produced over some period of time.

A product type example is NASA MODIS Level 1A 0.5 KM resolution data, or NOAA AVHRR Level 1b GAC data.

The second parameter is “product”. A product is an individual instance of a product type that is separately ingested, produced, archived, cataloged and / or distributed. For example, during a normal day, 255 MODIS Level 1A 0.5 KM products, separate instances of the MODIS Level 1A 0.5 KM resolution data product type, each containing five minutes’ worth of data, (called ‘granules’ in NASA parlance) are produced by the GSFC DAAC from Level 0 MODIS data. As another example, a single AVHRR Level 1b product (called an AVHRR Level 1b GAC ‘dataset’ in NOAA parlance) is produced by NOAA from raw AVHRR instrument data collected during the course of one Earth orbit. A user can order a single MODIS product or AVHRR GAC dataset, or a set of them that might contain data spanning a desired time period and / or covering a desired area of the Earth.

Products can be produced minute by minute (as with the MODIS example), daily, weekly, monthly, etc., and can include global coverage, small areas, etc., all depending on the product design. A product is a package of data, etc., that is separately cataloged, ingested, produced, archived, and / or distributed.

The third parameter is data volume, generally in terms of GB (gigabytes) per day (converted internally by the CET to TB (terabytes) per year).

C. Describing Ingest

The fundamental approach to describing your activity’s ingest data handling workload to the CET is to treat the flow of data into your data activity as being comprised of one or more streams of products, where each stream includes a series of products of one or more types. The workload for the stream is the sum of the workload values for each of the product types included in the stream (e.g. product instances ingested per day, volume ingested per day, summed over the product types making up the stream, and source). Since source and start and stop dates are associated with the flow of each product type, all of the types included in a stream need to have the same start and stop dates, and a common source.

As you consider describing your data activity for the CET, you may have in hand a detailed list of the data (i.e. product types and their characteristics) you know your new data activity will ingest from various sources. You can, if you wish, enter each product type as its own stream, or aggregate them into a convenient number of streams. The CET builds an aggregate workload from the streams you enter, so you do not need to enter information by individual product type unless you find it useful for your own purposes to do so.

On the other hand, if you do not have a detailed list of the data your new activity will be ingesting, or if you do not have detailed characteristics of some or all of the product types, you can create one or more general streams that approximately characterize the input workload, i.e. that include your overall estimates of product instance counts and volumes. A ‘ball park’ estimate of approximate annual averages, that you can later refine if you wish, can be used with the CET. Its estimates are generally not sensitive to small changes in its inputs.

D. Describing Processing

The approach to describing processing and operational distribution follows the approach used for ingest. For processing, the workload is treated as the generation of one or more product streams, each including one or more product types. As with ingest, you may know what suite of products your new activity will produce, and you can enter them into the CET as separate streams or aggregate them into streams that have a common start and stop date, and perhaps a common reprocessing plan (i.e. period between successive reprocessing of each type).

For each processing stream you can specify a retention period and a reprocessing interval, and the CET will estimate the effort associated with reprocessing according to your specified cycle by adding the appropriate workload (and its flow to archive).

E. Describing Distribution - Operational and By-Request

Operational distribution (as opposed to by-request distribution) is routine distribution of one or more product types to a particular destination (i.e. another data activity or user who receives the products routinely), over a specified time period. An example would be the flow of MODIS products from MODAPS to the GSFC, EDC, and NSIDC DAACs. As above, product types distributed to a common destination over a common time frame can be aggregated into a single distribution stream.

By-request distribution includes distribution made in response to user requests for products, including standing orders or subscription orders that last less than a year (standing orders for distribution of a stream of products to a user over a period greater than one year is considered operational distribution by the CET).

The CET will ask you for information characterizing your expectation for the by request workload, i.e. number of users you expect to have when your new data activity is mature, etc. You may specify a “ramp-up” to a maximum workload, e.g. a maximum number of users, products and volume to be delivered, or a “steady state” year by year constant workload (that might be appropriate for a data activity that will support a particular instrument team directly (while perhaps also performing a routine, operational distribution of its products to another data activity for more general access by users). If you specify a “ramp-up”, the CET will estimate a year-by-year ramp-up to the expected level.

F. Describing Supporting Services Functions

The CET includes two operating functions that involve services supporting users rather than handling of data and products. These are Documentation and User Support, which are assumed to be active during the data activity’s period of operations.

The Documentation functional area is intended to capture the effort associated with documentation of data and products, i.e., prepare for users the supporting information (of possibly many types including calibration parameters and algorithms, format descriptions, quality assessment, processing algorithms (and processing software), etc.) that they will need to have the fullest possible understanding of the data or product. At this time the CET uses two level of service parameters as the basis for an estimate of documentation effort.

The User Support functional area is intended to capture the effort associated with direct interaction, advice, guidance, and consultation with your data activity’s users. The CET develops a staffing estimate based on an estimate of the number of user contacts performed per year, which it develops from information you provide. A level of service parameter associated with outreach to users is also included.

2.2.3 Functional Area Information - Non-Operating Functions

The current version of the CET assumes that your data activity will be involved with all of the non-operational functional areas, including implementation, sustaining engineering, engineering support, technical coordination, and facility / infrastructure, at some level.

A. Describing Implementation

The CET produces estimates for implementation, including implementation effort, system purchase costs, COTS software license costs, and facility preparation based on the overall workload of your new data activity. The only additional inputs the CET seeks from you is a level

of service characterizing the software development to meet particular user needs that you will support. The CET will ‘size’ your data system in terms of new SLOC needed, etc., by comparing your new activity’s projected workload to the comparable activities. In estimating a system purchase cost, the CET will allow for a “Moore’s Law” effect; the comparable activity information is normalized to a base year and an estimate for your new activity is made in base year dollars and then projected to your planned implementation period.

Implementation effort, and derived labor cost, system purchase, COTS license purchase, and facility preparation are spread evenly over the implementation period you specify.

If you plan a major system redevelopment some years after your data activity becomes operational, you can consider an approach that is often used in cost estimation. That is to divide the data activity life cycle into two segments and make separate estimates for the two time periods involved. For example, if your new data activity will have a ten year mission life and you plan a major redevelopment in the sixth year, you could produce an estimate that would span the initial implementation and operation over the first five mission years, and then separately produce an estimate for the sixth through tenth years that would begin with the second round of implementation. (Note that in such a case you could continue overlapping operations during the second implementation.)

B. Describing Sustaining Engineering and Engineering Support

You will be asked only for a general level of service for each of these areas. The estimates made by the CET are based on its earlier estimates, such as the total SLOC to be sustained. These costs are spread evenly over the operating period that you specify.

C. Describing Technical Coordination

Technical coordination refers to work staff of your activity will do participating with staff from other NASA Earth science data activities in coordinating working groups, standards review processes, etc. You will be asked to identify those areas of coordination the staff of your data activity will support. The effort associated with technical coordination is in addition to the effort estimated for the basic work of your data activity.

D. Describing Miscellaneous Non-Staff Cost Items

This category is essentially a place for rounding up a few miscellaneous non-staff cost items that are not captured by the categories described above.

The first is a facility item, asking that you characterize your backup for your archive or working storage (if you plan to have any), that is used in conjunction with the earlier request for a value for the overall fraction of your data that will be backed up.

The other items are instances where you are asked to enter an annual budget that will be carried as specified into the life cycle cost estimate, for travel, data purchases (e.g. Landsat data from USGS or meteorological data from NOAA), or computing services (any form of data processing or handling service that you may procure from another source).

2.3 Interpreting and Using the CET’s Output Life Cycle Cost Estimate

This section describes the output produced by the CET, i.e. the life cycle cost estimate for a new data activity and additional related information, offers guidance on the use and interpretation of the CET output, and introduces the CET’s estimate review function.

The first and foremost point of guidance on using the CET’s life cycle estimate is to not assume that it is any way definitive or “The Answer” as to the projected cost of your new data activity. It should be regarded as one possible estimate, to be used alongside of estimates produced by other

means. The CET estimate is based on its own assumptions, the comparables information, and the input describing your data activity that you provide for it. Where some of your inputs are unavoidably uncertain but fall within probable ranges, you can produce a range of estimates by varying your inputs over their possible ranges.

The quality of the output of the CET will be no better than the quality of the input, the comparables information and your description of your new data activity. Having said that, the projected life cycle effort and costs will not be perfect no matter how good the input information is. Estimates of effort and cost three to five to ten years into the future are not without hazard - for example shifts in operating paradigms to adapt to changing user needs and/or technology enhancements may render estimates based on even recent experience, let alone past years' history, less accurate.

Section 2.3.1 describes the Life Cycle Cost Estimate tables produced by the CET, Section 2.3.2 describes the CET's Quality Report, and Section 2.3.3 provides additional guidance on interpretation and use of the CET's estimates.

2.3.1 The CET Life Cycle Cost Estimate Output, An Example

Tables 1 and 2 below are an example of the life cycle cost estimate (LCE) that the CET will produce for your data activity. They show a life cycle cost estimate for a hypothetical data activity that begins implementation in 2010, begins operations in 2012, and operates through 2017. The name of the activity dataset that contains the description of the data activity is "Moxie" (and the content of the "Moxie" Activity Dataset is contained in Appendix A below).

Table 1 - Life Cycle Cost Estimate for "Moxie" Activity Dataset

Life Cycle Cost Estimate		Activity Dataset: Moxie								Produced: 09/04/08							
Mission Start Year:		2010	Operations Start Year:			2012	Mission Complete Year:			2017	Inflation Rate		4.0%				
Estimated Staffing Level	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Total	Pct.			
Management Staff FTE	0.48	0.48	2.09	2.27	2.45	2.56	2.64	2.68	0.00	0.00	0.00	0.00	15.64	8.6%			
Administrative Support FTE	0.18	0.18	0.41	0.45	0.48	0.50	0.52	0.53	0.00	0.00	0.00	0.00	3.24	1.8%			
Technical Coordination Staff FTE	0.00	0.00	0.75	0.75	0.75	0.75	0.75	0.75	0.00	0.00	0.00	0.00	4.50	2.5%			
Development Staff FTE	8.70	8.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	17.40	9.5%			
Technical / Science Staff FTE	0.00	0.00	8.08	9.09	10.11	10.70	11.12	11.40	0.00	0.00	0.00	0.00	60.50	33.2%			
Operations Staff FTE	0.00	0.00	7.47	8.24	8.98	9.41	9.72	9.92	0.00	0.00	0.00	0.00	53.75	29.5%			
Sustaining Engineering Staff FTE	0.00	0.00	2.44	2.44	2.45	2.45	2.45	2.45	0.00	0.00	0.00	0.00	14.67	8.1%			
Engineering Support Staff FTE	0.00	0.00	2.09	2.09	2.09	2.09	2.09	2.09	0.00	0.00	0.00	0.00	12.54	6.9%			
Estimated Total FTE	9.36	9.36	23.33	25.33	27.31	28.46	29.28	29.82	0.00	0.00	0.00	0.00	182.25				
Estimated Staff Costs, K\$	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Total	Pct.			
Management Staff Cost	83	87	395	447	502	545	584	618	0	0	0	0	3,261	10.8%			
Administrative Support Staff Cost	18	18	44	50	56	61	66	69	0	0	0	0	383	1.3%			
Technical Coordination Staff Cost	0	0	138	143	149	155	161	168	0	0	0	0	915	3.0%			
Development Staff Cost	1305	1358	0	0	0	0	0	0	0	0	0	0	2,663	8.8%			
Technical / Science Staff Cost	0	0	1311	1534	1774	1953	2111	2250	0	0	0	0	10,933	36.3%			
Operations Staff Cost	0	0	889	1019	1156	1260	1353	1436	0	0	0	0	7,113	23.6%			
Sustaining Engineering Staff Cost	0	0	396	412	429	446	464	483	0	0	0	0	2,632	8.7%			
Engineering Support Staff Cost	0	0	339	353	367	381	397	413	0	0	0	0	2,249	7.5%			
Total Estimated Staff Cost	1,406	1,463	3,512	3,958	4,433	4,801	5,136	5,437	0	0	0	0	30,146	88.2%			
Other Non-Staff Costs	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Total	Pct.			
System Purchase Cost	15	11	0	0	0	0	0	0	0	0	0	0	26	0.6%			
COTS Software License Cost	58	58	14	14	14	14	14	14	0	0	0	0	200	5.0%			
Facility Preparation and Support Cost	209	209	350	380	410	427	439	447	0	0	0	0	2,871	71.2%			
System Maintenance Cost	0	0	22	22	22	22	22	22	0	0	0	0	132	3.3%			
Network / Communications Cost	4	4	7	11	15	17	19	20	0	0	0	0	97	2.4%			
General Supplies Cost	24	24	45	48	51	53	54	55	0	0	0	0	354	8.8%			
Archive Media Cost	0	0	3	2	2	2	1	1	0	0	0	0	11	0.3%			
Distribution Media Cost	0	0	3	3	2	2	2	1	0	0	0	0	13	0.3%			
Travel Cost	25	25	25	25	25	25	25	25	0	0	0	0	200	5.0%			
Training Cost	11	11	26	9	10	11	11	11	0	0	0	0	100	2.5%			
Data Purchase Cost	0	0	2	2	2	2	2	2	0	0	0	0	12	0.3%			
Computer Services Cost	0	0	3	3	3	3	3	3	0	0	0	0	18	0.4%			
Total Estimated Non-Staff Costs, K\$	346	342	500	519	556	578	592	601	0	0	0	0	4,034	11.8%			
Total Estimated Cost, K\$	1,752	1,805	4,012	4,477	4,989	5,379	5,728	6,038	0	0	0	0	34,180				

Table 1 contains three parts. The first is an overall, year by year staff effort level table with a line for each of the functional areas addressed by the CET. The life cycle total for each line is included at the right, and the percentage that each line is of the overall total. The second part contains estimated cost lines corresponding to each of the staff effort lines in the first table. In addition to the row totals, there is a row of column totals ending with a life cycle cost staffing cost total and the percentage that the staffing cost is of the total estimated cost (in this example 88.2%). The third part is a table containing year-by-year estimates of “Other Costs” - all non-staff costs, including hardware, COTS software, facility costs, etc., with row and column totals and the percentage that non-staff costs are of the total estimated cost.

As shown, an inflation rate of 4% is applied to the staff costs. The CET user is asked for labor rates as of the first year of the project, the Mission Start Year, and for an inflation rate. The rates are inflated thereafter by the specified inflation rate. Estimated non-staff costs have various cost curves built into them, and are not also inflated by the specified rate.

The very bottom row contains the year-by-year and total life cycle estimated costs.

Table 2 is a breakdown of operations and technical (including science) staff effort levels shown in Table 1. The operations and technical totals shown in Table 1 are broken down by the operational functional areas described by the hypothetical “Moxie” Activity Dataset. The other functional areas are shown separately on the first page, Table 1 above. For example, development (implementation) and sustaining engineering include only technical staff and have their own distinct lines in Table 1.

Table 2 - Technical and Operational Staffing Breakdown for “Moxie” Activity Dataset

Life Cycle Cost Estimate, Cont.														
Activity Dataset:	Moxie													
Produced:	09/04/08													
Breakdown of Management, Technical / Science, and Operations Staffing for the Operating Functional Areas														
Mission Start Year:	2010													
Operations Start Year:	2012													
Mission Complete Year:	2017													
Estimated Staffing Level	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Total	Pct.
Management Staff FTE	0.48	0.48	2.09	2.27	2.45	2.56	2.64	2.68	0.00	0.00	0.00	0.00	15.64	
Activity Level Management	0.48	0.48	1.10	1.20	1.30	1.35	1.39	1.42	0.00	0.00	0.00	0.00	8.71	55.7%
Second Level Management	0.00	0.00	0.98	1.07	1.16	1.21	1.24	1.27	0.00	0.00	0.00	0.00	6.93	44.3%
Technical / Science Staff FTE	0.00	0.00	8.08	9.09	10.11	10.70	11.12	11.40	0.00	0.00	0.00	0.00	60.50	
Ingest	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0%
Processing	0.00	0.00	6.07	6.07	6.07	6.07	6.07	6.07	0.00	0.00	0.00	0.00	36.39	60.1%
Documentation	0.00	0.00	0.91	0.91	0.91	0.91	0.91	0.91	0.00	0.00	0.00	0.00	5.46	9.0%
Archive / Working Storage	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0%
Access and Distribution	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.0%
User Support	0.00	0.00	1.11	2.12	3.13	3.72	4.15	4.42	0.00	0.00	0.00	0.00	18.65	30.8%
Operations Staff FTE	0.00	0.00	7.47	8.24	8.98	9.41	9.72	9.92	0.00	0.00	0.00	0.00	53.75	
Ingest	0.00	0.00	0.60	0.60	0.60	0.60	0.60	0.60	0.00	0.00	0.00	0.00	3.61	6.7%
Processing	0.00	0.00	1.54	1.54	1.54	1.54	1.54	1.54	0.00	0.00	0.00	0.00	9.27	17.2%
Archive / Working Storage	0.00	0.00	1.77	1.77	1.77	1.77	1.77	1.77	0.00	0.00	0.00	0.00	10.61	19.7%
Access and Distribution	0.00	0.00	2.56	2.62	2.65	2.67	2.68	2.68	0.00	0.00	0.00	0.00	15.86	29.5%
User Support	0.00	0.00	1.00	1.71	2.42	2.83	3.13	3.32	0.00	0.00	0.00	0.00	14.41	26.8%

The unit of effort is an “FTE”, a full-time equivalent, representing a full year of work by either a single person or a combination of persons adding up to a single person’s effort. For example, a level of effort of four FTE’s might be in fact four people each working full time, or eight people each working half time, or any other combination of full and part time people adding up to the equivalent of four full-time people. Note that no assumption is made concerning the hours of operation of your new data activity, only an estimate of effort needed without regard to how you choose to schedule the operation of your activity.

Operations staff includes computer operators, operations monitors, etc., i.e. staff that directly operate or oversee the operations of your data activity’s supporting systems. Technical and

science staff associated with a functional area includes scientists or engineers or other technical staff responsible for functions such as integration of science software into a production environment, quality control, integration of a new stream of data to be ingested, etc. Management effort associated with a functional area is second-line management and administrative effort, as distinct from overall management and administration of your data activity.

2.3.2 CET Quality Report

The CET provides the Quality Report as an aid in interpreting and using the life cycle cost estimate. See Figure 2 below for an example of the Quality Report produced by the CET for the Moxie activity. The Quality Report is expected to be added to and revised as users provide feedback on what sorts of information would be helpful to them in assessing the usefulness of the estimates the CET produces for them. Indeed the quality report as it currently exists is a response to a user recommendation.

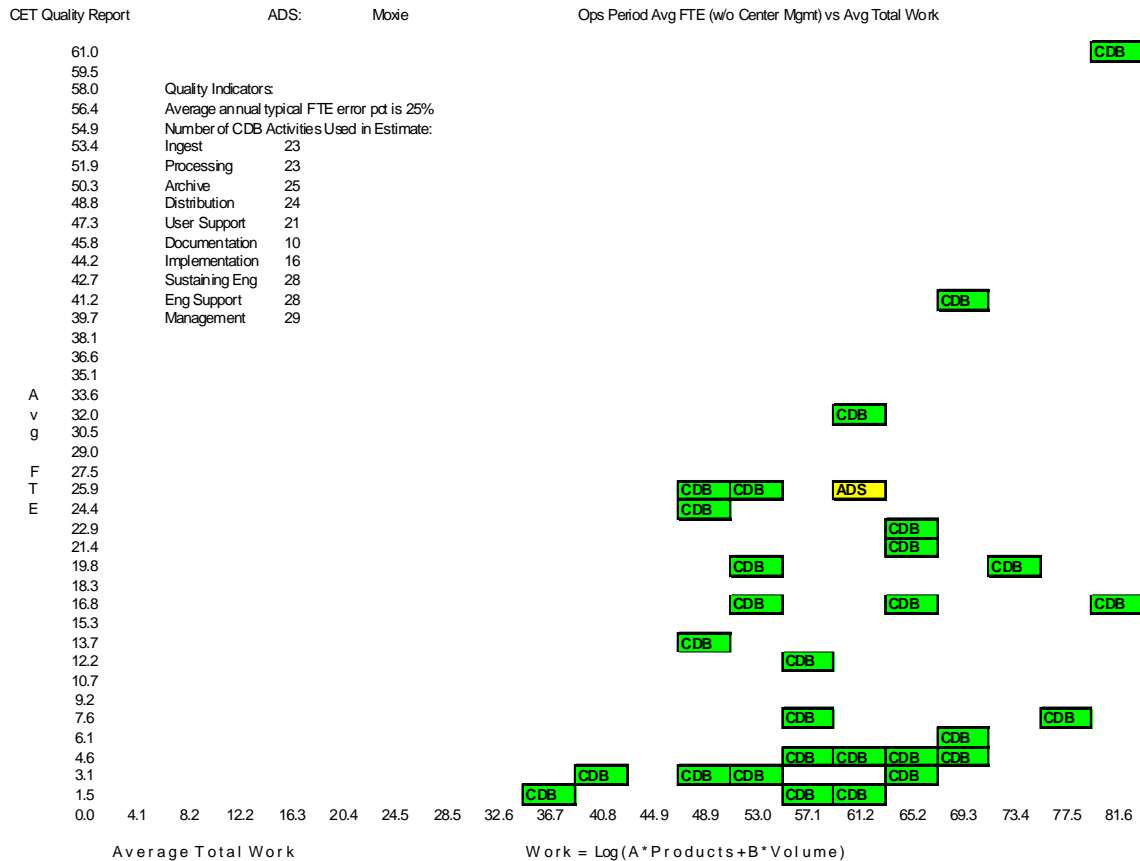
The first item included in the Quality Report is a graphical depiction of the relation between the new data activity, for which an estimate has just been produced, and the comparable data activities used by the CET. Since the estimate is based on a comparison of the new data activity with the comparable data activities, one basis for assessing the quality of the estimate is to see how like or unlike the new data activity is to the comparable activities. If it is similar to a number of comparable activities, the estimate is likely to be of higher quality than if the new activity is an outlier with respect to the comparable activities.

The similarity of the new activity to the comparable activities is indicated by plotting the new activity and the comparable activities on a scatter diagram whose axes are the annual average of the total operations period staff (less activity or Center level management) of the activities (estimated for the new activity, actual for the comparable activities) and an index of the average annual workload performed by the activities (based on information for the new activity provided by the CET user, i.e. the Activity Dataset, and actual for the comparable activities). The workload index used by the CET is a derived parameter called “Work” which is computed by the formula:

$$\text{Work} = \log(\text{base } e) [A * \text{Volume Handled} + B * \text{Products Handled}]$$

In the formula, “Volume Handled” is the average annual data volume ingested, produced, archived, and distributed, and “Products Handled” is the corresponding average annual count of products (i.e., product type instances, granules, etc.) ingested, produced, archived, and distributed.

Figure 2 - CET Quality Report, Example



Because the total staff includes staff such as implementation and sustaining engineering that are not directly related to ‘work’, and because these vary considerably across the comparable activities, and because indeed the staff directly related to work also varies for similar levels of ‘work’, there is some distribution of staff size for similar ‘work’ levels.

In the example shown in Figure 2, the Activity Dataset “Moxie” is indicated by the box labeled “ADS” while the comparable data activities are indicated in the boxes labeled “C”. (Note: as a matter of policy the identity of the comparable data activities used by the CET is kept confidential.)

The second item provided by the initial version of the Quality Report is the table named “Quality Indicators”. The first item in the table is an overall estimate of error that is based on independent testing of the CET against the existing comparable data activities. See Sections 2.3.3 and 2.3.4 below for a discussion of the interpretation and use of the estimates produced by the CET, including a discussion of the performance and probable errors. The error percentage reported, 25%, is an overall average percentage of typical error (which is absolute error so that over and under estimates do not cancel out) in the estimation of total staff effort reported by this testing. It is intended to be a conservative value that can be taken as a first, rough measure of the ability of the CET to estimate total life cycle effort for a data activity. Note that this overall error percentage is a single hard-coded value – the CET does not compute an estimated error for each estimate it makes.

The next items in the table are the number of comparable data activities used in producing the estimate for the new data activity described by the “Moxie” Activity Dataset. Since the CET operates on a functional area by functional area basis, the number of activities used is presented for each functional area and will vary from area to area. The variation will be due to two distinct factors. The first of these is the availability of usable information for each functional area for the comparable activities - an area might not apply to some, or the information about it may be incomplete for some. The second of these is the result of the CET’s screening of the data. The CET compares the comparable activity workload data to the projected workload for the new data activity, and screens out those comparable data activities that are most unlike the new data activity, in order to ensure that the estimate for the new data activity is based on similar comparable data activities. This screening is done on a functional area by functional area basis. (A second level of screening, at the parameter level, occurs within each functional area.) In principle, the greater the number of comparable activities used for a particular functional area, the better the estimate is likely to be.

2.3.3 Interpreting the CET’s Life Cycle Cost Estimate - Overall Notes

The goal of the CET is to provide a good overall year by year life cycle cost estimate for your new data activity. It builds its overall estimate functional area by functional area, and the life cycle cost estimate output worksheets (see Tables 1 and 2 above) show how the estimate breaks down by year, staff category, function, and non-staff items. The key components of the estimate are the estimated FTE levels, and as you will see the overall measure of the CET’s performance is its ability to estimate FTE. Within the overall estimate, estimates for individual functional areas will vary, and you can expect to see some areas that seem over staffed and some understaffed. This is due to differences in the approach to allocation of staff to functional areas by the comparable data activities used by the CET, as well as the inherent limitations of the CET’s estimating capability.

One overall note is that the CET life cycle cost estimate does not make any explicit allowance for re-use. So if you are estimating the cost of adding an additional data activity task to an existing base capability, you will have to discount the lines of the estimate appropriately. For example this can include system purchase and implementation effort, and operations effort, especially if the new activity represents a small overall increment to the current workload. The CET’s Estimate Reviewer function (see Section 3.13 below) is intended to help you make such adjustments to the CET’s estimate.

The accuracy of the CET has been assessed by a testing process that uses the comparable data activities as independent test cases. By this process, a data activity is removed from the comparables information (so that the results obtained for it are independent of any influence by it) and an estimate produced for it which is then compared to the actual information for that data activity, to see how well the estimated effort for implementation and operation of the activity corresponds to it’s actual effort. This process was performed for all of the comparable activities currently used by the CET, with the following results shown in Table 3 below:

Table 3 – CET Version 2.4 Performance Measurement

Number of Comparable Test Cases	29
Overall Annual Average Error Per Test Case	-.01 FTE
Overall Annual Average Pct Error	-0.1%
Average Annual Typical Test Case Error	2.03 FTE
Average Typical Percent Error	18.7%
Standard Deviation of Typical Test Case Error	1.74 FTE
Test Cases with Typical Error < 20%	21
Test Cases with Typical Error < 25%	22
Test Cases with Typical Error < 30%	22
Test Cases with Typical Error < 50%	24
Test Cases with Typical Error < 75%	25

This table provides some rough guidance for interpreting the estimates used by the CET. While the low average error suggests that the CET is generally free of systematic bias in its effort estimates, the typical (or absolute) error and its standard deviation shows that there was some variation in its results from case to case, and so offers a degree caution to the user.

Note that the typical error estimate is the annual average, computed over the whole life cycle, as if all costs are distributed evenly over the data activity life cycle. Since implementation costs are generally going to be concentrated in the early years of an activity's life cycle, the typical error in those years could be greater than average. The operating year typical errors could be less than the average, although the further out into the future the estimate goes, the more uncertainty there is that the underlying or implicit assumptions behind the estimate will continue to hold.

The grouping of test cases suggests a rough measure, for example if you feel you need an estimate of life cycle staffing that is within 20% accuracy, your chances are about 21 out of 29, or about 70% of getting that accuracy, while the odds that your estimate will prove to be within 25% are about 76%, within 30% about 76% and within 50% about 83%. (The costs corresponding to the estimated staffing are derived from the labor rates you provide.)

Non-staff costs are not included in the performance measurement.

What the foregoing suggests is that the best ways to use the CET's estimates are:

- a. To help you consider the full range of items that will contribute to the life cycle cost of your data activity;
- b. To assess what items might be the most important to estimate as well as you can, by seeing which items (especially of those of which you are least certain) your estimate is most sensitive to. You can do this by varying parameters and re-running the CET (i.e. by doing 'what if' cases) to see what effects are produced;
- c. To produce a strawman life cycle cost estimate that you do not accept as definitive but instead use as a first guess, which you modify freely where you see that it produces a result that you determine is not reasonable;

d. To produce an estimate that you can compare to or reconcile with estimates produced by other means. Where there is significant disagreement, you should use your own judgment to determine which is the better estimate.

2.3.4 Interpreting the CET's Life Cycle Cost Estimate - Itemized Notes

This section contains notes on interpretation and use of the CET's life cycle cost estimates that pertain to specific aspects of the estimate. Refer to Tables 1 and 2 above for an example of the CET output.

In some cases, such as those described below, you will want to make your own adjustments to the CET estimates. The CET's Estimate Reviewer tool (as described in Section 3.13 below) allows you to call up the CET's life cycle cost and effort estimate, and review and as needed adjust each estimate line, offering you suggested points to consider as you perform your review.

Alternatively, you can copy the CET outputs to other worksheets in a workbook of your own (you'll have to add formulas for column and row totals as you desire since the CET's totals are generated by the CET software). You can then adjust or manipulate the estimates however you wish.

In other cases, you may wish to make modifications to your activity dataset(s), re-run the CET to produce new estimates.

A. General

1. Implementation Period - note that the CET's implementation effort and cost estimates are spread evenly over the implementation period that you specify. This profile may not be realistic; you may have a sharper peak in an earlier year. Even if you are satisfied with the overall implementation estimates you may want to adjust the profile according to your actual plans, e.g. for system and/or COTS software license purchases.

2. Operations Period - effort and cost estimates for the operations period will vary with annual workload. If you have year to year upward and downward fluctuations in workload, you might find the resulting varying staffing profile to be unrealistic. In such a case you might wish to recast the estimate into a more level or gradually ramping profile.

3. Continuing Implementation - the sustaining engineering estimate line also covers a degree of continuing implementation after the implementation period. It is spread evenly over the operations period, but if you know you will have implementation effort in an out year you will want to adjust the profile accordingly. Also, the total effort in the sustaining engineering line will not be sufficient to account for a major system refresh or redevelopment, so if you want to plan for that you will have to add it (for example by adding in a substantial fraction of the initial implementation effort). If you plan a major system redevelopment some years after your data activity becomes operational, you can consider an approach that is often used in cost estimation. That is to divide the data activity life cycle into two segments and make separate estimates for the two time periods involved. For example, if your new data activity will have a ten year mission life and you plan a major redevelopment in the sixth year, you could produce an estimate that would span the initial implementation and operation over the first five mission years, and then separately produce an estimate for the sixth through tenth years that would begin with the second round of implementation. (Note that in such a case you could continue overlapping operations during the second implementation.)

B. Life Cycle Cost Estimate, Page 1 (Table 1 above)

1. The Management staff FTE and Administrative Support FTE lines (see the sample CET output, Table 1) are estimated by the CET based on the average experience of the comparable data

activities used by the CET, i.e. the average ratio of management staff to ‘working’ staff. You may want to adjust the management staffing to fit your own organization’s management approach.

2. Development Staff FTE - see the Implementation period note above.

3. Technical / Science Staff FTE and Operations Staff FTE - these are roll-ups of estimates the CET makes for the functional areas (see C below, Life Cycle Cost Estimate, Page 2). The roll-ups in this table are going to be better estimates than the estimates for individual functional areas (see below) because of the degree of uncertainty of the allocation of staff between functions, and the inherent variability of the comparables information in each functional area. As a result you should obtain an overall reasonable result, but individual lines might seem too large or too small. There is also uncertainty in the allocation of effort between operations staff and technical / science staff - you may, for example, plan to have your entire data activity staffed by technical / science staff. The only distinction that the CET draws is based on the labor rate you assign - you can assign a single labor rate for both categories.

4. Sustaining Engineering Staff FTE - see the note above on continuing implementation.

5. Engineering Support Staff FTE - the definition of this category is quite broad (see Sections A.2.10 and B.2.3 below) and you may wish to adjust the CET’s estimate if you see that it includes items that you may not need (e.g. that may be provided by your organization).

6. Estimated Staff Costs - these are driven by the labor rates that you provide, and the annual inflation rate that appears in the upper right hand corner of Table 1. Typical FTE error estimates can be recast as typical cost error estimates by application of an average overall labor rate.

7. System Purchase Cost and COTS Software License Cost - see the note on Implementation Period above.

8. Facility Preparation and Support Costs – the CET estimates of these costs do not include the costs of construction associated with a new facility or expanding a facility; if you plan construction you would have to add the associated costs. The intent is to include all items such as rent, utilities, support services such as security, furniture, etc. The estimation of recurring facility cost is based on a simple assumption of \$15K per FTE per year, an intermediate value from a set of actual values obtained from NASA and NOAA activities. It is possible that your organization’s overhead on labor covers some or all ongoing facility costs, in which case you would want to adjust the estimate accordingly, by deleting the \$15K per FTE per year cost from your estimate or reducing your labor rates by \$15K, or a different amount if your own institution has a different cost level.

9. System Maintenance Cost - this estimate is based on a fixed fraction of system purchase price and includes operating system upgrades, and in practice may well not be spread evenly over the operating period as shown.

10. Network / Communications Cost - these are estimated by the CET based on the total projected network volume (for ingest and distribution) for your activity and a current commercial rate; if you are able to obtain a better rate, your costs will be less. The rate assumed by the CET is \$3.6K per year for a full T1 capacity, which the CET assumed can be operated at 70% efficiency. The CET projects a 25% per year average reduction in the cost of network capacity, consistent with some published projections.

11. General Supplies Cost – these are estimated by the CET based on its staff estimate for your data activity; your organization may cover general supplies costs in its overhead on labor, in which case you will want to delete this cost.

12. Archive Media Cost and Distribution Media Cost - these are estimated by the CET based on volumes archived and distributed and the values for media unit capacity and price that you

accepted or provided. They do not allow for a change of media (such as adoption of blue laser DVDs when these become available) during the life of your activity; you would have to adjust the estimate if you wish to plan for such a change. The CET assumes a 15% per year reduction in the per GB cost of these media; you might project a greater reduction based on your particular choice of media.

13. Travel Cost, Data Purchase Cost, Computer Services Cost - these are as you specified; but since they are constant across the life of your activity you might want to change their profile according to your plans.

14. Training Costs – these are estimated by the CET based on its estimate of staffing for your data activity; your organization may cover training costs in its overhead on labor, in which case you will want to delete this cost.

C. Life Cycle Cost Estimate, Page 2 (Table 2 above)

1. Management Staff FTE - shows the estimated breakdown of management between overall activity management and second level management associated with functional areas. This may be helpful if you wish to adjust the estimate to be consistent with your organization's practices.

2. Functional Area Breakdown of Technical / Science Staff and Operations Staff (see also the note on these above, note 3 in part B); this breakdown of the CET estimate will enable you to spot any components of the estimate that you may judge to be way off, either because of a question of allocation between functional areas that might not affect the overall result, or a question of allocation between technical / science staff and operations staff, which while not affecting the total estimated effort would affect the total estimated cost because of differences in labor rates.

In some cases, your activity's operating policies may affect staffing. For example, the estimated user support staffing level and its split between operations and technical staff is based on an estimate for the number of user contacts per year and the scope of the user community (from general public to a small science team), but your activity may have a user support policy that will influence your staffing, and you will be able to adjust the estimate accordingly.

Beyond any question of misallocation, you may spot components of the estimate that you may judge to be just wrong, too high or too low, and you can adjust those components of the estimate accordingly. As noted above, the CET's ability to estimate individual functional areas is limited by the variability of the comparables information.

D. Notes on Re-Use

One of the considerations that will arise is the extent to which you might re-use existing software. The following are extracts from "Rapid Development," Steve McConnell, Microsoft Press, 1996, Chapter 33, Reuse (in Part III – Best Practices) that may be helpful to you.

"Create small, sharp components. You might be better off focusing your reuse work on the creation of small, sharp, specialized components rather than on large, bulky, general components. ... Data from the NASA Software Engineering Lab suggests that if a component needs to be modified by more than 25 percent, it's as cost-effective to develop a new one-off component as it is to reuse an old one (NASA 1990)."

"You can reuse opportunistically when you discover that an existing system has something in common with a system that you are about to build. You then save time by salvaging pieces of the existing system and using them in the new system. ... The bottom line on Opportunistic Reuse varies considerably, depending on how big the reuse opportunity is. Small amounts of Opportunistic Reuse will produce small savings. Large amounts of reuse could probably produce effort savings in the range of from 20 to 25 percent over the life of an entire project, assuming

that large amounts of code and design can be reused and that there is some staff continuity between the new program and the one being salvaged.”

“Keys to Success in Using Reuse:

1. Take advantage of personnel continuity between old and new programs.
2. Don’t overestimate your savings.”

3.0 Cost Estimation Toolkit Users' Guide

This section is the Users' Guide proper. It will lead you through the process of building an Activity Dataset, Modifying an existing Activity Dataset, and producing a life cycle cost estimate.

3.1 Setting Up the CET

This section describes the CET Package and how to get the CET installed and up and running on your PC or Macintosh computer.

3.1.1 What the CET Package Consists Of

The Cost Estimation Toolkit consists of one Excel workbook, "CET V2p4.xls" (for Cost Estimation Toolkit Version 2.4). The CET is an Excel VBA (Visual Basic for Applications) application. You must have Excel 2002, 2003 or 2007 to run the CET, or a corresponding Mac version. The CET is best run with Windows XP, but it will run with Windows 2000 or newer, and Apple OS 8 or newer. The CET has not yet been tested with Microsoft's new Vista operating system. The CET should be run with a display resolution of 1024x768. The CET detects which type of platform it is running on (PC or Macintosh) and adjusts its display handling accordingly. The CET Estimator running time varies with the capability of the platform it runs on, requiring about 5 seconds on a 3GHz PC platform.

The CET will communicate with you by means of "message boxes" or "user forms" which will be called simply "boxes" in this guide. In some cases you will be presented with boxes with blanks to be filled in. Where you must pick from a set of possible responses, the CET will display your choices. The CET will check your entries for errors and inconsistencies. It does so when you have completed all of the entries in a box (since only then can the CET check for inconsistencies between otherwise valid entries). The PC versions of the boxes are shown in this Users' Guide; the Macintosh boxes are different in style only.

The CET uses various Excel worksheets to contain its inputs and outputs and display backgrounds. When a "box" is displayed, a worksheet may be displayed in the background, and you can move the box about to look at the worksheet in the background (e.g. an activity dataset which you may modify with responses to the "box").

The life cycle cost estimate output generated by the CET is a pair of worksheets displaying the life cycle cost estimate for a data activity (see Tables 1 and 2 above). When each output worksheet is displayed, a box will offer you the option of printing it. If you are using a PC and Excel version 2002 or newer, you will be able to scroll around the worksheet while the box is displayed. Macintosh versions of Excel do not permit this.

When the Excel workbook is open, but the CET program is not running (either when the workbook is first opened before you run the CET or after you have exited from the CET), you will have access to some of the worksheets in the workbook.

The CET Workbook contains two types of worksheet: 1) worksheets that are internal to the CET, and 2) worksheets to hold Activity Datasets that you have defined. Some of the internal worksheets that are used for diagnostic information or intermediate results are hidden. The visible internal worksheets are password protected because the CET relies on formatting within the worksheets. The worksheets for Activity Datasets that you have entered, bear the names that you gave them (an example of an Activity Dataset, "Moxie" that was used to produce the example life cycle cost estimate output described in Section 2.3 and that is described in Appendix A, is provided with the package; you may discard it at your pleasure). Section 3.15, Accessing and Using the CET Worksheets, includes a table of the CET internal worksheets and notes / cautions on their use.

3.1.2 How to Install and Start the CET

Copy the CET workbook to a convenient folder. A good precaution is to make a working copy of the CET, thus preserving your option to revert to the original to produce a new working copy if you, for example, inadvertently delete a key CET internal worksheet (see below).

At this time, you have access to some of the worksheets contained in the CET workbook, and can inspect those that are visible and modify certain of them. See Section 3.15 for a brief discussion of the CET worksheets.

Macintosh Notes:

1. Screen Set-up: Set the Macintosh screen resolution to 1024x768 (not “stretched”) and with Excel’s “Full Screen” view selected on the Macintosh. Before running the CET, open the workbook and select the sheet named “Output_M”, and if necessary adjust the Excel window to display the full height and width of the worksheet.
2. Macintosh users should ensure that Excel has at least 16MB of memory allocated to it. This note applies to operating systems earlier than OS X.
3. In some cases, on your first attempt to open the CET workbook, when you click on the icon for the CET it may not open the workbook. If that happens, start up Excel itself, and then from within Excel open the workbook. Once you have opened the CET workbook from Excel and saved it and closed Excel, you will be able to open the CET workbook by clicking on its icon without first having to start Excel.

Activate the CET by opening the CET workbook and clicking on “Enable Macros” when Excel 2002 or 2003 displays a warning. With Excel 2007, the workbook will open in ‘Compatibility Mode’ with a security warning saying that macros have been disabled, accompanied by an options box. Click on the options box, select “Enable this Content” and click “OK”.

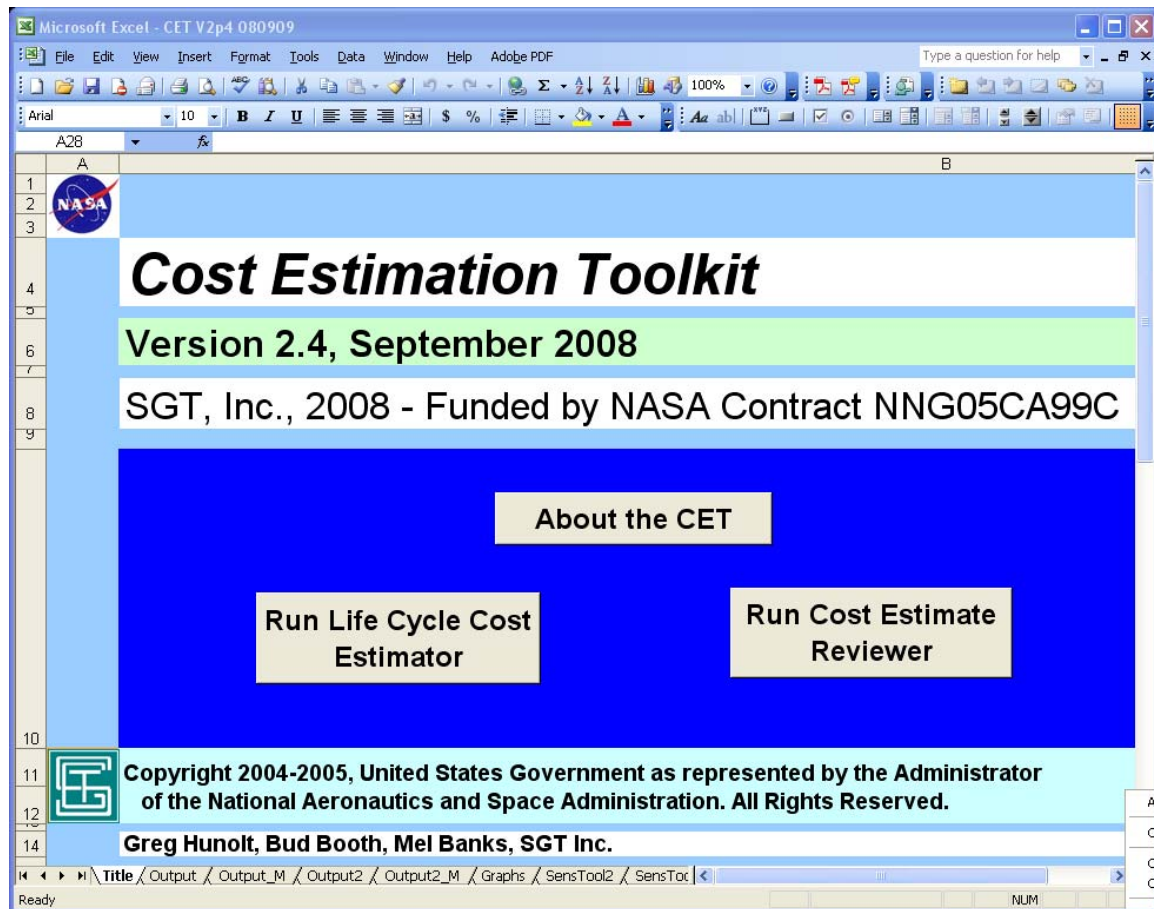
When the workbook is opened, the CET title worksheet should be displayed, as shown in Figure 3 below. (If it is not, click on the tab called “Title”.) The Title sheet contains three buttons, one called “About the CET” that provides some background information including points of contact for help and information, and two “Run” buttons, each of which when clicked launches a CET tool:

1. Button “Run Life Cycle Cost Estimator” starts the CET Estimator tool (see next Section 3.2. next below).
2. Button “Run Cost Estimate Reviewer” starts the CET Reviewer tool (see Section 3.13 below).

Note that you can always access this page and the “run” buttons by clicking on the “Title” tab.

One final note: the CET’s internal operations involve copying of worksheets. For reasons best known only to Excel, on rare occasions the worksheet copy operation fails and an error results. If this should happen, you will see an error message and you will be able to make a controlled exit from the CET program. You must save and close the Excel workbook, and re-open the workbook, thus restarting Excel, to clear the error condition – if you re-run the CET without closing and re-opening Excel, the error will recur.

Figure 3 - CET Title Worksheet



3.2 Beginning: Selecting an Activity Dataset

The first screen you will see when you have started the CET with the title screen in background is an opening box in the foreground titled "Select an Activity Dataset" as shown in Figure 4 below:

Figure 4 - Select Activity Dataset Box

Build a New Activity Dataset or Select an Existing Activity Dataset

An Activity Dataset is a package of information describing your Data Activity.
You may have several Activity Datasets, each saved under its own name,
i.e., alternate descriptions of your Data Activity.

☒ **Define New Activity Dataset**

Enter the name for a new Activity Data Set:
(Case sensitive)

☐ **Work with an Existing Activity Dataset**

Select an Existing Activity Dataset from list:

Moxie

Proceed **Exit from CET**

The box offers you two options, either to define a new Activity Dataset or to work with an existing (i.e., previously saved as a worksheet) Activity Dataset (e.g., in Figure 4 above, the sample Activity Dataset named “Moxie” provided with the CET). Working with an existing Activity Dataset can include making changes to it, and either saving new and old versions or replacing the old version with a new version.

A good practice, whether you are working on a new Activity Dataset or modifying an existing Activity Dataset, is to periodically save the workbook to save your work as you go along.

If you want to define a new Activity Dataset, select that option and enter the name for the new Activity Dataset in the blank, and then click “Proceed”. Note that the name is case sensitive, i.e. Fred and fred are different names. Go to Section 3.3 for the next steps.

If you want to work with an existing Activity Dataset, select that option and choose the name of the existing Activity Dataset from the list of names in the box (e.g. as “Moxie” is highlighted in the example), and then click “Proceed”. Go to Section 3.4 for the next steps.

If you want to leave the CET, click on “Exit from CET”.

3.3 Define New Activity Dataset

This section describes how you may define a new Activity Dataset. You will be asked for information describing the new activity as a whole, and then for information for each of the operational functional areas that your activity will include, and then for the non-operational functional areas. The description of entry of functional area information will step through all of the operational functional areas, though you would only be asked to enter information for those operational functional areas that you select, and the non-operational functional areas.

You will notice that some of the information items will be marked as “Advisory – No Effect on Estimate”. These are items that may have significant effort associated with them, but for which the CET has no basis for making an estimate of that effort. An example is “Science Software Service Options” which allows you to characterize the types of software (e.g. research or operational product generation, data integration, data mining) that your activity will accept and run for users. What you select can have a significant effect on your staffing, but the CET has no basis for making an estimate of that effect. Although your entry for “Science Software Service

Option” will not have any effect on the estimate, it serves as a reminder for you, and you can use the CET Reviewer tool (described in Section 3.10 below) to adjust the CET’s estimate to be commensurate with your plans. The Reviewer tool will remind you of this.

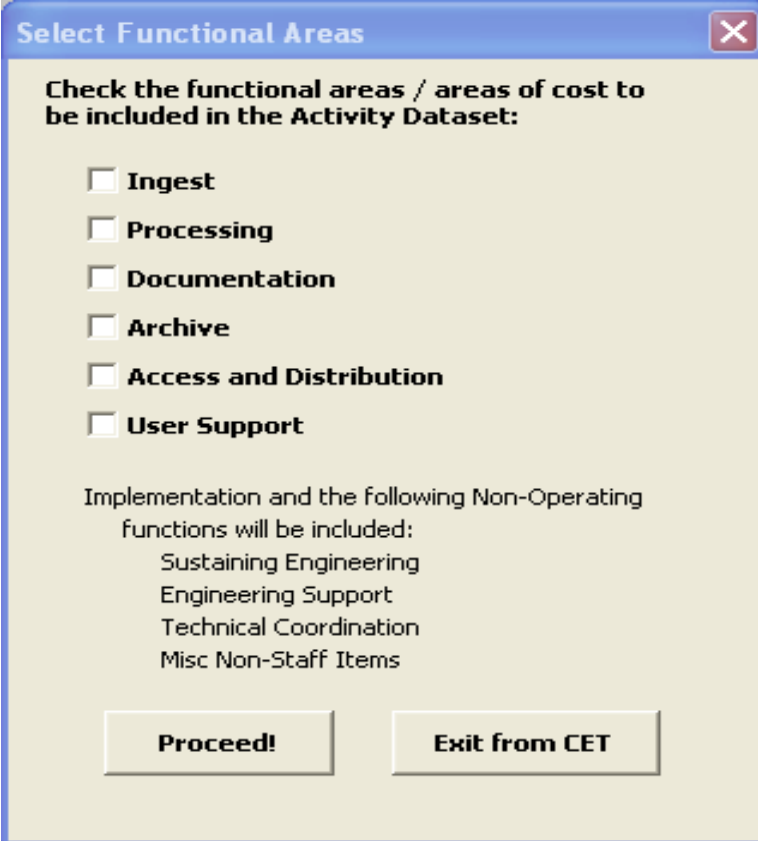
All of the numeric entries are integers or decimal (floating point) numbers entered with no commas (commas make Excel think you intend multiple entries, with results that tax the sense of humor).

All of the functional areas are described in Appendix A below.

3.3.1 New Site/Activity Level Information

The next thing you will see after clicking “Proceed” is a box labeled “Select Functional Areas” containing a check list of all of the possible operating functional areas / areas of cost that you can include in the new activity dataset. See Figure 5 below for an example, and Table B1 in Appendix B below for additional information:

Figure 5 - Select Functional Areas Box



Select Functional Areas

Check the functional areas / areas of cost to be included in the Activity Dataset:

- ☐ **Ingest**
- ☐ **Processing**
- ☐ **Documentation**
- ☐ **Archive**
- ☐ **Access and Distribution**
- ☐ **User Support**

Implementation and the following Non-Operating functions will be included:

- Sustaining Engineering
- Engineering Support
- Technical Coordination
- Misc Non-Staff Items

Proceed! **Exit from CET**

Select the appropriate functions for your data activity, and click “Proceed”. You’ll be shown a list of your choices and asked to confirm them or, if you wish, make corrections or changes. When you’ve confirmed the entries, the CET proceeds to the next step. Selecting “Exit from CET” to exit the CET if you wish. Note that the non-operating functions, including Implementation, will be included automatically.

After you’ve confirmed your function choices you will see a box titled “Activity Dataset - Data Activity Level Information”, as shown in Figure 6 below (with sample entries - when you are

entering new information the fields will be blank) and Tables B2 and B3 in Appendix B for additional information.

Figure 6 - Data Activity Information Box

Activity Dataset - Data Activity Level Information

Enter the information below for your Data Activity:

Mission/Project Start Calendar Year: 2010
(Four digits, e.g. 2004)

Period of Initial Implementation, Years: 2

Calendar Year of Operations Start: 2012

Duration of Operations, Years: 6

Enter Labor Rates

Fully Loaded Labor Rates. K\$/Yr, as of Mission/Project Start CY

Management Staff Labor Rate: 175

Admin Support Staff Labor Rate: 100

Development / Engineering Staff Labor Rate: 150

Technical / Science Staff Labor Rate: 150

Operations Staff Labor Rate: 110

Inflation Rate to Apply after Start Year: 4 %, e.g. 3.0

Accept **Exit from CET**

Fill in the blanks with information describing your data activity (which may describe your site, if your site will perform one data activity, or a new data activity to be added to the overall work of your site), beginning with whole-year life-cycle checkpoints that apply to your activity as a whole, and fully loaded annual labor rates for several labor categories and an inflation rate that will be used with your labor rates. (The labor rates and inflation rate are used in the final step of the cost estimation process to compute labor costs from estimated FTE's). ("Fully loaded" labor rates include all overheads your activity or host institution applies to labor cost. See the Activity Checklist, Appendix B, for information on these items.)

Once you have entered the information (don't hit carriage return after entering a value; use the tab or mouse to move to the next blank), click on "Accept", or you may click on "Exit from CET" to exit from the CET. When you click on "Accept", your entries will be checked, and you will be

asked to review, correct them as needed, and confirm them. After you've confirmed them, the CET proceeds to entry of information for the selected functional areas you selected previously.

3.3.2 New Ingest Information

As discussed in Section 2.2 above, ingest information is entered as descriptions of one or more 'ingest streams', where an ingest stream consists of either a flow of instances of a single data or product type or a flow containing instances of a group of data or product types that have common characteristics (e.g. start and stop year, source, mode of delivery) allowing them to be aggregated into a single ingest stream for the purposes of the CET. This is intended to save you from having to enter each data or product type individually where possible.

You will see the box shown below as Figure 7 for entry of information describing a single ingest stream (with sample entries; when you are entering new information, the fields will be blank). See also Table B4 in Appendix B below for additional information.

Figure 7 - Ingest Stream Information Box

Activity Dataset - Ingest Stream Information

Enter the following information for a single Ingest Stream - single data type or product type, or aggregate of types:

Name of Ingest Stream	<input type="text" value="Stream_A"/>	Product types with common characteristics and source can be aggregated (case sensitive)
Number of Product Types in Stream	<input type="text" value="2"/>	Enter 1 if no aggregation, i.e. stream is a single product type
Calendar Year - Ingest Start, e.g. 2009	<input type="text" value="0"/>	Enter 0 if Same as Operations Start, or 4 digit year
Calendar Year - Ingest End, e.g. 2014	<input type="text" value="0"/>	Enter 0 if Same as Operations End, or 4 digit year
Source	<input type="text" value="Site_A"/>	Name of source (case sensitive)
How Received	<input type="text" value="1"/>	1 - Electronic, 2 - Media
Level of Service	<input type="text" value="1"/>	1 - Time critical ingest with immediate verification of data integrity 2 - Routine Ingest and verification w/o tight time constraints 3 - Ad-Hoc or intermittent ingest with verification 4 - Ad hoc, non-operational ingest
Retention Period, Years	<input type="text" value="0"/>	Enter 0 if Indefinite
Automation LOS	<input type="text" value="2"/>	1 - Routine or frequent manual action or intervention in ingest process 2 - Occasional manual action or intervention in ingest process 3 - Rare manual action or intervention in ingest process. e.g. "lights out"
Products of Stream Ingested per Day	<input type="text" value="2"/>	Sum over product types included in stream
Total Volume of Stream, GB per Day	<input type="text" value="1.5"/>	Sum over product types included in stream

Once you've entered information for one stream, the CET will ask if you wish to enter information for another, until the information is complete.

As noted above, the box indicates, a single ingest stream can include a single product or data type, or it can include an aggregate of product or data types when these have all other properties in common (except for number of products/day and volume/day which are summed over the data/product types being aggregated).

Fill in the blanks for an ingest stream, making changes as you go as needed, and then click on "Accept this Stream". Note that the Name of Ingest Stream is case sensitive. When you click on

“Accept this Stream” you entries will be checked for errors. Information about the individual items is contained in the Activity Dataset Checklist (Appendix B below).

When you click “Accept this Stream”, a box titled “Add/Modify Ingest Information” is provided that displays the information for the ingest stream you have just entered. You can confirm that the information is OK by clicking “Yes”, or that a correction is needed by clicking “No”. If you click “No”, the ingest stream information box is brought back with your previous entries shown. You can make any needed corrections and again click “Accept this Stream”.

After the ingest stream is accepted, you are asked if information for another ingest stream is to be entered. Click “Yes”, to repeat the process described above.

If, after entering one or more ingest streams, you enter “No”, indicating that entry of ingest information is complete, the CET proceeds to the next selected functional area.

You may also back out of the entry of an ingest stream, or correction/modification of an ingest stream, by clicking on “Back Out of this Stream”. You will go back to either being asked if you want to enter a new ingest stream, or delete or modify an ingest stream, as appropriate.

3.3.3 New Processing Information

Processing information (see also Section 2.2 above) is entered as some overall items of information describing your site’s processing operation, followed by descriptions of one or more ‘processing streams’, where a processing stream (as was the case above for an ingest stream) consists of either a production flow of instances of a single product type or a production flow containing instances of a group of product types that have common characteristics (e.g. start and stop years) allowing them to be aggregated into a single processing stream for the purposes of the CET. This is intended to save you from having to enter each data or product type individually where possible. If you do not have information on specific product types you can enter overall estimates under a generic name.

You will first see a box titled: “Activity Dataset - Data Activity Level Processing Information” as shown in Figure 8 below (with sample entries). See also Table B5, Part A in Appendix B below for additional information.

Figure 8 - Data Activity Processing Information Box

Activity Dataset - Data Activity Level Processing Information

Science Software Service Options: (Advisory - No Effect on Estimate)

1 - Data Activity accepts operational product, research product, data integration / data mining, etc., software.
2 - Data Activity accepts research product, data integration / data mining, etc., software.
3 - Data Activity accepts both operational product and research product software.
4 - Data Activity accepts only research product software.
5 - Data Activity accepts only operational product software.

Calibration Validation LOS: Enter 1 if Data Activity performs special efforts in calibration-validation, otherwise 0.

Accept

The box asks you for entry of two processing parameters that are independent of individual processing streams.

The first parameter is Science Software Service Options, and you are offered a choice of entries. Note that this parameter is “advisory” – what you enter will not have any effect on the estimate,

but you may wish to use the Reviewer tool to adjust the estimate of you plan a significant effort for options 2 through 5.

The second parameter is Calibration-Validation LOS, which you can use to indicate that the Data Activity has a special responsibility for calibration and validation, perhaps including cross-calibrating different data types. Make the choices appropriate for your activity and click “Accept”. Your entries will be checked for errors, and you will be asked to review, correct if necessary, or confirm your entries.

When you’ve confirmed your entries, a box asks if you wish to enter information describing a processing stream.

Your entry of processing stream information proceeds as described above for ingest streams. You will be presented with a box for entry of information for an individual processing stream, as shown in Figure 9 below (with sample entries; when you are entering new information, the fields will be blank). See also Table B5, Part B in Appendix B below for additional information.

Figure 9 - Processing Stream Information Box

Activity Dataset - Processing Stream Information

Enter the following information for a single Processing Stream - single product, or aggregate of product types:

Name of Processing Stream	Level_2	Product types with common characteristics can be aggregated (case sensitive)
Number of Product Types in Stream	12	1 - no aggregation, i.e. if stream contains a single product type N - Number of product types included in this stream
Calendar Year - Production Start	0	0 - Same as Operations Start, or 4 digit year, e.g. 2009
Calendar Year - Production End	0	0 - Same as Operations End, or 4 digit year, e.g. 2014
Science Algorithm Software Source	1	1 - In-House, 2 - From an Outside Source, e.g. Instrument Team
Product Science Quality Assurance	1	1 - In-House Function, 2 - Performed by outside group
Processing Automation LOS	2	1 - Routine or frequent manual action / intervention in production process 2 - Occasional manual action or intervention in production process 3 - Rare manual action or intervention in ingest process. e.g. "lights out"
Production Mode	1	1- Operational, 2 - Non-Operational
Operational Production Mode, if applicable	1	1 - Scheduled/Routine, 2 - On-Demand, 0 - Not Applicable
Operational Processing LOS, if applicable	2	1 - 2 days or less, 2 - 7 days, 3 - 30 days, 0 - not applicable
Non-Operational Processing LOS, if app.	0	1 - Specific targets, 2 - General goals, 3 - No goals, 0 - N/A
Product Retention Period (each version)	0	0 - Indefinite, N - Retention period in years
Products in this Stream generated per day	24	Enter number (can be decimal fraction, e.g. 14.5/day for 1 per orbit in some cases, or .033 per day for 1 per month)
Volume of this Stream generated, GB/day	3	Number, can be decimal fraction, e.g. 2.34
Reprocessing LOS for this Stream	4	1 - Reprocessing is scheduled, retain previous version; 2 - Reprocessing is scheduled, replace previous version; 3 - Reprocessing is unscheduled, done on time available basis; 4 - No Reprocessing.
Reprocessing Interval for this Stream	0	N - Nominal Scheduled Reprocessing Interval in whole Years, e.g. 3; 0 - Not applicable - no scheduled reprocessing (for Reproc LOS 3 or 4).

Accept this Processing Stream **Back Out of this Processing Stream**

As noted above, a single “processing stream” can include a single product type, or it can include an aggregate of product types when these have all other properties in common (except for number

of products/day and volume/day which are summed over the data/product types being aggregated). If you do not have information on specific product types you can enter overall estimates under a generic name.

Fill in the blanks, making changes as needed, and click on “Accept this Processing Stream”. (Note that the Name of Processing Stream is case sensitive.) When you do so, your entries will be checked for errors.

When you click “Accept this Processing Stream”, a box titled “Add/Modify Processing Information” is provided that displays the information for the processing stream you have just entered. You can confirm that the information is OK by clicking “Yes”, or indicate that a correction is needed by clicking “No”. If you click “No”, the processing stream information box is brought back with your previous entries shown. You can make any needed corrections and again click “Accept this Stream”.

After the product stream is accepted, you are asked if you wish to enter information for another processing stream. Click “Yes”, to repeat the process described above.

If, after entering one or more processing streams, you enter “No”, indicating that entry of processing information is complete, the CET proceeds to the next selected functional area.

You may also back out of the entry of a processing stream, or modification of a processing stream, by clicking on “Back Out of this Processing Stream”. You will go back to either being asked if you want to enter a new processing stream, or delete or modify a processing stream, as appropriate.

3.3.4 New Documentation Information

The information you need to provide for the documentation functional area are two level of service parameters that apply to your activity as a whole. The box you will see is shown in Figure 10 below, with sample entries. See also Table B6 in Appendix B below for additional information.

Figure 10 - Documentation Information Box

The screenshot shows a dialog box titled "Activity Dataset - Documentation Information". It has a blue header bar with a close button (X) in the top right corner. The main area is light yellow. There are two sections, each with a label on the left and a dropdown menu. The first section is "Documentation Level of Service" with a dropdown showing "1". To the right of the dropdown is a list of three options: "1 - Data and product holdings documented to standard for long term archiving;", "2 - Data and product holdings documented to standard for current use;", and "3 - Documentation is only as received from the data provider.". The second section is "User Comment Level of Service" with a dropdown showing "2". To the right of the dropdown is a list of three options: "1 - Data and product documentation routinely updated with user comments;", "2 - Data and product documentation occasionally updated with user comment;", and "3 - Data and product documentation rarely updated with user comment.". At the bottom center is a button labeled "Accept".

Make the choices that provide the best overall characterization of your activity. Make your selections and click “Accept”. You’ll be asked to review, correct if necessary, and confirm your selections, and then to proceed to the next functional area.

3.3.5 New Archive Information

You will see an information entry box for the Archive functional area with seven blanks to be filled in, as shown in Figure 11 below (with sample entries). See also Table B7 in Appendix B below for additional information

Figure 11 - Archive Information Box

Activity Dataset - Archive Information

Archive Purpose:
2

1 - Indefinite Archive
2 - Long Term (Multi-Year) Storage
3 - Temporary Short-Term Working Storage

Archive Media Unit Capacity, GB:
50

Approximate capacity of each tape or disk, GB.
Accept the default shown, or enter another value.

Archive Media Unit Price, \$:
100

Accept the default shown, or enter another value,
in dollars, e.g. 25 for \$25.00 or 75.50 for \$75.50.

Archive Backup Fraction, Pct:
75

E.g., for 75% enter 75

Archive Media Re-Use Plan:
4

0 - Do not Re-Use Archive Media;
1 - Always Re-Use Archive Media when deleting, e.g. when
reprocessing and replacing old products with new version;
N - Do not Re-Use Archive Media after N years, at least 2.

Archive Monitoring LOS:
2

1 - 10% per year random screening
2 - 5% per year random screening
3 - 1% per year random screening
0 - No random screening

Archive Entry / Exit Screening LOS:
(Advisory - No Effect on Estimate)
1

1 - Entry and Exit Screening
2 - Entry Screening Only
0 - No Entry or Exit Screening

Pre-Existing Archive Information (Total Contents of Archive as of ADS Start, Archive Purpose must be 1 or 2):

	Product Types	Products	Volume, TB
In-Active (No Ingest or Production):	0	0	0
Active (Ongoing Ingest or Production):	0	0	0

Accept Archive Information

You can find additional information on the items in the Activity Dataset Checklist (Appendix B below). Note that the first item asks you to define the overall purpose of your activity's "archive", ranging from indefinite storage and preservation of data and product holdings to temporary working storage - not a true 'archive' in the literal sense of the term, perhaps, but necessary for the purpose of cost estimation. Even an activity that only produces and distributes products will require some amount of working storage capacity.

Note that "Archive Entry/Exit Screening LOS" is an advisory parameter. You entry will not affect the estimate, but you can adjust the estimate for this as needed later using the Reviewer.

Note that you can "initialize" your activity's archive with data holdings that you may be inheriting from an earlier data activity or otherwise plan to start out with. These can be either inactive data, i.e. data that is no longer being collected and ingested, or active data that will continue to be ingested by your activity. In either case the aggregate total number of product types, products (i.e. the sum of product instances across all of the types), and volume (you must enter all three parameters for active and/or inactive data). This will be helpful if your new activity is a major re-implementation / modernization of an existing activity. Note that if you enter pre-existing archive information, the Archive Purpose must be 1 or 2.

Fill in the blanks, and click on “Accept Archive Information”. Your entries will be checked for errors. You’ll be asked to review, correct if necessary, and conform your entries, and then to move on to the next functional area.

3.3.6 New Access and Distribution Information

Access and distribution information consists of three parts. The first part describes the overall access and distribution service to be provided by your data activity, which can include operational distribution, by request distribution, or a mix of both. Distribution by media or network (including both traditional FTP-like modes and system-to-system access is included).

Operational distribution is regular, on-going distribution of a defined stream of instances (e.g. discrete products, granules, etc.) of one or more product types (a ‘distribution stream’) to a particular destination. The items collected or aggregated together as a distribution stream, for the purpose of cost estimation, have enough in common that entry of them as separate items is not necessary. Note that for the purpose of cost estimation standing order or ‘subscription’ distribution where the term is a year or more is treated as operational distribution. Operational distribution is described as a set of one or more distribution streams (similar to ingest and processing streams).

By Request distribution is distribution in response to a user request, ranging from a single request to a standing order or subscription request with a term of less than one year. By Request distribution is described by a set of scope and sizing parameters that provide the basis for a projection of year by year activity by the CET.

The first box you will see is one asking for four parameters that provide an overall description of the access and distribution services your activity will provide, as shown in Figure 12 below (with sample entries). See also Table B8, Part A in Appendix B below for additional information.

Figure 12 - Distribution Information Box

Activity Dataset - Distribution Information

Access and Distribution Scope: 1
 1 - Public Access to all users
 2 - Access to Science and/or Applications community
 3 - Access to a limited team of scientists or applications users

Access and Distribution Service Modes: 1
 1 - Operational and By Request and Subscription
 2 - By Request and Subscription
 3 - By Request only - no Operational Distribution
 4 - Operational Distribution only - no By Request Distribution
 By Request includes web access;
 Subscription refers to standing orders.

Distribution Means LOS: 2
 1 - Distribution is only electronic, i.e. by network;
 2 - Distribution is predominantly by network (more than 50% of volume);
 3 - Distribution is predominantly by media (more than 50% of volume);

Supporting Data Services LOS: 2
 (Advisory - No Effect on Estimate)
 1 - Services provided for most data and products
 2 - Services provided for less than half of data and products
 3 - Services provided for a few selected data and products only

Accept

Fill in the blanks, and click on “Accept”. Your entries will be checked for errors. You’ll be asked to review, correct if necessary, and confirm your entries, and then to move on to the next part of access and distribution.

If you indicated in the second box that your activity will perform operational distribution (Access and Distribution Mode 1 or 4), then the next box you will see will ask you for information describing your operational distribution stream(s).

Note that “Supporting Data Services LOS” is an advisory parameter. Your entry will not affect the estimate, but you can adjust the estimate for this as needed later using the Reviewer.

Your entry of operational distribution stream information proceeds as described above for ingest or processing streams. You will be presented with a box for entry of information for an individual distribution stream, as shown in Figure 13 below (with sample entries). See also Table B8, Part B in Appendix B below for additional information.

Figure 13 - Operational Distribution Stream Information Box

Activity Dataset - Operational Distribution Stream Information

Enter the following information for a single Distribution Stream - single product, or aggregate of product types:

Name of Distribution Stream	<input type="text" value="Level_2X"/>	Product types with common characteristics can be aggregated (case sensitive)
Number of Product Types in Stream	<input type="text" value="4"/>	1 - No aggregation, i.e. if distribution stream is a single product type N - Number of products included in distribution stream
Destination for Distribution Stream	<input type="text" value="Center_A"/>	Enter name of destination activity or organization (case sensitive)
Distribution Mode	<input type="text" value="1"/>	1 - Network, 2 - Media
Distribution Stream Start Calendar Yr	<input type="text" value="0"/>	0 - Same as Operations Start or 4 digit calendar year, e.g. 2009
Distribution Stream End Calendar Yr	<input type="text" value="0"/>	0 - Same as Operations End, or 4 digit calendar year, e.g. 2014
Dist. Stream Total Products /Day	<input type="text" value="24"/>	Total count for all product types in aggregate if applicable
Dist. Stream Total Volume per Day, GB	<input type="text" value="3"/>	Total daily volume, GB, for all product types in aggregate

As noted above, a single “distribution stream” can include a single product type, or it can include an aggregate of product types when these have all other properties in common (except for number of products/day and volume/day which are summed over the data/product types being aggregated).

Fill in the blanks, making changes as needed, and click on “Accept this Distribution Stream”. (Note that the Name of Distribution Stream is case sensitive.) Your entries will be checked for errors.

When you click “Accept this Distribution Stream”, a box is provided that displays the information for the distribution stream you have just entered. You can confirm that the information is OK by clicking “Yes”, or that a correction is needed by clicking “No”. If you click “No”, the distribution stream information entry box is brought back with your previous entries shown. You can make any needed corrections and again click “Accept this Distribution Stream”.

After the distribution stream is accepted, you are asked if information for another distribution stream is to be entered. Click “Yes”, to repeat the process described above.

You may also back out of the entry of a distribution stream, or modification of a distribution stream, by clicking on “Back Out of this Distribution Stream”. You will go back to either being asked if you want to enter a new distribution stream, or delete or modify a distribution stream, as appropriate.

If, after entering one or more distribution streams, you enter “No”, the CET proceeds to By-Request distribution, unless you have indicated that your activity will not perform By-Request distribution (i.e., if you have selected Access and Distribution Service Mode 4, Operational Distribution Only).

If your activity will perform By Request distribution, you will next see a By Request distribution box with eight blanks to be filled in, as shown in Figure 14 below (with sample entries). See also Table B8, Part C in Appendix B below for additional information.

Consult the Activity Dataset Checklist (Appendix B below) for more information on these items. Fill in the blanks, and click on “Accept”. After you click “Accept”, the CET checks your entries for errors. You’ll be asked to review, correct if necessary, and confirm your entries, and then proceed to the next selected functional area.

Figure 14 - By Request Distribution Information Box

Activity Dataset - By Request Distribution Information

Information Describing 'By Request' Distribution:

By Request distribution profile	<input type="text" value="1"/>	0 - No by request distribution 1 - Ramp up over mission life 2 - Level / constant over mission life
Expected number of users when services are well established	<input type="text" value="3500"/>	E.g. 4000
Estimated average number of requests per user, per year	<input type="text" value="1.25"/>	E.g. 1.5
Estimated average number of products per request	<input type="text" value="137"/>	E.g. 6
Out of every 10 products distributed, estimated number (0-10) that go by network	<input type="text" value="5"/>	E.g. 8
Out of 10 units of volume distributed, estimated number (0-10) that go by network	<input type="text" value="8"/>	E.g. 5
Response time for requests delivered by network:	<input type="text" value="3"/>	1 - 10 sec for system-system access 2 - 10 sec for FTP or equivalent 3 - 10 minutes for FTP or equivalent 4 - 24 hours for FTP or equivalent
Response time for requests delivered on media:	<input type="text" value="2"/>	1 - Ready to ship in 3 days 2 - Ready to ship in 7 days 3 - Ready to ship in 30 days 0 - Not Applicable, No Media

Accept

3.3.7 New User Support Information

Information needed for the User Support functional area consists of a set of parameters that define the scope and approximate sizing of the By Request distribution service that your activity will provide. At this time, the CET does not estimate user support effort for Operational Distribution. Consult the Activity Dataset Checklist (Appendix B below) for additional information.

Note that “Outreach Activity Level of Service” is an advisory parameter. Your entry will not affect the estimate, but you can adjust the estimate for this as needed later using the Reviewer.

You will see an information entry box with three blanks for information describing user support services to be provided by your activity, as shown in Figure 15 below (with sample entries). The purpose of the user multiplier and the average number of contacts per user per year is to allow estimation of the number of user contacts per year, as explained on the form. As noted the starting point is the number of by-request users expected for the activity. For example, from Figure 14 above, the total number of by-request users when the activity has ramped up is expected to be 3,500. Of that group of 3,500 users, 80% (the User Multiplier of 0.8) are expected to contact user support that year, or 2,800 (i.e. $3,500 \times 0.8$). Those 2,800 users are each expected to make an average of 0.6 contacts to user support (Average Contacts per User per Yr), for a total of 1680 user contacts (i.e. 2800×0.6). This number of user contacts, computed for each year of operations, is the basic workload parameter for the CET’s estimate of user support staff.

See also Table B9 in Appendix B below for additional information.

Figure 15 - User Support Information Box

Activity Dataset - User Support Information

Enter User Support Information:

Outreach Activity Level of Service (Advisory - No Effect on Estimate)

1 - Outreach includes training sessions, expanded booth support at 4 conferences/yr, and production and distribution of outreach material;
2 - Outreach includes expanded booth support at 4 conferences/yr, and production and distribution of outreach material;
3 - Outreach includes booth support at 4 conferences/yr, and production and distribution of outreach material;
4 - Outreach includes production and distribution of outreach material.

User Multiplier Used to estimate number of persons who contact user support per year, which is the estimated number of by request distribution users multiplied by the User Multiplier, which can be a value like 1.25 or .85.

Average Contacts per User per Yr The average number of times per year a user would be expected to contact user support; enter a value like 1.25 or 0.75. This number will be multiplied by the estimated number of users who contact user support to compute the expected number of user contacts per year.

Accept

Fill in the blanks, and click “Accept”. You’ll be asked to review, correct if necessary, and confirm your entries. After you do, the CET will move on to the next selected functional area.

3.3.8 New Non-Operating Functional Area Information

The Non-Operating functional areas include Sustaining Engineering, Engineering Support, Implementation, Technical Coordination, and Miscellaneous Non-Staff Cost items. The information the CET needs from you is divided over three information entry boxes.

The first information entry box you will see asks for information about sustaining engineering, engineering support, and implementation), as shown in Figure 16 below (with sample entries).

See also Table B10 (implementation), B11 (sustaining engineering), and B12 (engineering support) in Appendix B below for additional information.

Figure 16 - Implementation and Engineering Information Box

Activity Dataset - Implementation and Engineering

Implementation - Applications Software (Software developed to meet specific user needs)

Applications Software Development LOS (Advisory - No Effect on Estimate)

- 1 - Data mining/integration, custom science products, data manipulation tools;
- 2 - Custom science products, data manipulation tools;
- 3 - Data manipulation tools.
- 4 - None.

Sustaining Engineering (Software maintenance, enhancement)

Sustaining Engineering LOS

- 1 - No or very infrequent interruptions in service allowed;
- 2 - Occasional interruptions in service allowed;
- 3 - Interruptions in service are a secondary concern.

Engineering Support (System admin, network admin, DB admin, etc.)
(System engineering, test engin., CM, etc.)

Engineering Support LOS (Advisory - No Effect on Estimate)

- 1 - No or very infrequent interruptions in service allowed;
- 2 - Occasional interruptions in service allowed;
- 3 - Interruptions in service are a secondary concern.

Accept

Note that “Applications Software Development LOS” and “Engineering Support LOS” are advisory parameters. Your entries for these will not affect the estimate, but you can adjust the estimate as needed later using the Reviewer.

Additional information about the items needed on this box can be found in the Activity Dataset Checklist (Appendix B below). Fill in the blanks, and click on “Accept”. Your entries will be checked for errors. You’ll be asked to review, correct if necessary, and confirm your entries, and then to proceed to the second Non-Operational form.

The second box covers technical coordination, and asks you to indicate those areas in which your activity will participate in NASA Earth science data activity coordination, as shown in Figure 17 below (with sample entries). See also Table B13 in Appendix B below for additional information.

Figure 17 - Technical Coordination Information Box

Activity Dataset - Technical Coordination

Indicate areas in which your activity coordinates with other data activities:

- ☒ Metrics
- ☐ Technology Infusion
- ☒ Data Life Cycle
- ☐ Re-Use
- ☐ Standards
- ☐ Levels of Service
- ☒ User Services

Accept

Check the appropriate boxes, and click “Accept”. You will be asked to review, correct if necessary, and confirm your entries, and to move on to the third Non-Operational box.

The third box covers Miscellaneous Non-Staff Cost items, including Backup Archive Facility LOS, and additional cost areas: budget for data purchase, budget for computer services, and budget for travel, as shown in Figure 18 below (with sample entries). See also Table B14 in Appendix B below for additional information.

Figure 18 – Miscellaneous Non-Staff Cost Information Box

Activity Dataset - Misc Non-Staff Items

Backup Archive Facility LOS: **(Advisory - No Effect on Estimate)**

1 - Secure off-site back-up archive facility;
2 - On-site but separate backup archive facility;
3 - Backup storage with primary storage;
0 - No Backup storage.

Other Costs:

Annual budget for data purchase: Enter \$K, e.g. 25

Annual budget for computing services: Enter \$K

Annual budget for Travel: Enter \$K

Accept

These items are described in Appendix B, Section 3.5 and Table B14.

Note that “Backup Archive Facility LOS” is an advisory parameter. Your entry will not affect the estimate, but you can adjust the estimate for this as needed later using the Reviewer.

Fill in the blanks, and click “Accept” to complete the entry of information for the Non-Operational functional areas. Your entries will be checked for errors. You’ll be asked to review, correct if necessary, and confirm your entries.

When information for the last selected functional area has been entered and confirmed, the CET will ask you if you wish to generate a life cycle cost estimate for your activity, using the now complete Activity Dataset. Go to Section 3.5 for the next steps.

3.4 Work With an Existing Activity Dataset

At this point, you have selected an Activity Dataset name from the list, and clicked “Proceed”, or you’ve chosen to do a “what if” exercise on the current Activity Dataset after producing a life cycle cost estimate.

You may make changes to the Activity Dataset you selected, and then produce a life cycle cost estimate for it, or proceed directly to producing a cost estimate without changing it.

If you make changes, in effect creating a new version of the Activity Dataset, you will be able to either save the new version under a new name, or replace the old version with the modified version under the current name.

The CET will ask if you would like to make any modifications to the Activity Dataset. To proceed directly to producing an estimate, click “No”.

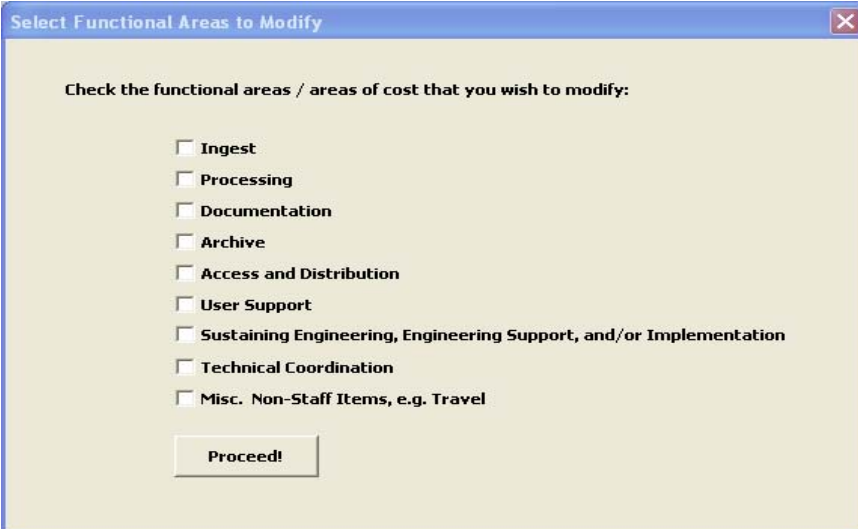
3.4.1 Modify Activity / Site Level Information

If you click “Yes” indicating that you wish to make changes, you will next see a box titled “Activity Dataset - Data Activity Level Information”. This is the same box you used when defining a new Activity Dataset, except that now, instead of being blank, the text fields contain the information for the existing Activity Dataset that is to be modified. See the example in Section 3.3.1, Figure 6 above.

Make any changes you desire, and select “Looks Good, Proceed”, or “Exit from CET” to exit.

When you click “Proceed” you will see a box displaying a list of the functional areas asking you to check those that you wish to make changes to, as shown in Figure 19 below:

Figure 19 - Select Functional Areas to Modify Box



Select Functional Areas to Modify

Check the functional areas / areas of cost that you wish to modify:

- ☒ Ingest
- ☐ Processing
- ☐ Documentation
- ☐ Archive
- ☐ Access and Distribution
- ☐ User Support
- ☐ Sustaining Engineering, Engineering Support, and/or Implementation
- ☐ Technical Coordination
- ☐ Misc. Non-Staff Items, e.g. Travel

Proceed!

If you wish to add a new operational function to your Activity Dataset, you can do so by checking that function for modification, as was done for Ingest in the example shown.

Once you have made your choices, click on “Proceed”. You’ll be asked to review a list of the functions you checked, make corrections if necessary, and confirm your choices. Then you will be lead through modification of those functions that you selected.

3.4.2 Modify Ingest Information

You can either add one or more additional ingest streams, modify one or more existing ingest streams, or delete one or more ingest streams.

a. Add an Ingest Stream

You will first be asked if you wish to add a new ingest stream. If you do, click “Yes”, and you will see same ingest stream information box that you used when you entered a new Activity Dataset, with blank text fields for you to fill in. See the example in Section 3.3.2 above, Figure 7. When you have entered the information describing the new ingest stream, click “Accept this Stream”. The process proceeds through a verification step as you saw when you entered a new Activity Dataset, and you may enter as many new ingest streams as you wish. After you enter each one, you are asked if you wish to add another ingest stream, and if you click “Yes” the process just described is repeated.

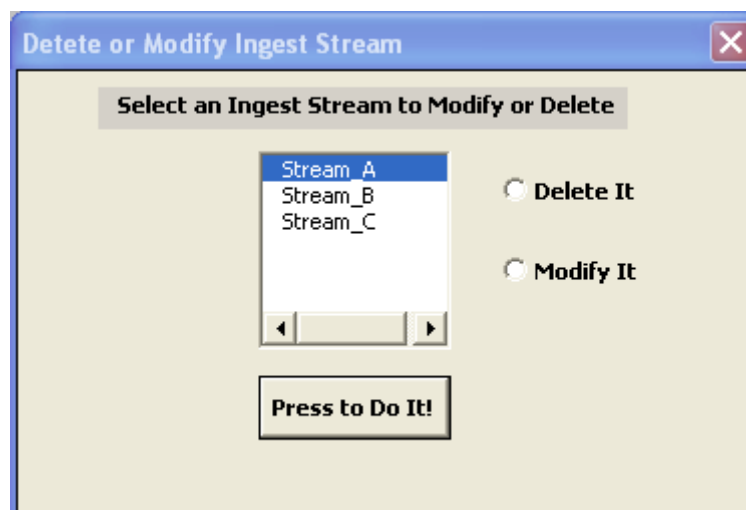
Once you’ve finished entering new ingest streams (if any) you will be asked if you want to make changes to, or delete, an existing ingest stream. Click “No” to proceed to the next selected functional area. Click “Yes” to modify or delete an ingest stream.

You may also back out of the entry of the new ingest stream, or correction/modification of an ingest stream, by clicking on “Back Out of this Stream”. You will go back to being asked if you want to add an ingest stream.

b. Modify an Ingest Stream

If you have clicked “Yes”, you will be presented with a box titled “Delete or Modify Ingest Stream” that contains a list of the current ingest streams, “Delete It” or “Modify It” options and a “Press to Do It” button, as shown in Figure 20 below:

Figure 20 - Delete or Modify Ingest Stream Box



Highlight the ingest stream to be modified or deleted.

If you want to modify the ingest stream, select the “Modify It” option, and click “Do it Now”. You will then see an “OK” box confirming the name of the ingest stream to be modified, and then the box for ingest stream information entry, with fields in the box containing the current

information for the ingest stream you selected for modification, and a button labeled “Accept this Stream”, as shown in Figure 7 above. Make the changes you wish, and then click “Accept this Stream”.

You will then be asked if you want to make changes to or delete another ingest stream. If you click “No”, indicating that modification of ingest information is complete, the CET proceeds to the next functional area you selected for modification.

You may also back out of the entry of a modified ingest stream, or correction/modification of an ingest stream, by clicking on “Back Out of this Stream”. You will go back to being asked if you want to delete or modify an ingest stream.

c. Delete an Ingest Stream

If you want to delete an ingest stream, select the “Delete It” option and click “Press to Do It”. You’ll be given an “OK” box confirming the name of the ingest stream to be deleted, and you’ll be asked if there is another ingest stream to be modified or deleted.

If you click “No”, indicating that entry of ingest information is complete, the CET proceeds to the next functional area you selected for modification.

3.4.3 Modify Processing Information

You can either add one or more additional processing streams, modify one or more existing processing streams, or delete one or more processing streams.

a. Add a Processing Stream

First, you will be asked if you wish to add a new processing stream. If you do, click “Yes”. You will then see the processing stream information box (that same box as used when a processing stream for a new Activity Dataset is being entered) with blank text fields for you to fill in. See the example, Figure 9 in Section 3.3.3 above. Enter the information describing the new processing stream, and click “Accept this Processing Stream”. You’ll be given an opportunity to make corrections, and then you will be asked if you wish to add another processing stream, and if you do, click “Yes” and the process just described will be repeated.

You may also back out of the entry of a new processing stream, or correction/modification of a processing stream, by clicking on “Back Out of this Processing Stream”. You will go back to being asked if you want to add a processing stream.

If you click “No”, you will be asked if you wish to modify the Activity level processing flags (science software LOS and calibration-validation LOS). If you do, click “Yes”. The same box that you used to enter the flags originally is then displayed, with the current values of the flags shown. See Figure 8. Make any desired changes and click “Accept”.

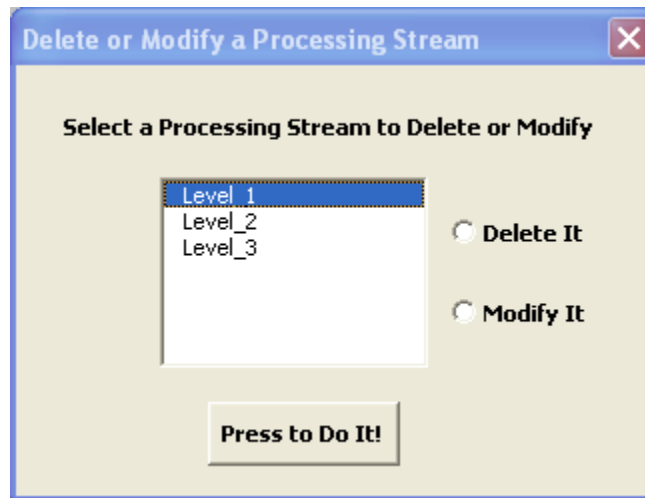
After you’ve made any changes to the processing flags you will be asked if you wish to modify or delete a processing stream.

If you click “No”, indicating that modification of processing information is complete, the CET proceeds to the next functional area you selected for modification.

b. Modify a Processing Stream

If you wish to make changes to, or delete, one or more processing streams, click “Yes”. You will see a box titled “Delete or Modify Processing Stream” that contains a list of the current processing streams, “Delete It” or “Modify It” options and a “Press to Do It” button, as shown in Figure 21 below:

Figure 21 - Delete or Modify Processing Stream Box



Highlight the processing stream to be modified or deleted.

If you want to modify the selected processing stream, check the “Modify It” option, and click “Press to Do It”. You will see an “OK” box confirming the name of the processing stream to be modified. Then you will see the entry box for the processing stream. This is the same box that you used for entry of information for a new processing stream, except that the blanks in the box contain the current information for the processing stream you selected for modification. See the example, Figure 9 in Section 3.3.3 above.

Makes the changes you wish, and then click “Accept this Processing Stream”. You’ll then be asked if you want to modify or delete another processing stream. If you do, click “Yes”, and the process described above will be repeated.

If you don’t want to modify or delete a processing stream, click “No”, indicating that modification of processing stream information is complete. The CET will proceed to the next selected functional area.

You may also back out of the entry of a modified processing stream, or correction/modification of a processing stream, by clicking on “Back Out of this Processing Stream”. You will go back to being asked if you want to delete or modify a processing stream.

c. Delete a Processing Stream

If you want to delete the selected processing stream, check the “Delete It” option and click “Press to Do It”. You’ll be given an “OK” box confirming the name of the processing stream to be deleted, and you’ll be asked if there is another processing stream to be modified or deleted.

If you are finished, click “No”, to indicating that modification of processing information is complete. The CET will proceed to the next functional area you selected for modification.

3.4.4 Modify Documentation Information

You will see a box containing the values for the two level of service parameters you entered previously. See the example, Figure 10 in Section 3.3.4 above. Make any changes you wish, and click “Accept” to proceed to the next functional area you selected for modification.

3.4.5 Modify Archive Information

You will see the information entry box for the Archive functional area with ten fields containing the information you entered previously. See the example, Figure 11 in Section 3.3.5 above. You can find additional information on the items in the Activity Dataset Checklist (Appendix B below). Make any changes you wish, and click on “Accept Archive Information” to move on to the next functional area you selected for modification.

3.4.6 Modify Access and Distribution Information

As described above (Section 3.3.6) Access and Distribution information consists of three parts, a description of the overall access and distribution service to be provided by your activity, operational distribution if any, and by request distribution.

The first box you will see is one describing the overall access and distribution services your activity will provide. It contains the four values you entered previously. See the example, Figure 12 in Section 3.3.6 above. Make any changes you wish, and press “Accept” to move on to the next part of access and distribution.

If your activity includes operational distribution, you can either add one or more additional operational distribution streams, modify one or more existing operational distribution streams, or delete one or more operational distribution streams.

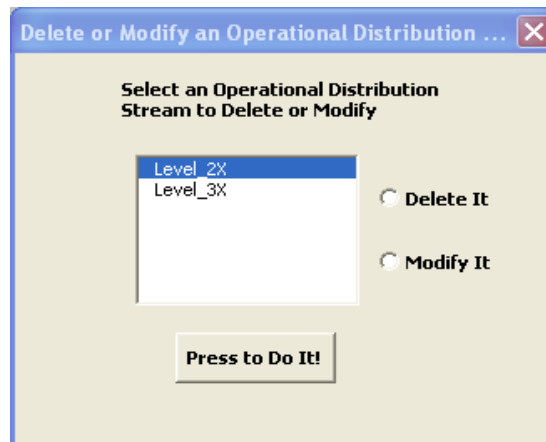
a. Add an Operational Distribution Stream

If you indicated that your activity will perform operational distribution, then the CET will ask if you wish to add an additional operational distribution stream. Click “Yes” if you wish to add an additional stream, and the CET will present you with the operational distribution stream information entry box, with blank fields for you to fill in. Provide the needed information (see the example, Figure 13 in Section 3.3.6 above) and click on “Accept this Distribution Stream”. You will be given opportunity to review and revise your entries.

You may also back out of the entry of a new distribution stream, or correction/modification of a distribution stream, by clicking on “Back Out of this Distribution Stream”. You will go back to being asked if you want to add a distribution stream.

Once you have accepted the new distribution stream(s) and indicate that you do not wish to enter another new distribution stream, the CET will ask if you wish to modify or delete an existing operational distribution stream. If you wish to do so, click “Yes” and the CET will display a box displaying a list of the current operational distribution streams, as shown in Figure 22 below:

Figure 22 - Delete or Modify Operational Distribution Stream Box



Select the distribution stream you wish to modify or delete, choose “Delete It” or “Modify It” as you wish, and click on “Press to do it!”.

b. Modify an Operational Distribution Stream

If you’ve chosen to modify an operational distribution stream, the CET will display the operational distribution stream box with the information you previously entered for the selected operational distribution stream. See the example, Figure 14 in Section 3.3.6 above. Make any changes you wish, and click “Accept this Distribution Stream”. After giving you a chance to review and revise your entries, you’ll be asked if you wish to modify or delete another operational distribution stream, and if you wish to do so you can repeat the process just described by clicking “Yes”. Otherwise, click “No” and the CET will move on to by-request distribution information.

You may also back out of the entry of a modified distribution stream, or correction/modification of a distribution stream, by clicking on “Back Out of this Distribution Stream”. You will go back to being asked if you want to delete or modify a distribution stream.

c. Delete an Operational Distribution Stream

If you’ve chosen to delete an operational distribution stream, you’ll be asked to confirm that choice.

d. Modify By Request Distribution

You will next see the By Request distribution information entry box with eight fields containing the information you provided previously. See the example, Figure 14 in Section 3.3.6 above. Make any changes you wish, and click on “Accept”. After you click “Accept”, the CET proceeds to the next functional area you selected for modification.

3.4.7 Modify User Support Information

You will see an information entry box with six fields containing information you entered previously to describe the user support services to be provided by your activity. See the example, Figure 15 in Section 3.3.7 above. Make any changes you wish, and click “Accept”. After you do, the CET will move on to the next functional area you selected for modification.

3.4.8 Modify Non-Operating Functional Area Information

The first information entry box you will see contains information you entered previously about sustaining engineering, engineering support, internal support (part of facility / infrastructure but included here for convenience), and implementation (development). See the example, Figure 16 in Section 3.3.8 above. Make any changes you wish, and click on “Accept” to proceed to the second Non-Operating box.

The second box covers technical coordination, and indicates those areas in which you previously indicated that your activity will participate in ESE coordinating activities. See the example, Figure 17 in Section 3.3.8 above. Make any changes you wish, and click “Accept” to move on to the third Non-Operating box.

The third box covers the Facility / Infrastructure area, and displays the information that you entered previously. See the example, Figure 18 in Section 3.3.8 above. Make any changes you wish, and click “Accept” to complete the modification of information for the non-Operational functional areas.

When modifications for the last functional area you selected have been completed, the CET will ask you if you wish to generate a life cycle cost estimate for your activity, using the newly modified Activity Dataset. Go to Section 3.5 for the next steps.

3.5 Produce Life Cycle Cost Estimate and Quality Report

When you have entered / modified information for all selected functional areas, the CET presents a box asking if you wish to proceed with a life cycle cost estimate.

If you would like a life cycle cost estimate, click “Yes”. Otherwise, click “No”, and proceed to Section 3.6 for the next steps.

If you click “Yes”, the CET will generate a life cycle cost estimate for your data activity, based on the information you have entered into the current active Activity Dataset.

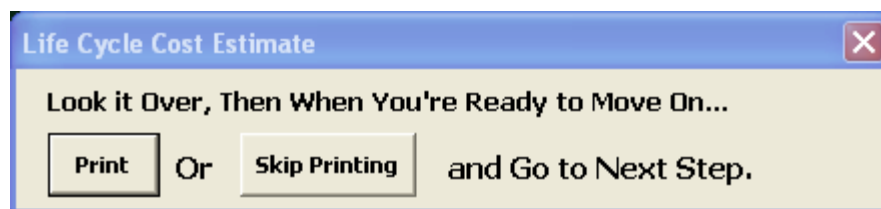
The time required by the CET to produce its estimate varies, e.g. from about five seconds on a 3GHz PC platform to 45 seconds on a 1 GHz Macintosh iMac platform.

The CET will produce your life cycle cost estimate, and as it executes, it reports its progress. When running on a PC, the CET displays its progress on the gray sheet beneath the CET logo as it works through the data activity functions. The CET also reports progress on the Excel status bar (located at the bottom of the Excel window; the progress messages appear at the left corner of the status bar).

As discussed in Section 2, the CET’s life cycle cost estimate is contained in two spreadsheets. The first contains the overall estimate of effort and cost for your activity, and the second contains a breakdown of estimated effort by functional area. The life cycle cost estimate is followed by a Quality Report. See the examples, Tables 1 and 2 and Figure 2 provided in Sections 2.3 and 2.4 above.

The first output spreadsheet is displayed along with a box titled “Life Cycle Cost Estimate” as shown in Figure 23 below. The box is located in the upper left hand corner of the screen so as to not obscure any part of the output worksheet.

Figure 23 - Life Cycle Cost Estimate Box



PC users (actually those using Excel 2002 or newer) may scroll around the output page before choosing to print it or not. Scrolling will not work for Mac users. When you click to print (you’ll get a standard Excel print window if you’ve clicked to print) or click to skip printing, the CET will display the second page of the life cycle cost estimate output with the same box, again offering you the choice of printing the second page of the output or skipping the print. When you’ve clicked print or skip printing the CET display the Quality Report, with the same box asking you to print or skip printing the Quality Report. When you have clicked print or skip printing, the CET will proceed to the next step, described in Section 3.6 below.

Note that the output worksheets are not erased when the CET run is completed and you have exited from the CET. After you exit the CET, but with the CET workbook open, you can tab to the output worksheets and examine them at your leisure, scrolling about at will, and print if you wish. Because the screen display characteristics of the PC and Macintosh are different, there are two sets of output worksheets and you must look at the set the CET uses for your platform. The

two versions of the first page are named “Output” for the PC and “Output_M” for the Macintosh, likewise the two versions of the second page are named “Output2” and “Output2_M”, and the two versions of the Quality Report are named “Quality” and “Quality_M”.

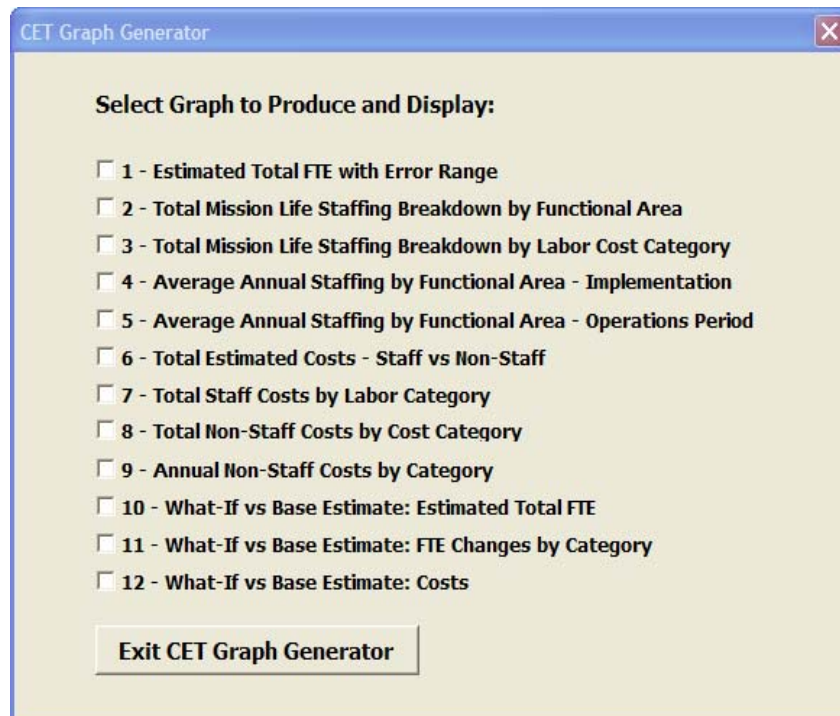
Note further that the two pages of life cycle cost estimate output are appended to the Activity Dataset.

3.6 Generate Graphs

After displaying the Quality Report, the CET will ask if you wish to generate graphs. If you click on yes, the CET displays the graph menu shown in Figure 24 below. The menu is a list of twelve graphs that can be generated from the results of the life cycle cost estimate the CET has just produced. When you check a graph, the CET will produce and display that graph along with a small box asking you to check on a button to proceed, or to print the graph and then proceed. When you’ve looked at the graph long enough, click on one of the proceed buttons. The CET then repeats the graph menu, and you may either select another graph to display or exit the CET Graph Generator, which takes you to the next step (see Section 3.7 below).

The graphs are discussed briefly below, followed by samples of each type of graph.

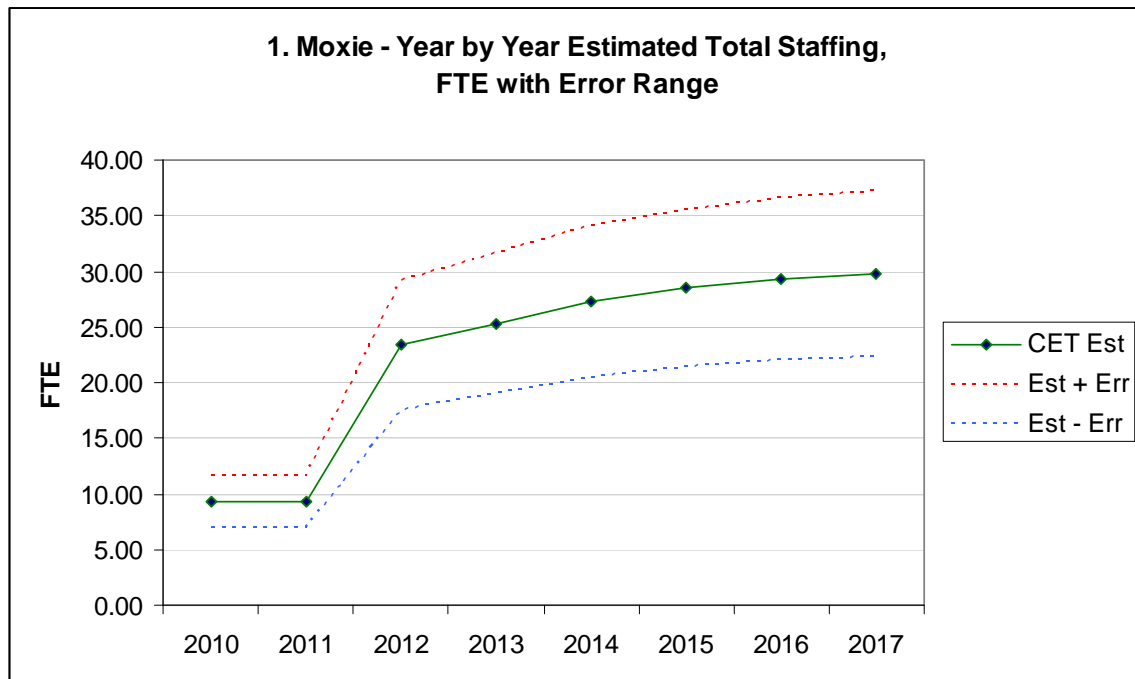
Figure 24 – Graph Menu



The graphs are line graphs or pie charts. They are displayed in color, and may be printed in either color or black and white depending on your selection of printer. Note, especially as an aid to interpreting the pie charts on a black and white printout, that the plotted categories are listed in the graph legend in alphabetical order, and the pie sectors follow that alphabetical listing in clockwise order beginning from the 12:00 position on the pie.

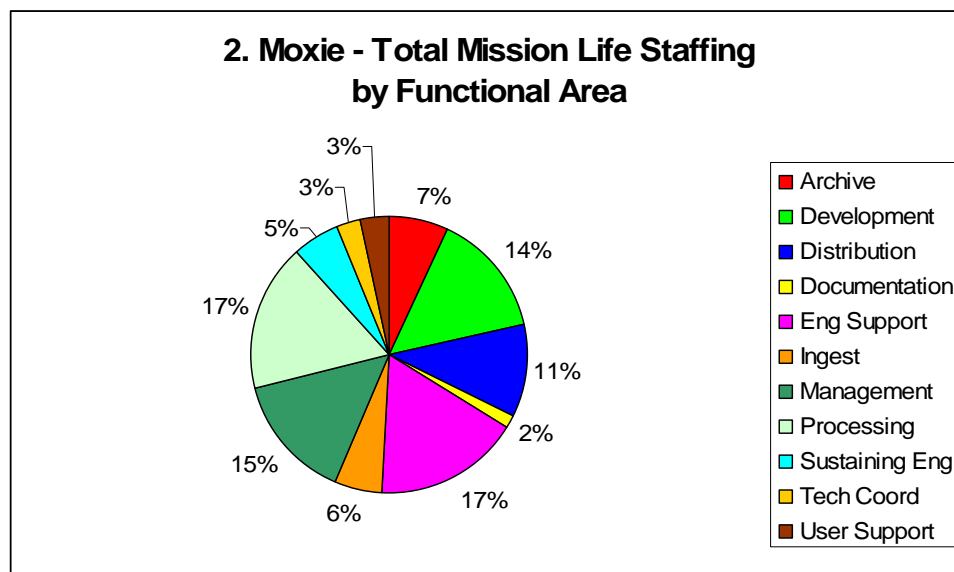
Graph 1, Estimated Total FTE with Error Range (see sample in Figure 25 below), plots the year by year total estimated staffing in FTE as a solid line, with two dashed lines marking the + or – 25% CET error bars. See Section 2.3.3 (page 18) for a discussion of the CET’s performance.

Figure 25 – Graph 1, Estimated Total FTE with Error Range



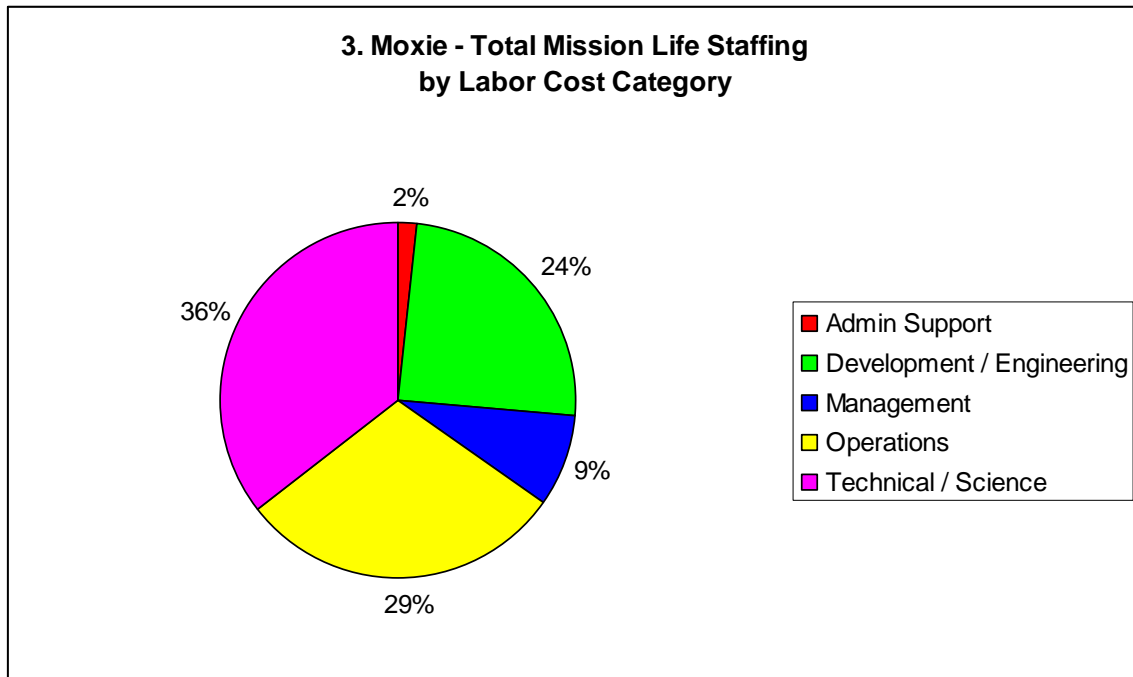
Graph 2, Total Mission Life Staffing Breakdown by Functional Area (see sample in Figure 26 below), is a pie chart that shows the percentage distribution of total life cycle FTE estimated for your activity across the eleven functional areas of (alphabetically) archive, development, distribution, documentation, engineering support, ingest, management, processing, sustaining engineering, technical coordination, and user support).

Figure 26 - Graph 2, Total Mission Life Staffing Breakdown by Functional Area



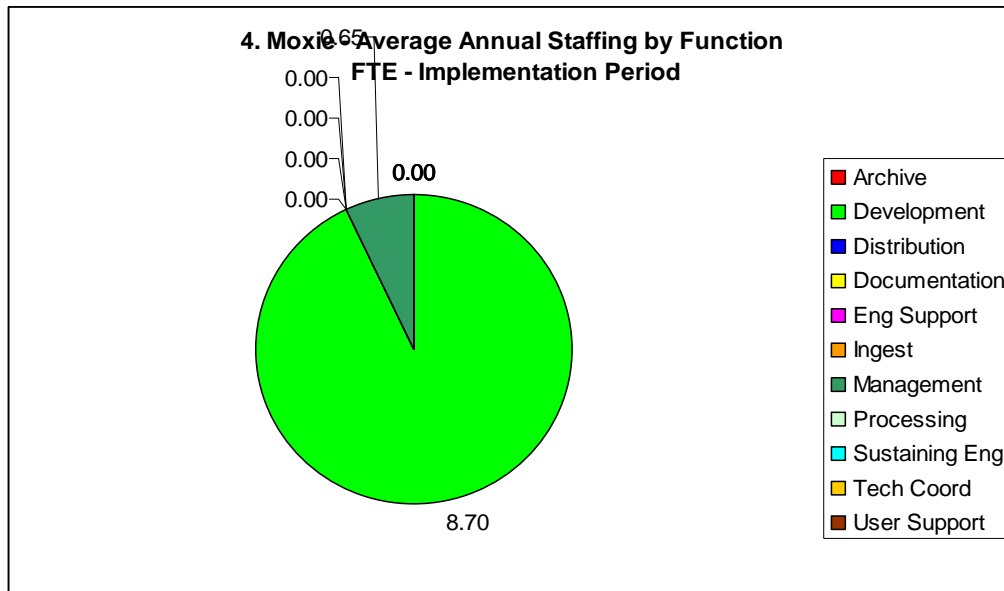
Graph 3, Total Mission Life Staffing Breakdown by Labor Cost Category (see sample in Figure 27 below), is a pie chart that shows the percentage distribution of the total life cycle FTE estimated for your activity across the five high level labor cost categories used by the CET: (alphabetically) administrative support, development / engineering, management, operations, and technical / science.

Figure 27 - Graph 3, Total Mission Life Staffing Breakdown by Labor Cost Category



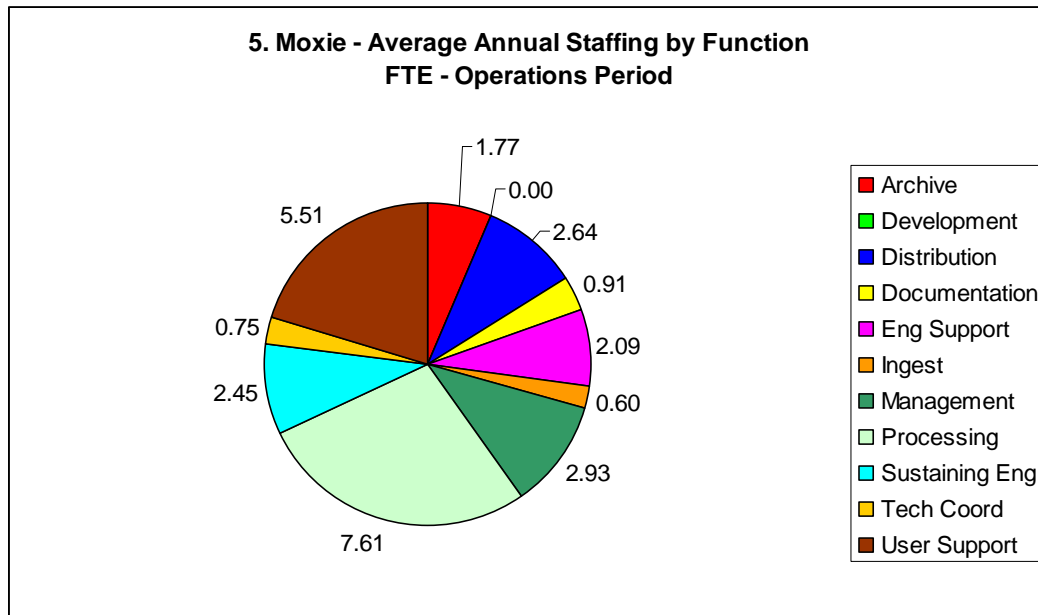
Graph 4, Average Annual Staffing by Functional Area – Implementation (see sample in Figure 25d below), is a pie chart that shows the estimated average staff level in FTE for each of the eleven functional areas (alphabetically - archive, development, distribution, documentation, engineering support, ingest, management, processing, sustaining engineering, technical coordination, and user support) during the implementation period. In most cases, there will be FTE during the implementation period in only two areas, development and management. In some cases development will partially overlap operations, and FTE in other areas will be shown.

Figure 28 - Graph 4, Average Annual Staffing by Functional Area - Implementation



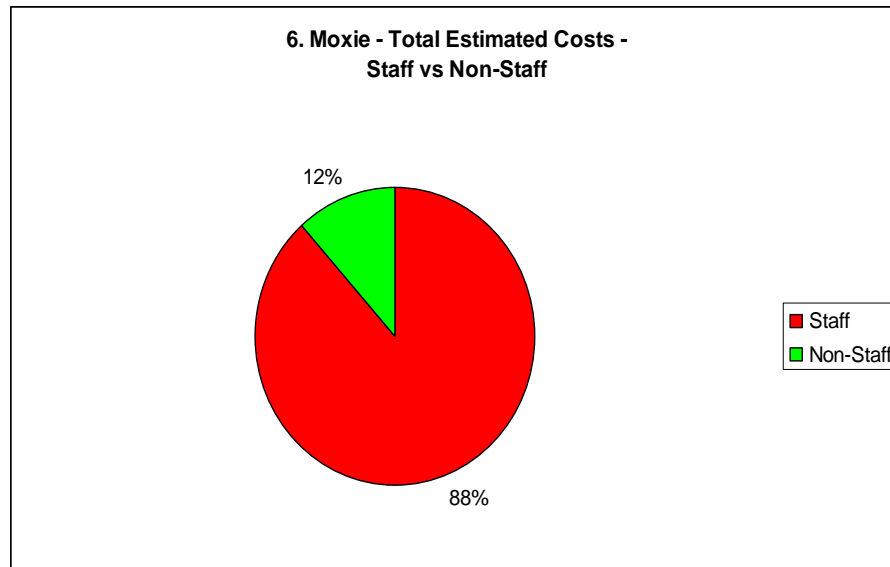
Graph 5, Average Annual Staffing by Functional Area – Operations Period (see sample in Figure 29 below), is a pie chart that shows the estimated average staff level in FTE for each of the eleven functional areas during the operations period.

Figure 29 - Graph 5, Average Annual Staffing by Functional Area – Operations Period



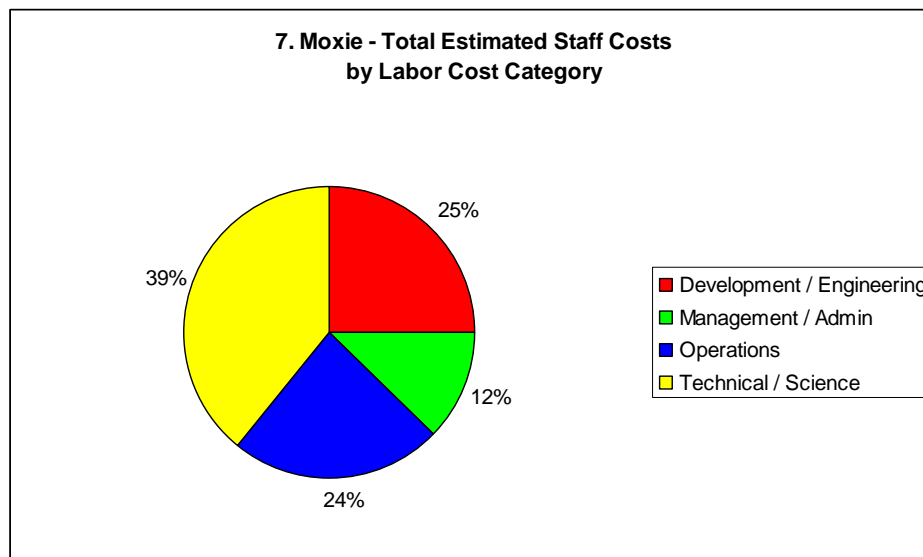
Graph 6, Total Estimated Costs, Staff vs Non-Staff (see sample in Figure 30 below), is a simple pie chart that shows the distribution of estimated life cycle costs between staff costs and all non-staff costs.

Figure 30 - Graph 6, Total Estimated Costs, Staff vs Non-Staff



Graph 7, Total Staff Costs by Labor Category (see sample in Figure 31 below), is a pie chart that shows the distribution of estimated life cycle staff costs across four labor cost categories, development / engineering, management / administrative support, operations, and technical / science.

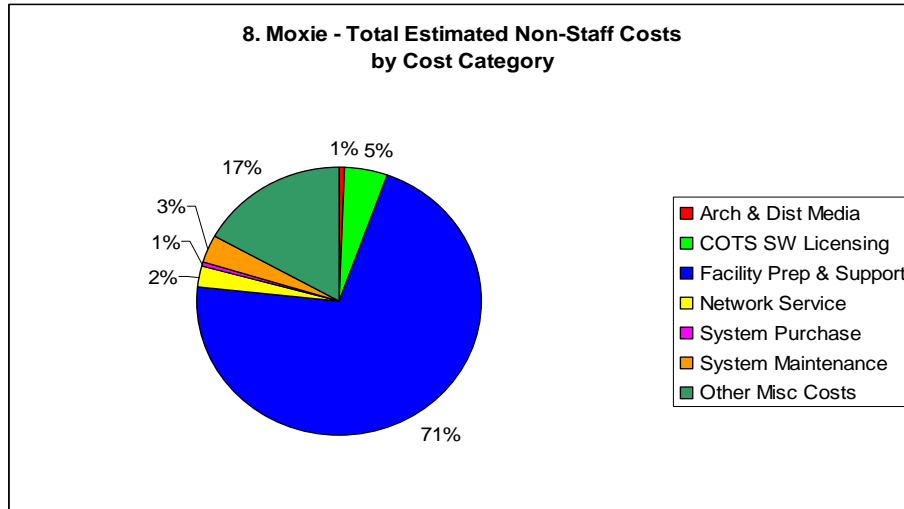
Figure 31 - Graph 7, Total Staff Costs by Labor Category



Graph 8, Total Non-Staff Costs by Cost Category (see sample Figure 32 below), is a pie chart that shows the distribution of estimated life cycle non-staff costs between the cost categories of (alphabetically) archive and distribution media costs, COTS software licensing costs, facility preparation and support costs, network service costs, system purchase cost (i.e. hardware and

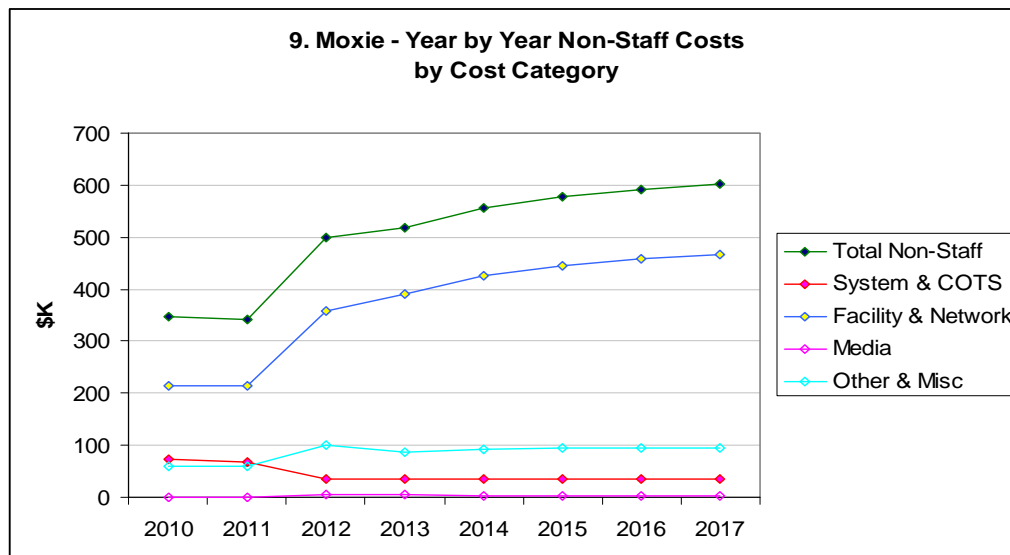
system software), system maintenance costs (i.e. hardware maintenance and system software support), and other miscellaneous costs (including travel cost, data purchase costs, computer service costs).

Figure 32 - Graph 8, Total Non-Staff Costs by Cost Category



Graph 9, Annual Non-Staff Costs by Category (see sample Figure 33 below), is a line graph that shows the year by year total non-staff cost and distribution of that total non-staff cost across five categories, system and COTS (system purchase, COTS software licensing, system maintenance), facility and network (facility preparation and support and network services costs), media (archive and distribution), and other miscellaneous costs.

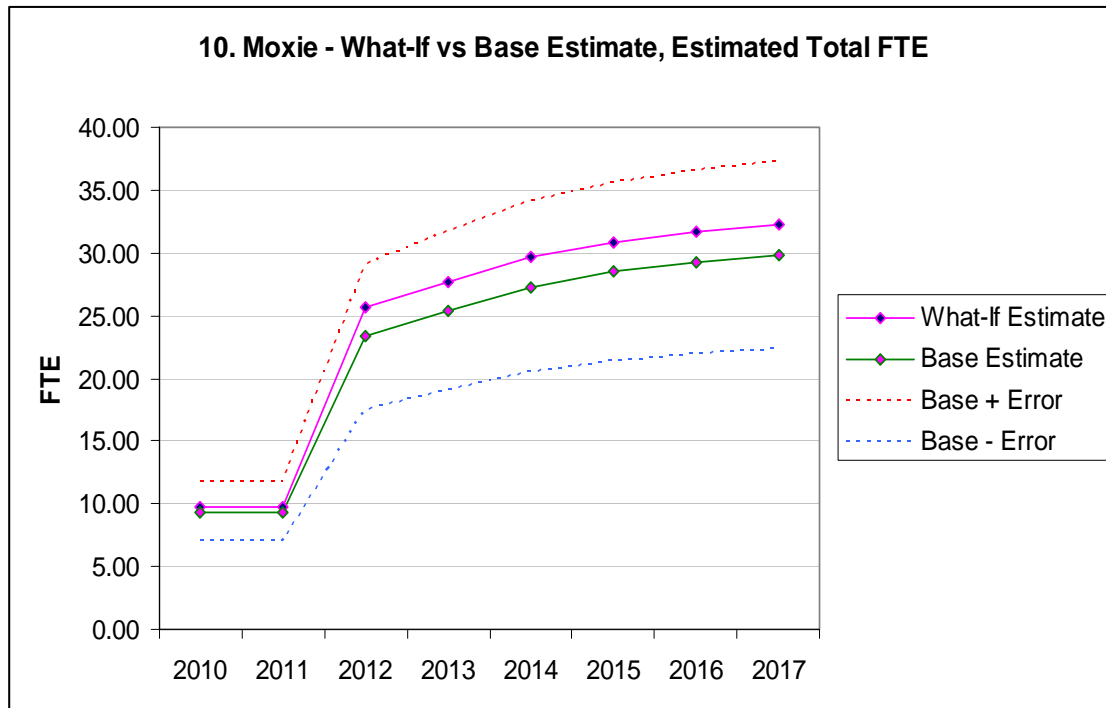
Figure 33 – Graph 9, Annual Non-Staff Costs by Category



Graph 10, What-If vs Base Estimate: Estimated Total FTE (see sample Figure 34 below), is a line graph that compares the year by year estimated total FTE for the last 'What-If' run against the original year by year base estimated FTE and its error range (i.e., it adds the 'What-if' estimated FTE to Graph 1). This graph, which is only available after a 'What-If' has been run, allows you to

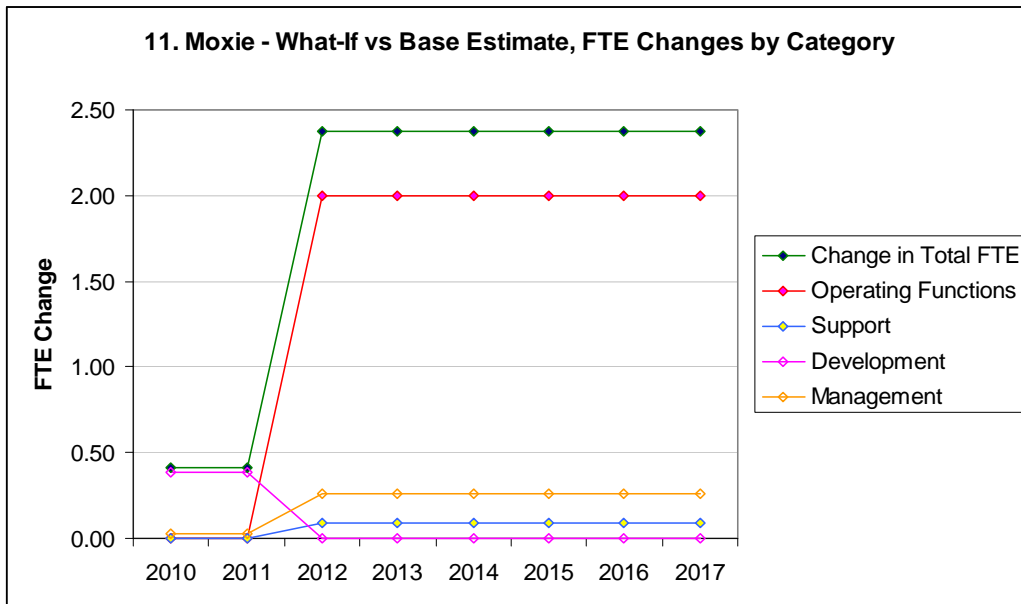
see, for example, if the ‘What-if’ changes are within the error range of the original FTE estimate or are large enough to push beyond the original FTE error range. The ‘What-If’ example shown in this graph and in graphs 11 and 12 below is the example discussed in Section 3.9 below.

Figure 34 - Graph 10, What-If vs Base Estimate: Estimated Total FTE



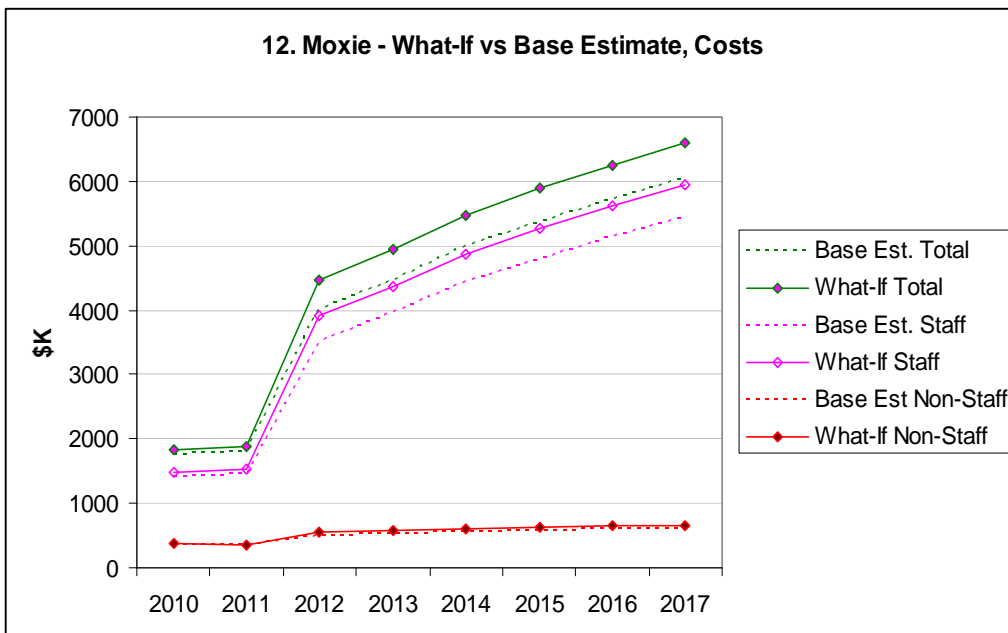
Graph 11, What-If vs Base Estimate: FTE Changes by Category (see sample Figure 35 below), is a line graph that compares the year by year changes in (‘What-If’ value less the base value) the total FTE, and five categories of FTE, operating functions (i.e. ingest, processing, archive, distribution, user support), support functions (sustaining engineering, engineering support), development, and management. This graph is only available if a ‘What-If’ has just been run.

Figure 35 - Graph 11, What-If vs Base Estimate: FTE Changes by Category



Graph 12, What-If vs Base Estimate: Costs (see sample Figure 36 below), is a line graph that compares the year by year total cost, staff cost, and non-staff cost for the last 'What-If' run to the corresponding year by year base estimated costs. This graph is only available if a 'What-If' has just been run.

Figure 36 - Graph 12, What-If vs Base Estimate: Costs



3.7 CET Sensitivity Test

After you have finished with (or have chosen not to generate) CET graphs, the CET asks if you wish to perform a sensitivity test. You may click ‘yes’ to do so or ‘no’ to proceed to the next step (see Section 3.8 below). The next sections discuss the background and purpose for the sensitivity test (3.7.1 below), how the sensitivity test complements the “What-If” capability (3.7.2 below), the sensitivity test menu (3.7.3 below), and the sensitivity test results (3.7.4 below).

3.7.1 Background and Purpose for the Sensitivity Test

The purpose of the sensitivity test is to help you understand the comparative sensitivity of the life cycle cost estimate produced by the CET to the various data activity workload parameters on which the estimate is based. Sensitivity for this purpose is measured by the percent change in the CET’s estimate produced by a specified percent change in the value of the workload parameters. Because the key cost driver for data activities in general is staff cost, and because staff costs are estimated by the CET by applying labor rates you provide to FTE estimated by the CET, the CET’s FTE estimates are the key. Therefore the sensitivity test measures the effect that changes in workload have on changes in estimated FTE, which in turn drive changes in cost.

The CET’s sensitivity test allows you to select a workload parameter and see a plot of the percent changes in the total estimated FTE for a range of changes in the workload parameter, from -50% (i.e. cutting the value in half) to 100% (i.e. doubling the value). Running the test for a sequence of workload parameters lets you see to which ones the CET’s FTE estimate is most sensitive to variations in. This can tell you, for example, which workload parameters it is most advisable for you to get the best handle you can on in order to have greater confidence in the CET’s estimate.

The sensitivity test takes into account the fact that a change in one workload parameter may ripple down into other parameters, e.g. a change in volume ingested will induce a change in volume archived. Also changes in workload parameters will ripple into changes in implementation and sustaining engineering.

3.7.2 Sensitivity Test and the What-If Function

The CET’s sensitivity test complements but does not replace the CET’s ‘What-If’ function (described below in Section 3.9).

The CET’s life cycle cost estimate is based on a comparison between your data activity (as described in the Activity Dataset that you entered) and information about comparable data activities. The points of comparison are various workload parameters associated with the operating functional areas, especially ingest, processing, and distribution. These are used to develop a set of regression equations, from which the best equation is selected, for estimating the FTE as a function of your data activity’s workload parameters, functional area by functional area. The sensitivity test uses the equation for each functional area that is used to produce the base estimate, varying the workload parameters and obtaining corresponding FTE estimates, functional area by functional area. The greater the departure of the value of the workload parameter from its base value, the less accurate will be the resulting FTE estimate, which is why the range of variation of workload parameters for the sensitivity test is constrained to -50% to +100%.

The ‘What-If’ function, on the other hand, takes your revisions to one or more workload parameters (or labor rates or other items) and re-runs the process by which the regression equations are developed, tested, and the best one selected. For this reason, if you compare the results of a sensitivity test on a 100% increase in the value of workload parameter A with the result of a What-If run with the same increase in the same parameter, you will obtain a different result because a slightly-to-significantly different regression equation will have been used. But

the ‘What-If’ result should be a better result precisely because the whole estimating process has been re-run. Therefore you can think of the sensitivity test as a quick way to examine the sensitivity of the CET’s FTE estimate to ranges of variation in a number of workload parameters, and perhaps pick some for full-scale ‘What-If’ tests where you see that the sensitivity is greatest and perhaps where your uncertainty is greatest. The ‘What-If’ tests will also produce life cycle cost estimates as well as FTE estimates, and allow you to compare both to the base estimate. Certainly, to see the effect of variations outside of the -50% to +100% range in a workload parameter you will need to run a ‘What-If’.

3.7.3 Sensitivity Test Menu

If you respond with by clicking on “yes” when the CET asks if you wish to run a sensitivity test, the CET presents you with a menu of workload parameters as shown in Figure 37 below. When you check a workload parameter to test, the CET will generate the sensitivity test results as shown in Figure 38 below (which will be discussed in the next section). When you’ve examined the results, and click on ‘Proceed’ or “Print and Proceed”, the CET will print the results if selected and return you to the sensitivity test menu, allowing you to select another workload parameter to test or to exit the sensitivity test and move on to the next step in the CET’s process (see Section 3.8 below). Note that on a Macintosh platform, when you click on ‘Proceed’ or “Print and Proceed” the CET will return you to the original question, asking you if you wish to do a sensitivity test.

Figure 37 lists a menu of twelve workload parameters that you may select from, grouped by the functional area (ingest, processing, or distribution) that they pertain to.

Figure 37 – Menu of Workload Parameters for Sensitivity Test

CET Sensitivity Test

Test the Sensitivity of the CET's FTE Estimate to variations of -50% to +100% in a selected Workload Parameter.

Select the Parameter below to Test, or Exit:

- ☐ Ingest - Product Types Ingested
- ☐ Ingest - Products Ingested
- ☐ Ingest - Volume Ingested
- ☐ Ingest - External Interfaces
- ☐ Processing - Products Types Generated
- ☐ Processing - Products Generated
- ☐ Processing - Volume Generated
- ☐ Processing - Product Types Integrated
- ☐ Processing - Product Types QA'd
- ☐ Distribution - Products Types Distributed
- ☐ Distribution - Products Distributed
- ☐ Distribution - Volume Distributed

Exit Sensitivity Test

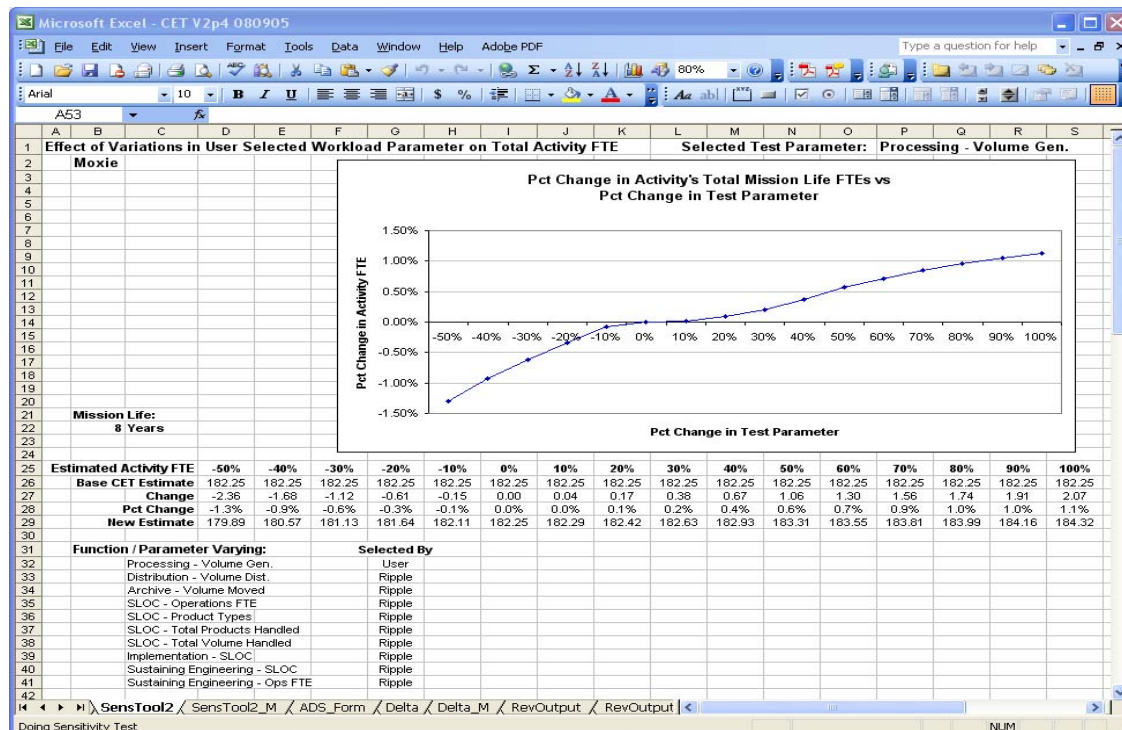
3.7.4 Example of Sensitivity Test Results

Figure 38 below presents an example of the results of a sensitivity test, a worksheet with an embedded graph. The upper left of the worksheet holds the sheet's title, below which appears the Activity Dataset's name, "Moxie" in this example. The upper right corner of the worksheet indicates which of the workload parameters you selected for this sensitivity test, 'Ingest - Products Ingested' in the example shown.

The graph plots the percent changes in total life cycle estimated FTE for the data activity (Y axis) for given percent changes in the workload parameter (X axis) in 10% steps through a range of -50% to +100%. The graph depicts the sensitivity of the total estimated FTE to variations in the selected workload parameter; i.e. the greater the percent change in the FTE to a pct change in the workload parameter, the greater the sensitivity of the FTE to the parameter. Note that while the X axis range will be fixed as described, the Y axis range will vary with the range of percent variation that is induced in the total estimated FTE. Therefore it will be important to check the Y axis range, since the shape of the curves will be similar but the Y magnitudes will vary according to the sensitivity of the total FTE to the workload parameter being plotted.

The table below the graph contains the values from which it was generated. The columns of the table contain data for each step of variation of the workload parameter. The first row of data is the base estimate of total FTE, which is constant. The second row is the change in total FTE resulting from the change in the workload parameter. The third row is the percent change in total FTE, i.e. the percent that the change in the second row is of the base value in the first row. The fourth row is the modified estimate of total FTE, i.e. the sum of the base value and the change. Below the table is a list of parameters. The first parameter in the list is the workload parameter selected for test by the user, "Processing Products Gen" in this example. To its right in the "Selected By" column is the word "User", indicating that the user selected this parameter. The other parameters in the list are those changed because of changes in the user selected parameter, the "ripple" effect mentioned above.

Figure 38 – Sample Sensitivity Test Results



3.8 Save Your Activity Dataset

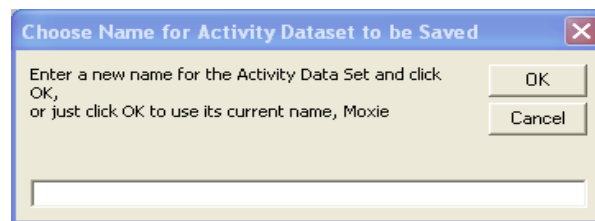
After a life cycle cost estimate has been produced, and you have printed one or both of the output spreadsheets or declined to do so, and have examined the quality report, plotted graphs and/or run a sensitivity test, the CET will check to see if the Activity Dataset is either new or a new modified version of an existing Activity Dataset. If it is a new one, you will be asked if you wish to save it, and for a name for it. If you used an existing Activity Dataset and made changes to it, and if you decide to save it, you will be able to either save it under a new name, keeping the previous version intact, or under its existing name, replacing the previous version. If you used an existing Activity Dataset without making any changes to it, it's already saved.

Note that if you exit the CET without saving a new or modified Activity Dataset it will be lost.

You are first presented a box titled “Save Activity Dataset” asking if you wish to save the current Activity Dataset (either a new one just created, or a previously saved one that you have reused or modified, and including the CET output just produced).

If you wish to save it, click “Yes”. You will see a box labeled “Choose Name for Activity Set to be Saved”, as shown in Figure 39 below:

Figure 39 - Choose Name for Activity Dataset to be Saved Box



You may either enter a new name and click “OK”, or just click “OK” leaving the name blank in order to use the existing name (e.g., “Moxie” in the example shown above) for the Activity Dataset. Note that the name of the Activity Dataset is case sensitive.

If you choose to save the Activity Dataset under a new name, the previous version of the Activity Set will remain in place under its old name.

If you choose to save the new version of the Activity Dataset under its existing name, the old version of the Activity Dataset will be deleted.

Note: If saving a new Activity Dataset you exit the CET and also want to exit Excel, in order to save any new Activity Datasets that you have created you must either do a “File Save” or “Save As” to save the now modified Excel workbook before you close it, or answer “Yes” when you close the workbook and Excel asks you if you want to save the changes you’ve made to it. If you do not, when you next start Excel you will find yourself working with the CET workbook as it existed when you last opened it.

3.9 Run a “What If” on Your Current Activity Dataset

You will next be asked if you wish to run a “What-If exercise on xxx, the current Activity Dataset”, i.e., if you would like to make a change or changes to the Activity Dataset you’ve just produced a life cycle estimate for, and to produce a new estimate to see the effect of the changes. The name (‘xxx’) of the current Activity Dataset will be shown – note that if you do a sequence of ‘what-if’s’ they will all be with reference to the original Activity Dataset even if you save one or more of the ‘what-if’ results under new names. If you click yes, you will be asked for the modifications you wish to make, as described in Section 3.4.

When the “What-If” is completed, after displaying the two output pages, the CET will display a “delta” page titled “What-If Differences” showing the differences between the “what-if” results and the original result, with increases highlighted in blue and decreases highlighted in red, and again the CET will offer you the option of printing or skipping the printing of this page. When you click either to print or skip printing, the CET will display the quality report. An example of a “what-if” differences report is presented in Table 4 below.

Note that when you run a “What If”, the CET re-runs its cost estimating model, including redoing the regression curve fitting process. Changes you make in workload parameters may cause the CET to select a different curve (e.g. exponential or logarithmic instead of linear). If that happens, the CET may generate a “What-If” estimate that contradicts expectations – for example a small decrease in estimated staffing in response to an increase in workload instead of the increase in staffing you would expect. This should not happen frequently, and when it does the overall estimate should remain within the CET’s error bars of 20% - 25%.

Table 4 - “What-If Differences” for the “Moxie” Activity Dataset

What-If Differences	Activity Dataset: Moxie				What-If #1				Produced: 09/04/08					
Mission Start Year:	2010	Operations Start Year:				2012	Mission Complete Year:				2017	Inflation Rate		4.0%
Estimated Staffing Level	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Total	
Management Staff FTE	0.02	0.02	0.22	0.22	0.22	0.22	0.22	0.22	0.00	0.00	0.00	0.00	1.36	
Administrative Support FTE	0.01	0.01	0.04	0.04	0.04	0.04	0.04	0.04	0.00	0.00	0.00	0.00	0.27	
Technical Coordination Staff FTE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Development Staff FTE	0.38	0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.76	
Technical / Science Staff FTE	0.00	0.00	1.43	1.43	1.43	1.43	1.43	1.43	0.00	0.00	0.00	0.00	8.60	
Operations Staff FTE	0.00	0.00	0.65	0.65	0.65	0.65	0.65	0.65	0.00	0.00	0.00	0.00	3.93	
Sustaining Engineering Staff FTE	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.02	0.00	0.00	0.00	0.00	0.13	
Engineering Support Staff FTE	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Estimated Total FTE	0.41	0.41	2.37	2.37	2.37	2.37	2.37	2.37	0.00	0.00	0.00	0.00	15.05	
Estimated Staff Costs, K\$	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Total	
Management Staff Cost	4	4	41	43	45	47	49	50	0	0	0	0	282	
Administrative Support Staff Cost	1	1	5	5	5	5	5	6	0	0	0	0	32	
Technical Coordination Staff Cost	0	0	0	0	0	0	0	0	0	0	0	0	0	
Development Staff Cost	57	60	0	0	0	0	0	0	0	0	0	0	117	
Technical / Science Staff Cost	0	0	438	479	522	557	590	621	0	0	0	0	3,207	
Operations Staff Cost	0	0	-98	-119	-141	-157	-172	-184	0	0	0	0	-871	
Sustaining Engineering Staff Cost	0	0	3	4	4	4	4	4	0	0	0	0	23	
Engineering Support Staff Cost	0	0	0	0	0	0	0	0	0	0	0	0	0	
Total Estimated Staff Cost	63	64	391	412	434	455	476	497	0	0	0	0	2,792	
Other Non-Staff Costs	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Total	
System Purchase Cost	7	5	0	0	0	0	0	0	0	0	0	0	12	
COTS Software License Cost	0	0	0	0	0	0	0	0	0	0	0	0	0	
Facility Preparation and Support Cost	10	10	36	36	35	36	36	36	0	0	0	0	235	
System Maintenance Cost	0	0	12	12	12	12	12	12	0	0	0	0	72	
Network/ Communications Cost	0	0	1	1	0	1	0	0	0	0	0	0	3	
General Supplies Cost	1	1	4	4	4	3	3	3	0	0	0	0	23	
Archive Media Cost	0	0	0	1	0	0	1	1	0	0	0	0	3	
Distribution Media Cost	0	0	0	0	0	0	0	0	0	0	0	0	0	
Travel Cost	0	0	0	0	0	0	0	0	0	0	0	0	0	
Training Cost	0	0	3	1	1	1	1	1	0	0	0	0	8	
Data Purchase Cost	0	0	0	0	0	0	0	0	0	0	0	0	0	
Computer Services Cost	0	0	0	0	0	0	0	0	0	0	0	0	0	
Total Estimated Non-Staff Costs, K\$	18	16	56	55	52	53	53	53	0	0	0	0	356	
Total Estimated Cost, K\$	81	80	447	467	486	508	529	550	0	0	0	0	3,148	

Table 4, Continued.

What-If Differences, Cont.			Activity Dataset: Moxie						What-If #1			Produced: 09/04/08		
Breakdown of Management, Technical / Science, and Operations Staffing for the Operating Functional Areas														
Mission Start Year:	2010		Operations Start Year:		2012	Mission Complete Year:			2017					
Estimated Staffing Level	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Total	
Management Staff FTE	0.02	0.02	0.22	0.22	0.22	0.22	0.22	0.22	0.00	0.00	0.00	0.00	1.36	
Activity Level Management	0.02	0.02	0.12	0.12	0.12	0.12	0.12	0.12	0.00	0.00	0.00	0.00	0.74	
Second Level Management	0.00	0.00	0.10	0.10	0.10	0.10	0.10	0.10	0.00	0.00	0.00	0.00	0.62	
Technical/ Science Staff FTE	0.00	0.00	1.43	1.43	1.43	1.43	1.43	1.43	0.00	0.00	0.00	0.00	8.60	
Ingest	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Processing	0.00	0.00	1.34	1.34	1.34	1.34	1.34	1.34	0.00	0.00	0.00	0.00	8.07	
Documentation	0.00	0.00	0.09	0.09	0.09	0.09	0.09	0.09	0.00	0.00	0.00	0.00	0.54	
Archive / Working Storage	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Access and Distribution	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
User Support	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Operations Staff FTE	0.00	0.00	0.65	0.65	0.65	0.65	0.65	0.65	0.00	0.00	0.00	0.00	3.93	
Ingest	0.00	0.00	0.19	0.19	0.19	0.19	0.19	0.19	0.00	0.00	0.00	0.00	1.14	
Processing	0.00	0.00	0.24	0.24	0.24	0.24	0.24	0.24	0.00	0.00	0.00	0.00	1.44	
Archive / Working Storage	0.00	0.00	0.22	0.22	0.22	0.22	0.22	0.22	0.00	0.00	0.00	0.00	1.34	
Access and Distribution	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
User Support	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

In the example above, four changes were made to the “Moxie” activity dataset (see Appendix C for details). A new ingest stream and a new processing stream were added, and two labor rate adjustments were made (technical science raised, operations lowered). The effects can be seen in Table 4. With the new streams, ingest and processing operations FTE increase a bit, and processing technical FTE increase a bit more (because for Moxie the product software and QA are in-house functions) and archive operations FTE increase a bit in response to the increased number of product types, products, and volume to be added to the archive. Development staff increases about 0.4 FTE per year. Other changes in management result from changes in “working” staff levels. Note that operations staff costs decrease as the assumed difference in labor rate (\$20K less per year) more than offsets the effect of the FTE increase.

Note that the “What-If Differences” page displays a “what-if” number. If you do another “what-if”, its “what-if” number will be incremented, and the differences will still be with reference to the original estimate. If you do a sequence of “what-if”s you can save one more of them as additional, separate Activity Data Sets under their own names. The “what-if”s will still always show changes from the original Activity Data Set that you began with.

As described above with respect to the output pages, the last “what-if” output worksheet will remain available to you after you exit the CET, but you must look at the version for your platform. The PC version of the what-if spreadsheet is named “Delta” and the Macintosh version is named “Delta_M”.

3.10 Review the CET’s Life Cycle Cost Estimate

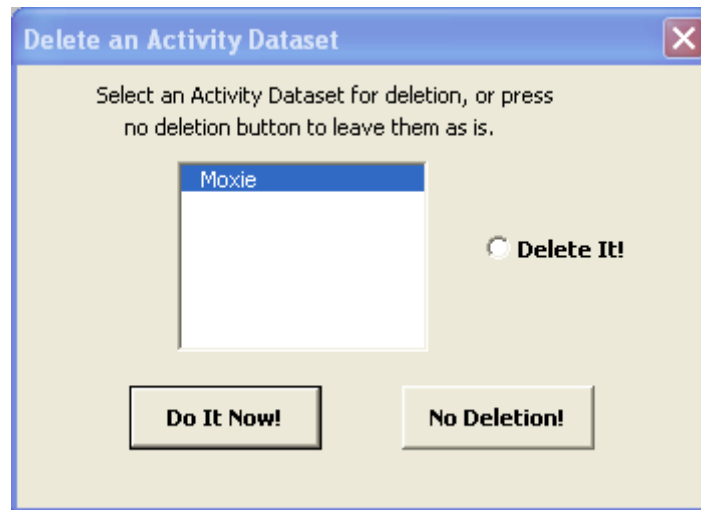
You will next be asked if you wish to review the CET’s life cycle cost estimate. If you wish to do a review, click “Yes”. See Section 3.13 for a description of the review process.

3.11 Delete No Longer Needed Activity Datasets

You will next be asked if you wish to delete any Activity Datasets. If you wish to delete one or more, or check the list of existing Activity Datasets to see if one or more are no longer needed, click “Yes”.

You will then see a box titled “Delete Activity Datasets”, which includes a list of the currently saved Activity Datasets, a “Delete It” option button, and “Do it Now” and “No Deletion!” buttons, as shown in Figure 40 below:

Figure 40 - Delete Activity Dataset Box



If you wish to delete an Activity Dataset, the highlight its name in the list (as was done for “Moxie” in the example shown above), select “Delete It” option, and click the “Do it Now” button. The CET will ask to confirm that the Activity Dataset (named in the box) is to be deleted; click “Yes” to proceed.

If you select “No Deletion!” the CET proceeds to the CET exit box described in the next section.

3.12 Exiting from the Cost Estimation Toolkit

When you are finished saving and/or deleting Activity Datasets, you will be asked if you wish to exit from the CET. Click “Yes”, to exit, and the CET will present a farewell box. It will delete the worksheet containing what was the active Activity Dataset (which you have previously either saved under its own name, or consigned to its fate).

Note: If saving a new or modified Activity Dataset, or deleting an Activity Dataset, you exit the CET and also want to exit Excel, in order to save any new Activity Datasets that you have created, or any deletions, you must either do a “File Save” or “Save As” to save the now modified Excel workbook before you close it, or answer “Yes” when you close the workbook and Excel asks you if you want to save the changes you’ve made to it. If you do not, when you next start Excel you will find yourself working with the CET workbook as it existed when you last opened it.

If you wish to do more work with the CET, click “No”. The CET will return to the title screen and beginning point as described in Section 3.1 above.

3.13 Reviewing the CET Life Cycle Cost Estimate

The CET’s Estimate Review Tool allows you to review, and as needed modify, the life cycle cost estimate produced by the CET to take into account planned technology re-use, the extent to which the new data activity is built on an existing institutional base, to adjust phasing of effort or expenditures, to smooth out year by year variations in effort caused by workload fluctuations (e.g. due to reprocessing), and other factors unique to your particular situation that are not taken into account by the general estimation capability provided by the CET (see Section 2.3.4 above for discussion of other possible considerations).

The CET’s review estimate tool is either started during the sequence of steps that follow after a CET estimate has been made, or can be started from the button on title page. In that case, the review tool will run on the most recent estimation output (resident in the Output worksheets).

The CET estimate review tool is not a duplication of the ‘what-if’ capability provided by the CET – the ‘what if’ function re-runs the CET effort / cost estimation model based on your revised inputs, while the CET estimate review tool simply allows you to review and manually modify the CET’s output. If in the course of your review you decide you want to change one or more of the mission or workload parameters or labor rates you will have to re-run the CET to produce a new estimate, which you can then again review.

3.13.1 CET Estimate Review Categories

The CET estimate review tool groups the lines that comprise the life cycle cost estimate into categories as shown in Tables 5 and 6 below. Each category, as shown in the tables, includes a number of FTE or cost lines from the estimate output tables produced by the CET (see Tables 1 and 2 above). You are given the opportunity to review and, as needed, modify each line within the selected category. The CET estimate review tool updates the estimate as you proceed, and when you have completed your review the tool displays and, if you wish, prints, your revised estimate and a “delta” worksheet showing the differences of your revised estimate from the CET’s original estimate. Note that you will be able to review effort (FTE) levels but not derived staff costs; if you wish to change your labor rates you will have to re-run the CET, modify you activity dataset and produce a revised estimate (which you may then review as needed).

Table 5 – CET Estimate Effort Categories and Lines Each Contains

CET Estimate FTE Category	Estimate FTE Lines
Development, Sustaining Engineering, Engineering Support	Development FTE
	Sustaining Engineering FTE
	Engineering Support FTE
Technical Coordination	Technical Coordination FTE
Ingest Operating Function	Ingest Operations FTE
	Ingest Technical FTE
Processing Operating Function	Processing Operations FTE
	Processing Technical FTE
Archive Operating Function	Archive Operations FTE
	Archive Technical FTE
Access and Distribution Operating Function	Distribution Operations FTE
	Distribution Technical FTE
User Support Operating Function	User Support Operations FTE
	User Support Technical FTE
Documentation	Documentation Technical FTE
Management	Activity Level Management FTE
	Second Level Management FTE
	Administrative Support FTE

Table 6 - CET Estimate Cost Categories and Lines Each Contains

CET Estimate Cost Category	Estimate Cost Lines
System Related Non-Staff Costs	System Purchase Cost
	System Maintenance Cost
	COTS Software License Cost
	Facility Preparation and Support Cost
	Network / Communications Cost
Activity Support Non-Staff Costs	General Supplies Cost
	Archive Media Cost
	Distribution Media Cost
	Travel Cost
	Training Cost
	Data Purchase Cost
	Computer Services Cost

3.13.2 Running the CET Estimate Reviewer

When you click on the “Run Cost Estimate Reviewer” button on the CET’s title sheet (see Figure 3 above) you will be presented with a menu of estimate categories (as outlined in Tables 5 and 6 above) from which you may select, and final radio button to tell the tool that you have completed your review / revision. See Figure 41 below. The CET Estimate Reviewer, when activated from the title page, works with the most recent estimate produced by the CET (which is resident in the Output worksheets).

The review process works category by category; you select a category, review the lines within that category, and return to the menu in Figure 41.

Figure 41 - Estimate Review Category Menu

Life Cycle Cost Estimate Reviewer - Menu

Select Portion of CET Estimate to Review and Revise, or Move On!

- ☐ Development and Sustaining Engineering, Engineering Support
- ☐ System Related Non-Staff Costs
- Operating Functions:
 - ☐ Ingest
 - ☐ Processing
 - ☐ Archive / Working Storage
 - ☐ Access and Distribution
 - ☐ User Support
 - ☐ Documentation
- ☐ Management / Administrative Support
- ☐ Technical Coordination
- ☐ Activity Support Non-Staff Costs
- ☐ Review Completed - Proceed to view and/or print Revised Estimate

Select a category and click on its radio button. You will immediately be presented with a screen inviting you to review the first estimate line within the selected category, the Review Estimate Line Box, which will be referred to as the Review Box. See Figure 42 below:

Figure 42 – Review Estimate Screen including Cost Estimate Line Review Box

The screenshot shows a Microsoft Excel window titled 'CET V2p4 080905'. The background worksheet displays a table of staff FTE data for various roles like Management Staff, Administrative Support, Technical Coordination, etc. Overlaid on this is a 'Cost Estimate Line Review Box' dialog. The dialog has a yellow background and contains the following elements:

- Estimate Line:** A dropdown menu set to 'Development FTE'.
- Operations Start Year:** A dropdown menu set to '2012'.
- Calendar Year:** A row of buttons for years from 2010 to 2017.
- Activity:** A row of buttons for 'Impl' and 'Ops'.
- CET Estimate:** A text box containing '8.7'.
- Revised Estimate:** A text box containing '8.7'.
- Total:** A text box showing '17.4'.
- Review Considerations - Implementation Period:** A text box with instructions about development FTE assumptions.
- Review Considerations - Operations Period:** A text box with instructions about operations FTE assumptions.
- Buttons:** 'Make No Changes', 'Accept Changes', and 'Revert to CET Estimate'.

As shown above, you will see the Review Box, and in the background you will see the revised estimate worksheet, positioned so that the portion of the CET life cycle cost estimate page that includes the estimate line appears above the Review Box. When you see this on your PC screen, the Development FTE line will be highlighted (even though it is not highlighted here).

The Review Box, starting from the top, identifies the estimate line up for review and the year in which your data activity's operations start.

Then the Review Box presents the CET's life cycle cost estimate in three lines. The first contains the calendar years spanning the life cycle of your data activity. The second describes the principal activity for each year, an initial period of implementation followed by a period of operations. The third contains the CET's estimated value for each year for the line up for review, either in FTE or in K\$ depending on the type of line. (As noted above, the staff cost lines are not reviewed.)

The CET Estimate line is followed by a Revised Estimate line, which initially repeats the values from the original CET estimate.

Below the Revised Estimate line are two text boxes. The first contains considerations applicable to the implementation period for you to think about as you review the estimate line. The second text box contains considerations applicable to the operations period. Of course, you may have other reasons than those shown for wishing to revise the CET's estimate for the line up for review.

Below the text boxes are three control buttons, labeled "Make No Changes", "Accept Changes", and "Revert to CET Estimate".

If you do not wish to make any changes to the estimate line up for review, click on the “Make No Changes” button. You will immediately be presented with the Review Box for the next estimate line for your review, or if that was the last line within the category you selected, you will be presented with the category menu (Figure 41 above).

If you wish to revise the CET estimate, enter your revisions directly into the Revised Estimate line. When you are finished, click on “Accept Changes”. You will see the CET output spreadsheet revised according to the changes you entered, with the line and annual totals of FTE (and the corresponding staff cost line and totals) or non-staff cost updated accordingly. You will be asked if you are happy with the revisions.

If you are happy with your revisions, click ‘yes’ to proceed to the next estimate line.

If you are not happy with your revisions, click ‘no’ to return to the Review Box. When you do, you will see your revisions as the “Revised Estimate”. You may decide that you like the changes after all, and click on “Make No Changes” to take them and go on. You may make further changes, and click on “Accept Changes”, and you will see the updated revised CET output worksheet as before. If you decide that you want to revert to the original CET estimate, click on “Revert to CET Estimate”, and you will see the revised CET output worksheet showing the original CET estimated values.

After you have reviewed the last estimate line within a category, the CET estimate review tool will display the estimate category menu (Figure 41 above). You may select another category to review, or by clicking on the last radio button you will see in sequence the revised cost estimate output pages and the “delta” page. This sequence is the same as you see at the conclusion of a CET run; you may view or print each of the pages.

After you have viewed and/or printed the revised estimate pages, the CET estimate review tool asks if you would like to save the revised estimate.

If you click ‘yes’, the tool will ask you for a name for the revised estimate. It will use the name you assign to identify three worksheets that will contain the revised estimate (by appending a page number to the name you assign). If there is already a revised estimate with the name you’ve given, the tool will ask if you wish to replace that revised estimate with the new revised estimate, or assign a different name for the new revised estimate.

After the estimate has been saved, or if you do not wish to save the estimate, the CET estimate review tool will ask if you want to delete any previously saved revised estimates. If you click ‘yes’, the tool will display a list of the revised estimates that you previously saved. You may highlight one and delete it. The tool will redisplay the list so that you can continue to delete revised estimates until you click on ‘No Deletion’.

The CET estimate review tool will then ask if you wish to exit the tool. If you click ‘yes’, the tool will terminate. If you click ‘no’, the tool will go back to its beginning and display the estimate category menu.

You can manually examine and print out any revised estimates you have saved.

Note that the original CET estimate pages are not changed by the CET’s estimate review tool.

3.14 Accessing and Using the Cost Estimation Toolkit Worksheets

When you open the CET Workbook, and before you run the CET program or after you have exited the CET program, the Excel workbook is open, allowing you access to the CET worksheets, both the Activity Dataset worksheets that you have created (and the sample Activity Dataset provided with the CET) and some the CET's internal worksheets.

The CET's internal worksheets are password protected. If you defeat the protection and make changes you may adversely affect the CET. The solution is to first have saved a 'clean' copy of the CET workbook as received, and then to revert to it, i.e. make yourself a new working copy of the CET.

3.14.1 CET Internal Worksheets

Table 7 describes the CET internal worksheets that you will see in the CET workbook, including any cautions appropriate for them. You can safely examine or print any of the internal worksheets used by the CET.

The internal worksheets are identified by the names that appear on the tabs associated with each worksheet in the CET workbook. All but two are PC and Macintosh versions of the CET output worksheets. The other two, ActiveADS (when present) and ADSForm, might be of interest to you. ActiveADS, when present, will contain the most recent Activity Dataset that the CET Estimator worked with, while ADSForm is the internal Activity Dataset template which includes the row definitions for the ADS worksheets generated by the CET when you enter the information describing your data activity.

If you have made yourself a working copy of the CET into which you've entered your Activity Datasets, and if a static CET internal worksheet gets changed or deleted, you may recover by opening an unmodified copy of the CET and copying the original worksheet to your working copy of the CET. Alternatively you could make a fresh working copy of the CET and copy your Activity Dataset worksheets to the fresh working copy.

Table 7 - CET Visible Internal Worksheets

Name	Description	Notes (all below are password protected)
ActiveADS	Current active Activity Dataset	Sheet and content created as needed by CET.
ADS Form	Blank Activity Dataset	Static, CET uses the row names, data must be empty.
Output	LCE Cost Estimate Output, pg 1, PC format	Content created by the CET Estimator tool when requested.
Output_M	LCE Cost Estimate Output, pg 1, Macintosh format	Content created by the CET Estimator tool when requested.
Output2	LCE Cost Estimate Output, pg 2, PC format	Content created by the CET Estimator tool when requested.
Output2_M	LCE Cost Estimate Output, pg 2, Macintosh format	Content created by the CET Estimator tool when requested.
RevOutput	Revised LCE Cost Estimate Output, pg 1, PC format	Content created by the CET Reviewer tool when requested.
RevOutput_M	Revised LCE Cost Estimate Output, pg 1, Macintosh format	Content created by the CET Reviewer tool when requested.
RevOutput2	Revised LCE Cost Estimate Output, pg 2, PC format	Content created by the CET Reviewer tool when requested.
RevOutput2_M	Revised LCE Cost Estimate Output, pg 2, Macintosh format	Content created by the CET Reviewer tool when requested.
Delta	“What-if” Difference Table, PC format	Content created by the CET Estimator tool and/or Reviewer tool when requested.
Delta_M	“What-if” Difference Table, Macintosh format	Content created by the CET Estimator tool and/or Reviewer tool when requested.
Quality	CET LCE Quality Report, PC format	Content created by the CET Estimator tool.
Quality_M	CET LCE Quality Report, Macintosh format	Content created by the CET Estimator tool.
Graphs	Graphs Worksheet	Content created by the CET Estimator tool when requested.
Sensitive	Intermediate Sensitivity Results	Content created by the CET Estimator tool when requested.
SensTool2	Sensitivity Results, PC Format	Content created by the CET Estimator tool when requested.
SensTool2_M	Sensitivity Results, Macintosh Format	Content created by the CET Estimator tool when requested.
SensToolTemp2	Sensitivity Results Template, PC Format	Static
SensToolTemp2_M	Sensitivity Results Template, Macintosh Format	Static
Title	The CET Title Page	Includes CET run buttons, identification
Background	Background page	Includes CET Logo and shows progress
Gray	Background page	Plain gray background used by Reviewer

A few key notes:

1) The Life Cycle Cost Output page worksheets, for original and reviewed/revised estimates are left with the results from the last CET Estimator or Reviewer run when you exit the CET. You may access and examine them or print them manually, copy them to another workbook or, for example, a Word document. The same applies to the Delta worksheet that contains the “what-if” or reviewer comparison information.

2) You can run the CET's Reviewer tool from the button on the title worksheet at any time, after at least one CET Estimator run – it will use whatever is in the CET's life cycle cost estimate pages (Output and Output2 or their Mac equivalents) as its input, i.e. as the estimate to be reviewed.

3) The Graphs worksheet will hold the last graph you generated when you are finished working with an activity data set. If you wish to copy a graph to, for example, a Word document, you can generate the graph, exit the CET, tab to the Graphs worksheet, select (click on) the graph and copy it to the Word document.

3.14.2 Activity Dataset Worksheets

You can, with care, inspect and make manual modifications to your Activity Dataset worksheets. When you do so you bypass the error and consistency checking that the CET does on entries made via the user data entry forms. For example, you may copy an Activity Dataset worksheet, give it a new name (i.e., rename the worksheet tab), and then make changes to the new worksheet, thus creating a new Activity Dataset. When you next run the CET, you will see your new Activity Dataset on the list of existing Activity Datasets that you can work with. You may also construct an Activity Datasets outside the CET, insert a blank worksheet in the CET workbook, copy the external ADS to the blank worksheet (or do an equivalent worksheet copy or move operation), and change the tab to the name of your new Activity Dataset.

Note that the year-by-year by request distribution data is computed by the CET from the control parameters you provide via the data entry box (see Section 3.3.6 and Figure 14 above). You can modify the year-by-year data manually, but if you only modify the control parameters the year-by-year values will not be changed. To have them updated by the CET you must activate the CET and modify the control parameters via the by-request distribution data entry box.

3.14.3 Saved Revised Life Cycle Cost Estimate Worksheets

When you run the CET Reviewer tool and elect to save the revised estimate, it is saved in three new worksheets. You will have assigned a name for the revised estimate, e.g. "Fudge". The revised estimate will be saved in three worksheets names "FudgePG1", "FudgePG2", and "FudgePG3", containing the content from "RevOutput", "RevOutput2", and "Delta" (or their Macintosh equivalents) respectively. You can inspect and print these worksheets manually as you wish. You can modify them, but note that the worksheets do not contain any row or column totaling formulas (the computations are done by VBA procedures when the revised estimates are generated).

Appendix A - CET DSP Reference Model - Functional Areas

This appendix describes the Data Service Provider (DSP) Reference Model, a functional model of a generic data service provider, and discusses the model's functional areas and their use in the CET. Section 2.1 above describes how the reference model serves as the underpinning for the the CET.

A.1 Data Service Provider Reference Model

The DSP reference model has three related aspects:

- 1) A set of 'functional areas' that collectively comprise the full range of functions that a data service provider might perform and the areas of cost that must be considered by the cost estimation by analogy model and embodied in the CET.
- 2) A set of parameters for each functional area that constitute a quantitative description of the workload, staff effort, and any other factors that contribute to cost for that area, additional 'roll-up' parameters that sum items such as staff effort across the functional areas, and other parameters like labor rates that are required for cost estimation. The parameter definitions are included in the CET Technical Description Document.
- 3) A set of requirements and levels of service for each functional area. See Appendix B, Activity Dataset Checklist, for reference to the levels of service as used by the CET.

These three aspects of the model are closely coupled to ensure the internal consistency of the model. The set of functional areas is the underpinning; the model parameters and requirements / levels of service are organized according to the functional areas. The requirements / levels of service and the model parameters are coupled in that the definitions of the requirements / levels of service embody model parameters. This integration of the three aspects of the model is intended to ensure that estimated costs produced by the CET are driven by and traceable to requirements to the fullest extent possible.

The intent of the descriptions of the functional areas in Section A.2 below and the corresponding requirements / levels of service (see Working Paper 5, "General Data Service Provider Reference Model - Requirements / Levels of Service") is to provide a reasonably full description of the abstract ESE data service provider, and to reflect the concerns expressed in the February, 2002, SEEDS community workshop and other community feedback. The ability of the cost estimation by analogy approach to reflect the full range of detail described in the functional areas and requirements / levels of service will be limited by the information available for the comparable data activities and the feasibility of reasonable assumptions where information is not available. This is reflected in the reference model's parameter set.

A.1.1 Functional Areas - Areas of Cost

The functional areas of the reference model are defined in Section A.2. Some of the areas are not strictly speaking "functional" in nature (such as 'facility / infrastructure') but are needed to ensure that all significant cost areas are included.

The functions / areas of cost span the full life cycle from implementation through operations. Implementation includes capital and staff costs associated with developing, implementing, integrating and testing the data service provider's data and information system, and facility start-up / preparation costs. Implementation is assumed to be spread over a specified number of years. Implementation can overlap the start of operations. Implementation can also recur during the operating period, allowing for system expansion, enhancement, or replacement, i.e. 'technology refresh'. Operations includes hardware maintenance, sustaining engineering, operations staff,

supplies (e.g. storage and archive media), recurring facility costs, etc., for the expected lifetime of the activity.

A.1.2 Reference Model Parameters

The parameters of the reference model are defined in detail in the CET Technical Description Document (and previously in Working Paper 4, “General Data Service Provider Reference Model - Model Parameters”).

The scope of the parameters spans implementation and operations, year by year over the specified lifecycle of the data service provider, and includes cost elements as well as workload factors and high level system configuration information. The implementation and operations parameters are broken down into outputs to be provided by the model, internal (derived) parameters used by the model, and inputs required by the model.

The cost estimation relationships (CER’s) used by the CET have been derived from information describing actual data centers or other data service providers comparable to future ESE data service providers. Raw information received from the data service providers has been mapped to the standard reference model parameter set to provide an internally consistent set of parameters for the CET’s estimating process.

The information describing comparable data activities is used to derive the cost estimating relationships that allow estimation of the outputs given the inputs for independent test cases, as described in Section 2.3 above.

A.1.3 Reference Model Requirements / Levels of Service

The requirements / levels of service of the reference model are presented in Working Paper 5, “General Data Service Provider Reference Model - Requirements / Levels of Service”.

The general data service provider reference model maps to a general requirements template, a statement of requirements / levels of service for a generic data service provider, in which the requirements / levels of service are defined for all of the functional areas included in the model.

The reference model includes two functional areas, Search and Order and Instrument / Mission Operations that are not included in the CET. The Search and Order area is not included because sufficient information for it distinct from Access and Distribution was not available for the comparable data activities. The Instrument / Mission Operations area was not included because too few comparable data activities perform those functions.

The requirements / levels of service are a template in that they contain placeholders for quantitative parameters that will be defined for a specific instance of a data service provider. For example, a requirement in the template might be that “the data service provider shall provide an archive capacity of [number TB]”. A data service provider of a type that would include providing an archive would have that item in its template. If the mission of the data service provider required that it archive certain data streams and generated products that would accumulate to a total volume of 100 TB, then that value would be inserted into the template, with the result being a specific requirement for that data service provider (i.e., “the data service provider shall provide an archive capacity of 100 TB”) that could then be used in the process of generating a cost estimate for the data service provider.

A.2 Functional Areas

This section describes the functional areas / areas of cost that comprise the Data Service Provider Reference Model and how they are used by the CET. They describe the full range of functions of an abstract general data service provider. It is unlikely that an actual ESE data service provider would perform in all of the functional areas; different ones would perform in different subsets of

the full set, and would perform at different levels (i.e. provide different levels of service) within functional areas. The CET allows the use to select functional areas and levels of service appropriate to his/her new data activity.

The functional areas are primarily focused on operating activities of the data service provider. The data service provider also has additional responsibilities that require high levels of expertise in science in the discipline(s) supported by the provider, data management expertise, and information technology expertise. The Management area (see Section A.2.13 below) includes lead, site-level responsibilities in these areas, and the Technical Coordination area (see Section A.2.11 below) includes coordination with other ESE data service providers and broader communities in these areas.

The intent is not to provide exhaustive descriptions in great detail of every possible aspect of each of the functional areas, but rather to describe key aspects of each that are significant cost drivers, or are of greatest concern to either users or data service provider operators or planners.

The following sections present working definitions of the functional areas that make up the data service provider reference model.

A.2.1 Ingest

The ingest functional area includes receiving, reading, quality checking, cataloging, of incoming data (including metadata, documentation, etc.) to the point of insertion into the archive. Ingest can be manual or electronic with manual steps involved in quality checking, etc.

Incoming data can be received from external sources or internally generated. Ingest can include format conversion, metadata extraction, or other preparation of incoming data for archive or use within the data service provider. Ingest includes verifying that all data made available for ingest has been successfully ingested, with exceptions tracked and accounted for. Ingest must be accomplished in a timely manner as needed to meet mission requirements of the data services provider.

Ingest is included in the CET, as described in Sections 2 and 3 above.

A.2.2 Processing

The processing functional area includes the generation and quality checking of new derived data products from data or products that have been ingested, or previously generated, generally on a routine, operational basis. Operational processing can be on demand as well as scheduled. Operationally generated products are often ‘standard products’ characterized by a peer reviewed, validated, reasonably stable, ‘science quality’ processing algorithm.

Processing includes ad hoc, non-operational generation of products that can include responding to requests for data mining or generation of special subsets. Processing includes process control (production planning, scheduling, monitoring, etc.) as well as product generation per se. Processing also includes reprocessing of new versions of previously generated products, either according to a reprocessing schedule or plan, or as allowed within a specified overall reprocessing capacity.

Where science or applications needs require simultaneous measurements from multiple instruments, processing performed by a data service provider can include data integration - mapping parameters from different sources to a common spatial / temporal base.

Processing can also include ‘data mining’, where software may search through many of the holdings of a data service provider for items meeting certain criteria.

The data service provider may receive the software that embodies product generation algorithms from outside developers (e.g. some Terra instrument teams for the DAACs currently) who are

responsible for the initial delivery and for delivering updated versions. Where quality, especially science quality, of products remains the responsibility of an outside developer, processing includes supporting quality checking by the science software developer. Support provided by the data service provider for integration and test of this ‘science software’ is included as an activity under processing. In cases where a data service provider develops algorithm software, that effort (i.e. development, integration, and test) is included under Implementation.

The data service provider may also accept software from science or applications users to produce a research product, perform data integration, or perform data mining.

Processing is included in the CET, as described in Sections 2 and 3 above.

A.2.3 Documentation

The documentation functional area includes the development (or upgrading of received) data and product documentation (including user guides, catalog interfaces, etc.) to meet any NASA adopted documentation standards, including catalog information (metadata), user guides, etc., through consultation with data providers, algorithm developers, flight projects, etc. Knowledge capture is a critical concern - the data service provider must be committed to pro-actively capture knowledge of instruments, calibration, processing history, etc., from its data sources (e.g. instrument teams).

NASA adopted documentation standards may include FGDC (Federal Geographic Data Committee) metadata standards, documentation standards for long term archiving, Algorithm Theoretical Basis Documents (or equivalent, which must reflect ‘as-built’ algorithms), Data Software Interface Specifications, etc. When science needs require that multiple versions of a product be held, the documentation of each version must include the provenance information (e.g. processing algorithm) peculiar to it.

Documentation should include comments received from users on their experience with the data and products (product accuracy, usability, etc.), perhaps in the form of FAQ’s (Frequently Asked Questions) for products, both from scientists on staff or working closely with the provider and from the general user community.

Documentation should include read software and other appropriate tools for data access kept current with commonly available technology. Documentation includes maintenance and refresh according to best industry practice or NASA policy.

Documentation needs will evolve, e.g. information relevant to intellectual property rights may be needed.

Documentation is included in the CET, as described in Sections 2 and 3 above.

A.2.4 Archive

The archive functional area includes the insertion of data into archive storage, and data stewardship - management, handling and preservation of data, metadata, and documentation within a data service provider’s archive. Inserted data can include data ingested from sources external to the site, or data/products generated on-site.

Data stewardship / preservation includes quality screening of data entering and exiting the archive, quality screening of archive media, tested and verified backup and restoration capability, and accomplishing migrations from one type of media to another.

Insertion into the archive can be electronic or manual (e.g. hanging tapes on a rack or popping them into a robotic silo).

Archive is included in the CET, as described in Sections 2 and 3 above.

A.2.5 Search and Order

The search and order functional area includes providing access to catalog information (a range of descriptive information to aid in selecting data and products) and a search and order capability to users, and receiving user requests for data.

“Search and order” in this context is used in a very broad sense; search and order includes support for system to system interactions as well as conventional search and order by users directly. For example, system to system interactions might include a program running on a user platform accessing the data service provider system directly, locating a needed product, and executing a protocol (e.g. for user registration, security) to gain access to it.

“Search”, whether by a user directly or through a system-system interaction, implies applying criteria that might include geophysical parameter(s), spatial-temporal coverage, specific product names, etc., to the metadata describing available data and products and returning to the user listings supplemented by descriptive information of those data or product types and instances that meet the criteria.

“Order” implies a request/permission step, regardless of how implemented (e.g. manual or automated), where a request for a set of data or product instances, perhaps the results of (or a selected subset of the results of) a search, is processed and accepted or denied.

Search and order can include providing local user interface and capability and/or providing an interface to a broader based, cross-site search and order capability (e.g. DAACs supporting search and order via the EOS Data Gateway).

Search and Order is not explicitly addressed by the CET since it was not possible to clearly separate or distinguish data activities’ effort and workload for Search and Order from Access and Distribution.

A.2.6 Access and Distribution

The access and distribution functional area includes fetching the requested data from the archive, performing any subsetting, resampling, reformatting / format conversion (e.g. to a GIS (Geographic Information System) format), reprojection, or packaging, and providing the end product to the user by electronic means or on physical media.

“Access” is included to embrace a service allowing a program running on a user platform to access data and products from the data service provider directly, through an appropriate protocol, perhaps as a seamless extension of the system to system search and order described above.

Access and distribution can be performed on an operational basis, meaning in part that a data service provider will formally commit to terms of service in a level of service agreement or equivalent.

Access and distribution is an area likely to see substantial evolution in the next five to ten years, perhaps especially if distributed computing comes into play on a significant scale. Highly automated access techniques, software agents, and new tools for data discovery, access, integration from multiple sources, etc., will become available.

Note: Success from a user point of view may be even more dependent on a product’s format than the speed of its delivery (what if a product is delivered in 30 seconds but the form is such that a user needs to spend several hours to be able to use it, versus a product delivered in 30 minutes but in a form that can be used directly?). The data service provider should take care to offer formats (whether as a default or an option) that are directly useable by the largest possible fraction of its user community.

Access and Distribution is included in the CET, as described in Sections 2 and 3 above.

A.2.7 User Support

The user support functional area includes support provided in direct contact with users by user support staff, including responding to queries, taking of orders, staffing a help desk (i.e., staff awaiting user contacts who can assist in ordering, track and status pending requests, resolve problems, etc.), etc. User support staff includes science expertise to assist users in selecting and using data and products.

The demands on user support will increase with the proliferation of data types, data sources, and tools for users, continuing or increasing the need for highly trained user support staff even as user interactions become more automated and more automated user support aids become available (beginning with on-line documentation, FAQ, etc.).

User support also includes outreach to potential new users and education / training for current or potential new users.

User support should also be a channel for feedback from the users to the data service provider, whether comments on particular data or products or on the provider's services and support.

User support includes coordination of user support guidelines and practices across the network of ESE data service providers and with other data centers as needed to support the ESE science and applications program - see Technical Coordination, Section A.2 below.

Access and Distribution is included in the CET, as described in Sections 2 and 3 above.

A.2.8 Instrument / Mission Operations

The instrument / mission operations functional area includes monitoring instrument and spacecraft performance, generating instrument and spacecraft commands, and event scheduling (using NASA or other appropriate operational mission management services).

The CET does not address Instrument / Mission operations because very limited information for only a small number of data activities was available.

A.2.9 Sustaining Engineering

Sustaining engineering includes maintenance and enhancement of custom applications software (including any science software embodying processing algorithms developed by the site).

Sustaining Engineering is included in the CET, as described in Sections 2 and 3 above.

Sustaining engineering includes all development after the initial implementation period. This can be a limitation in the case of a new data activity with a lengthy planned operating life that is intended to include a substantial or complete overhaul / refresh of its data system elements; the effort and cost of such a refresh can not be estimated by the CET. There is an approach to producing a life cycle that does include a major system redevelopment some years after a data activity becomes operational. That approach is to divide the data activity life cycle into two segments and make separate estimates for the two time periods involved. For example, if a new data activity will have a ten year mission life with a major redevelopment in the sixth year, an estimate could be produced that would span the initial implementation and operation over the first five mission years, and then a second, separate estimate could be produced for the sixth through tenth years that would begin with the second round of implementation.

A.2.10 Engineering Support

Engineering support includes some or all of the following as applicable at a particular site: systems engineering, test engineering, configuration management, coordination of hardware maintenance by vendors, COTS procurement, installation of COTS upgrades, system administration, database administration, network/communications engineering, and security.

Engineering support is internal, directed toward the internal operation of the data service provider.

Engineering Support is included in the CET, as described in Sections 2 and 3 above.

A.2.11 Technical Coordination

Technical coordination includes participation with other data service providers in NASA coordinating processes, including coordination on data management, data stewardship (including standards for content of life cycle data management plans), standards and best practices (including quality assurance standards and practices), interfaces, common metrics, and interoperability (e.g. for data access and integration).

This area includes coordination on evolution of the overall NASA Earth science data service architecture (including an examination of the changing needs of the NASA Earth science and applications program and the consequent impacts on the roles, missions, and services of data service providers).

Technical coordination includes coordinating user support guidelines and practices across the network of NASA Earth science data service providers and with other data centers as needed to support the NASA Earth science and applications program.

Technical coordination includes participation in processes to coordinate production and delivery of products between data service providers.

Technical coordination includes cooperating with other NASA data service providers in representing NASA in broader community processes in areas such as standards, interoperability, data management, security, etc.

Technical coordination, which by its nature includes engineering, is directed outward, supporting the data service provider as one element of a system of cooperating centers.

Technical Coordination is included in the CET, as described in Sections 2 and 3 above.

A.2.12 Implementation

Implementation includes development of, and making operational, the data and information system capabilities required by the data service provider to perform its mission, including design and implementation of the data system (hardware and system software) and applications software. Implementation can recur during the operating period as systems are expanded or replaced.

In addition to a major implementation effort, implementation can include ongoing applications software development. Implementation can include development of software tools for use by users to unpack, subset, or otherwise manipulate products provided by the data service provider.

In some cases applications software will include product generation software embodying science algorithms, e.g. to produce a product to meet a particular user need. Applications software can include software to perform a 'data mining' or data integration operation to meet a user need.

Implementation is included in the CET, as described in Sections 2 and 3 above. The CET user defines an 'implementation period' over which the implementation effort and cost are estimated. Continuing implementation after the implementation period is covered by the CET under sustaining engineering (see Section A.2.9 above).

A.2.13 Management

Management includes management and administration at the data service provider level ("front office") and direct management of functional areas. Management also includes staff with overall

responsibility for internal and external science activities, information technology planning, and data stewardship.

Management includes planning information technology upgrades / technology refreshes, based on assessments of changing mission or user needs and availability of new technology.

Management includes developing data stewardship practices, performing data administration with science advice (via the User Advisory Group and other appropriate bodies), developing and maintaining life cycle data management plans (which address data migrations).

Management also includes coordinating the science activities within the data service provider and its interaction with the NASA Earth Science and broader science community, including a visiting scientist program, collaboration among NASA Earth Science data service providers to support science needs, annual Enterprise peer review, and support for its User Advisory Group (which includes representation from the science, applications, education, etc., communities as appropriate for a given data services provider) and any other ESE or broader advisory activities that may be appropriate.

Management also includes participation in NASA management processes, strategic planning, coordination with other NASA or NASA-sponsored data centers and activities beyond NASA (e.g. NOAA or USGS data centers).

Management includes performing supervisory, financial administration, and other administrative functions.

Management is included in the CET as described in Sections 2 and 3 above.

A.2.14 Facility / Infrastructure

Facility / Infrastructure includes provision and maintenance of a fully furnished and equipped, environmentally controlled, physically secure facility to house data service provider staff, systems, and data and information holdings, including a backup facility for its data and information holdings. An off-site backup facility would be one sufficiently removed from the data service provider's primary site such that a fire, tornado, or other event that destroys the primary site would be very to extremely unlikely to also destroy the backup site (a risk analysis would be performed on a site by site basis).

This area includes resource planning, logistics, supplies inventory and acquisition, and facility management.

This area includes maintenance of system and site security according to established NASA security policies and practices.

Facility / Infrastructure also includes a variety of non-staff cost factors such as supplies, facility lease and utility costs and other similar overhead costs, hardware maintenance, COTS licenses, etc.

The CET includes Facility / Infrastructure as described in Sections 2 and 3 above. The effort associated with resource planning, logistics, supplies inventory and acquisition, and facility management is included by the CET in Engineering Support; information gathered from comparable activities did not allow a finer separation.

Appendix B - Activity Dataset Checklist

The following is a checklist of the information about your data activity that you need to build an Activity Dataset. The checklist is intended as an aid to help you prepare for the actual entry of information.

The required parameters required for each functional area are described below, with valid values given for items such as levels of service. You may omit any operational functional area not applicable to your activity. In all cases, entry of a 0 (zero) indicates that a parameter is not applicable to your activity or 'none of the above' when options are provided.

B.1 Data Activity Level ADS Information

This section describes the information for the Activity Dataset that pertains to your new data activity as a whole.

The Activity Dataset name is the name you wish to give to this Activity Dataset - you may create and store multiple activity datasets to help you evaluate different options, do trade-offs of requirements vs estimated effort levels and costs, etc.

Table B1 lists the operating functional areas, from which you may select all or those which apply to your new data activity. The CET box for selection of functional areas for your activity's ADS is shown in Figure 6, Section 3.3.1 above.

Table B1 - ADS Operating Functional Areas

Ingest	Receipt and capture of incoming data, verifying data integrity, possibly including a format conversion.
Processing	Generation of derived products from input data, either operationally (scheduled or on-demand) or on an ad hoc basis.
Documentation	Ensuring documentation to standards of products generated or received, cataloged, archived, and distributed.
Archive	Storage and preservation of data and product holdings, ranging from long term to temporary working storage.
Access and Distribution	Distributing data to users by network or media, perhaps including fully automated system-to-system access.
User Support	Providing assistance to users in understanding data, products, services, etc., in selecting and using data and products.

Table B2 below lists the parameters that characterize the key milestones in the life cycle of your new data activity. Note that the Implementation Period may overlap the operations period, i.e. the year of operations start can occur during the implementation period. Implementation may continue during operations, for the mission life, perhaps including a major system refresh, as appropriate. The CET can not perform an estimate for a major system refresh or redevelopment; see Section 2.2.3 A for a way this can be accomplished. The CET box for selection of milestone parameters for your activity's ADS is shown in Figure 7, Section 3.3.1 above.

Table B2 - Data Activity Life Cycle Milestones for ADS

Mission / Project Start Calendar Year	Calendar year (e.g. 2005) when your activity begins implementation.
Implementation Period, Years	The number of years from the beginning of implementation (see above). to the end of initial implementation. Implementation may continue during operations, for the mission life, perhaps including a major system refresh, as appropriate.
Calendar Year of Operations Start	The calendar year when operations will first begin, which may be prior to when full operational capability is reached.
Duration of Operations, Years	The number of years that your activity will be in operation, or the span of operating years for which a cost estimate is needed.

Table B3 describes the data activity labor rate information you will need to enter into the CET. These are fully loaded labor rates you estimate would apply to your activity, in your location, in \$K/Year, as of the Implementation Start CY above. "Fully loaded" includes all overheads applied to base labor rates by your activity or host institution. The CET box for selection of labor rates for your activity's ADS are shown in Figure 7, Section 3.3.1 above.

Table B3 - Data Activity Labor Rates for the ADS

Labor Rates:	
Management Staff Labor Rate	An approximate overall rate for management and activity level coordination (e.g. project manager, project system engineer, project scientist).
Administrative Support Labor Rate	An approximate overall rate for administrative support staff.
Development / Engineering Staff Labor Rate	An approximate overall rate for software developers / maintainers and related engineering, engineering support including system engineering, system administration, database administration, resource planning, network management, facility support, etc.
Technical / Science Staff Labor Rate	An approximate overall rate for technical and science staff, involved with ESE coordination, activity level coordination, and associated with operational functional areas.
Operations Staff Labor Rate	An approximate overall rate for operational level staff in operational functional areas.
Inflation Rate	Inflation rate to be applied to labor costs, assuming Mission / Project Start Calendar Year as base year.

B.2 ADS Information for Operating Functional Areas

This section addresses the ADS parameters needed by the CET to describe the ingest, processing, documentation, archive, access and distribution, and user support functional areas. In each case tables will address all of the parameters needed, with choices for levels of service defined, and any notes on individual parameters as needed.

You will notice that some of the information items will be marked as "Advisory – No Effect on Estimate". These are items that may have significant effort associated with them, but for which the CET has no basis for making an estimate of that effort. An example is "Science Software Service Options" which allows you to characterize the types of software (e.g. research or operational product generation, data integration, data mining) that your activity will accept and

run for users. What you select can have a significant effect on your staffing, but the CET has no basis for making an estimate of that effect. Although your entry for “Science Software Service Option” will not have any effect on the estimate, it serves as a reminder for you, and you can use the CET Reviewer tool (described in Section 3.10 above) to adjust the CET’s estimate to be commensurate with your plans. The Reviewer tool will remind you of this.

B.2.1 Ingest ADS Information

Table B4 describes the information needed by the CET for the Ingest functional area. You need to describe each ingest stream that your activity will receive and ingest. Each ingest stream can consist of a single product (or data) type, or a group or aggregate of product types that are received from a common source. The CET box for selection of ingest stream parameters for your activity’s ADS is shown in Figure 8, Section 3.3.2 above.

Table B4 - Ingest Functional Area ADS Parameters

Name of Ingest Stream	The Ingest Stream Name is the name of a product type or name assigned to an aggregate (group) of product types that are being treated as a single ingest stream for the purpose of the estimate. Members of the aggregate should have the items below (after the count of product types in the aggregate) in common, except for product counts and volumes that are summed over the product types included in the aggregate.
Number of Product Types in Ingest Stream	The number of different product types included in this ingest stream (aggregate or group).
Calendar Year - Ingest Start	Calendar year (e.g. four digits, 2005) when ingest of this stream begins, or 0 to indicate operations start year.
Calendar Year - Ingest End	Calendar year when ingest of this stream ends, or 0 to indicate last overall operations end year.
Source	Name of the source of the ingest stream (case sensitive - the CET will count distinct sources as external interfaces)
How Received	How the ingest stream is received: 1 - electronic (by network), or 2 - by media.
Level of Service	Level of service associated with this ingest stream: 1 - Time critical ingest with immediate verification of data integrity; 2 - Routine ingest and verification without tight time constraints; 3 - Ad-hoc or intermittent ingest with verification; 4 - Ad-hoc, non-operational ingest.
Retention Period, Years	The number of years that the products from this ingest stream will be retained in the archive, or 0 if indefinite.
Automation LOS	Level of service associated with this ingest stream: 1 - Routine or frequent manual action or intervention in ingest process; 2 - Occasional manual action or intervention in ingest process; 3 - Rare manual action or intervention, e.g. “lights out”.
Products of Type/Agg Ingested / Day	The number of individual products ingested per day in this ingest stream, summed over product types included.
Volume of Type/Agg Ingested / Day	The total daily volume in GB for all products received in this ingest stream, summed over product types included.

B.2.2 Processing ADS Information

Table B5 below describes the information needed by the CET for the Processing functional area. The table is in two parts. Part A, “Overall”, contains the parameters that apply to the Processing area as a whole. The CET box for selection of overall processing parameters for your activity’s ADS is shown in Figure 9, Section 3.3.3 above.

Part B of Table B5, “For Each Processing Stream” contains information you will be asked for describing each processing stream that your activity will produce. Each processing stream can consist of a single product (or data) type, or a group or aggregate of product types that are produced together. The CET box for selection of processing stream parameters for your activity’s ADS is shown in Figure 10, Section 3.3.3 above.

Table B5 - Processing Functional Area ADS Parameters

Part A - Overall Parameters:	
Science Software Service Options (Advisory – No Effect on Estimate)	Indicate how your activity will work with outside sources (e.g. instrument teams or research / applications users needing custom products) of applications software: 1 - Your activity will accept operational product, research product, data integration / data mining (etc.) software; 2 - Your activity will accept research product, data integration / data mining (etc.) software; 3 - Your activity will accept both operational product or research product software; 4 - Your activity will accept only research product software; 5 - Your activity will accept only operational product software; Note: 'data mining (etc.) software' can include intelligent search agents, etc.
Calibration Validation LOS	Enter a 1 if your activity will make a special effort in calibration - validation, perhaps including cross-calibration of different data types, otherwise enter a 0.

Part B - For Each Processing Stream:	
Name of Processing Stream	The name of a product type or name assigned to an aggregate (group) of product types that are being treated as a single processing stream for the purpose of the estimate. Members of the aggregate should have the items below (after the count of product types in the aggregate) in common, except for product counts and product volumes that are summed over the product types included in the aggregate or stream.
Number of Product Types in Processing Stream	The number of different product types included in this processing stream (aggregate or group).
Calendar Year - Production Start	Calendar year (e.g. four digits, 2005) when production of this stream begins, or 0 to indicate operations start year.
Calendar Year - Production End	Calendar year when production of this product stream ends, or 0 to indicate last overall operations end year.
Science Algorithm Software Source	1 if the algorithm software for this product stream will be developed in-house, 2 if it will be received from an outside source.
Product Science Quality Assurance	Enter 1 if science quality assurance for this product stream will be performed in-house, 2 if by an outside group

Part B - For Each Processing Stream, Continued:	
Processing Automation LOS	Level of service associated with this processing stream: 1 - Routine or frequent manual action or intervention in production process; 2 - Occasional manual action or intervention in production process; 3 - Rare manual action or intervention, e.g. "lights out".
Production Mode	The production mode for this stream, either 1 if operational, or 2 if not operational.
Operational Production Mode, If Applicable	If this processing stream will be operational, enter 1 if production is on a routine (time or event driven) schedule, or 2 if production is on-demand.
Operational Processing LOS, If Applicable	If this processing stream will be operational, enter a nominal production cycle time: (Where production cycle time for a typical or nominal granule runs from receipt of inputs to completion of production.) 1 if 2 days or less; 2 if about 7 days; 3 if about 30 days or longer.
Non-Operational Production LOS, If Applicable	If this processing stream will be non-operational, enter a level of service indicator: 1 if products will be produced according to specific completion targets; 2 if products will be produced according to general goals; 3 if there will be no production goals (e.g., in a research environment when a new product is being experimented with).
Product Retention Period (each version)	The number of years that products (each version) from this product stream will be retained in the archive, or 0 if indefinite.
Products Generated per Day	The number of individual products generated per day in this processing stream, summed over product types included.
Volume Generated, GB/Day	The total daily volume in GB for all products generated in this processing stream, summed over product types included.
Reprocessing LOS	A level of service for reprocessing for this product stream: 1 - Reprocessing will be scheduled and accomplished at the interval given by the Reprocessing Plan parameter; 2 - Reprocessing will be accomplished according to general goals 3 - Reprocessing will be performed on a time available basis, i.e. no extra capacity, no plan. 4 - No Reprocessing.
Reprocessing Interval	The nominal number of years between reprocessing expected for this product stream.

B.2.3 Documentation ADS Information

Table B6 below contains the parameters needed by the CET for the Documentation functional area. The CET box for selection of documentation parameters for your activity's ADS is shown in Figure 11, Section 3.3.4 above.

Table B6 - Documentation Functional Area ADS Parameters

Documentation LOS	<p>Characterize the overall documentation practice planned for your activity:</p> <ul style="list-style-type: none"> 1 - Data and product holdings will be documented to the standard for long term archive (i.e., long term science use); 2 - Data and product holdings will be documented to a level suitable for current science use. 3 - Data and products will include documentation as received from their provider or developer.
User Comment LOS	<p>Characterize the approach your activity will take to including user comments in documentation:</p> <ul style="list-style-type: none"> 1 - Your activity will routinely update data and product documentation with user comments. 2 - Your activity will occasionally update data and product documentation with user comments. 3 - Your activity will rarely (or never) update data and product documentation with user comments.

B.2.4 Archive ADS Information

Table B7 below contains the parameters needed by the CET for the Archive functional area. The CET box for selection of archive parameters for your activity's ADS is shown in Figure 12, Section 3.3.5 above.

Table B7 - Archive Functional Area ADS Parameters

Archive Purpose	<p>Characterize the purpose of your activity's data storage:</p> <ul style="list-style-type: none"> 1 - Your activity will store and preserve data and products indefinitely, as a long term archive. 2 - Your activity will store and preserve data and products for many years prior to transfer to a long term archive. 3 - Your activity will store data on a temporary, short term basis, prior to transfer to another activity or distribution ("working storage").
Archive Media Unit Capacity, GB	Approximate capacity of each tape or disk, GB. You may accept the default value of 50 GB or enter another value.
Archive Media Unit Price, \$	Enter the price of a single unit (e.g. tape or disk) of the archive media you plan to use, or accept the default price of \$100. You may enter a decimal number, i.e. 50.75 for \$50.75.
Archive Backup Fraction, Pct	Enter the approximate fraction of your primary archive that will be included in the backup, whether the backup is off-site or on-site, e.g. 50 = 50%.
Archive Media Re-Use Plan	<p>Indicate your activity's plan for re-using archive media freed up when data / products are deleted from the archive:</p> <ul style="list-style-type: none"> 0 – Do not re-use archive media (e.g. if media is not rewritable); 1 - Always re-use archive media (e.g. when reprocessing and deleting old products to replace with new reprocessed versions); 2 to N – Plan to re-use archive media but do not re-use archive media that is 2 to N years old; choose N based on the expected useful life of the media.

Table B7 - Archive Functional Area ADS Parameters, Con t.	
Archive Monitoring LOS	Characterize your activity's plans for monitoring the quality / integrity of its archive: 1 - 10% per year random screening of archive media (in addition to tracking normal operations); 2 - 5 % per year random screening; 3 - 1% per year random screening; 0 - No random screening.
Archive Entry/Exit Screening LOS (Advisory – No Effect on Estimate)	Characterize your activity's plans for monitoring its archive in the course of operations: 1 - Entry and exit screening (i.e. read after write when data is archived, verification of read when data is accessed); 2 - Entry screening only. 0 - No entry or exit screening.
Pre-Existing Archive Contents	The number of product types, number of products (of all types), and volume, separately for inactive (i.e. those no longer being ingested) product types and active product types (i.e. those that will continue to be ingested).

B.2.5 Access and Distribution ADS Information

Table B8 below contains the parameters needed by the CET for the Access and Distribution functional area. The access and distribution functional area includes two modes of distribution, operational and by request. Operational distribution includes delivery of defined sets of products (a distribution stream) to a specific user destination for more than a year. By request distribution includes delivery of products to a general community of users on a request basis, including making products available for access via web as well as FTP or other network delivery or via media.

Table B8 is divided into three parts. Part A contains parameters that apply to the access and distribution functional area as a whole. The CET box for selection of overall access and distribution parameters for your activity's ADS is shown in Figure 13, Section 3.3.6 above.

Part B of Table B8 contains the parameters that describe each operational distribution stream that your activity will distribute. Each distribution stream can consist of a single product (or data) type, or a group or aggregate of product types that are distributed operationally to the same destination. Members of the aggregate should have the items below (after the count of product types in the aggregate) in common, except for product counts and product volumes that are summed over the product types included in the aggregate. The CET box for selection of operational distribution stream parameters for your activity's ADS is shown in Figure 14, Section 3.3.6 above.

Part C of Table B8 contains the parameters that describe the by-request distribution service to be provided by your new data activity. The CET box for selection of by-request distribution parameters for your activity's ADS is shown in Figure 13, Section 3.3.6 above.

Table B8, Access and Distribution Functional Area ADS Parameters, Part A

Part A - Overall Distribution Parameters:	
Distribution Scope	<p>Indicate the planned availability of your activity's access and distribution capability, however implemented:</p> <p>1 - Your activity's access and distribution capability will be public, available to any and all users;</p> <p>2 - Your activity's access and distribution capability will be available to the science / applications communities you will serve;</p> <p>3 - Your activity's access and distribution capability will be available to a limited team of scientists or applications specialists (e.g., a flight project data system that supports a project PI team and hands its data off to another activity for archive and general distribution).</p>
Distribution Service Mode	<p>Characterize the access and distribution services to be offered by your activity:</p> <p>1 - Your activity will offer operational distribution, a subscription (standing order) service, and by request distribution;</p> <p>2 - Your activity will offer a subscription (standing order) service, and by request distribution;</p> <p>3 - Your activity will only distribute data in response to specific user requests (i.e. only by request).</p> <p>4 - Your activity will only offer operational distribution - no by request distribution.</p>
Distribution Means LOS	<p>Enter the level of service appropriate for your data activity:</p> <p>1 - Distribution is all electronic, i.e. only by network;</p> <p>2 - Distribution is predominantly by network (more than 50% of total volume distributed);</p> <p>3 - Distribution is predominantly by media (more than 50% of total volume distributed).</p>
Supporting Data Services LOS (Advisory – No Effect on Estimate)	<p>Indicate your activity's plans to offer supporting data services (e.g. subsetting, reformatting, remapping):</p> <p>1 - Your activity will provide appropriate supporting services for most of its data and products;</p> <p>2 - Your activity will provide appropriate supporting services for less than half of its data and products;</p> <p>3 - Your activity will provide appropriate supporting services for a few selected data and products;</p>

Table B8, Access and Distribution Functional Area ADS Parameters, Part B.

Part B - Operational Distribution:	
Name of Distribution Stream	The name of a product type or name assigned to an aggregate (group) of product types that are being treated as a single distribution stream for the purpose of the estimate.
Count of Product Types in Distribution Stream	The number of different product types included in this distribution stream (aggregate or group).
Destination for Distribution Stream	The name of the activity or organization that is the destination for this distribution stream.
Distribution mode	Enter 1 for distribution of this stream by network, 2 for distribution of this stream by media.
Distribution Stream Start Calendar yr	The calendar year (4 digit, e.g. 2005) in which this distribution stream becomes active.

Part B - Operational Distribution, Continued:	
Distribution Stream End Calendar yr	The calendar year (4 digit, e.g. 2005) in which this distribution stream ends.
Distribution Stream Products per day	The number of individual products distributed per day in this distribution stream, summed over product types included.
Distribution Stream Volume per day, GB	The total daily volume in GB for all products distributed in this distribution stream, summed over product types included.

Table B8, Access and Distribution Functional Area ADS Parameters, Part C.

Part C - By Request Distribution:	
By Request Distribution Mode	0 – No By Request Distribution; 1 – Ramp Up over Mission Life; 2 – Level Profile, Constant over Mission Life.
Expected number of users when services are well established	The expected number of active by request users in a peak year when services are well established.
Expected Average Number of Requests per User, per Year	The estimated approximate average number of requests per year by each by-request user.
Expected Average Number of Products per Request	The estimated approximate average number of products expected per request by by-request users.
Out of every ten products distributed, estimated number (0-10) that go by network	Out of every ten products distributed, how many are expected to be distributed by network?
Out of ten units of volume distributed, estimated number (0-10) that go by network	Out of every ten GB of data/products distributed, how many GB are expected to be distributed by network?
Response time for requests delivered by network	Characterize the nominal or average response times for network delivery anticipated for your activity (i.e. time until delivery of the first item in a request begins): 1 - about 10 seconds or less for system-to-system access; 2 - about 10 seconds or less for FTP or equivalent; 3 - about 10 minutes for FTP or equivalent; 4 - about 24 hours for FTP or equivalent.
Response time for requests delivered on media	Characterize the nominal or average response times for delivery on media anticipated for your activity (i.e. time until media is ready to leave your activity): 1 - Products to be delivered by media will be packed and ready to ship in about 3 days; 2 - Products to be delivered by media will be packed and ready to ship in about 7 days; 3 - Products to be delivered by media will be packed and ready to ship in about 30 days;

B.2.6 User Support ADS Information

Table B9 below contains the parameters needed by the CET for the User Support functional area. The CET selection box for documentation parameters for your new activity's ADS is shown in Figure 16, Section 3.3.7 above.

Table B9 - User Support Functional Area ADS Parameters

Outreach Activity LOS (Advisory – No Effect on Estimate)	Indicate the level of Outreach effort you plan for your activity: 0 – No user support. 1 - Will include training sessions, expanded booth support at 4 conferences/yr, production and dist of outreach material; 2 - Will include expanded booth support at 4 conferences/yr, production and distribution of outreach material; 3 - Will include booth support at 4 conferences/yr, production and distribution of outreach material; 4 - Will include production and distribution of outreach material; Note - 'expanded booth support' above includes extra booth space for mini-workshops and hands-on training experiences, extra staffing to prepare material and staff the booth to provide that.
Average Contacts per User /Yr	Estimate the average number of times a typical active user will contact user support (phone, email, etc.) per year.
User Multiplier	Estimate, as a multiplier to be applied to the estimated number of by request users when service is well established (see Table B8 part C above), the approximate number of active or potential users who will contact user support in a peak year when services are well established. For example, a 1.25 would mean 5 for every 4 active users, .75 would mean 3 for every 4 active users.

B.3 ADS Information for Non-Operating Functional Areas

This section addresses the ADS parameters needed by the CET to describe implementation, sustaining engineering, engineering support, technical coordination, and facility / infrastructure functional areas. In each case tables will address all of the parameters needed, with valid choices for levels of service defined, and any notes on individual parameters as needed.

B. 3.1 Implementation ADS Information

Table B10 below contains the parameter needed by the CET for the Implementation functional area. The CET selection box for the implementation parameter for your new activity's ADS is shown in Figure 16, Section 3.3.8 above.

Table B10 - Implementation Functional Area ADS Parameter

Applications SW Development LOS (Advisory – No Effect on Estimate)	Describe the development of applications software to meet particular user needs / requests that your activity will perform: 1 - Data mining, data integration, custom science products, data manipulation tools; 2 - Custom science products, data manipulation tools; 3 - Data manipulation tools; 4 - none.
---	---

B.3.2 Sustaining Engineering ADS Information

Table B11 below contains the parameters needed by the CET for the Sustaining Engineering functional area. The CET selection box for the sustaining engineering parameter for your new activity's ADS is shown in Figure 17, Section 3.3.8 above.

Note that Sustaining Engineering includes implementation performed after the end of the implementation period you specify for your new activity. This would not estimate a major re-development effort.

Table B11 - Sustaining Engineering Functional Area ADS Parameters

Sustaining Engineering LOS	Level of service for sustaining engineering (software maintenance, enhancement of existing systems): 1 - No or very infrequent interruptions in operational services are allowed; 2 - Occasional interruptions in operational services are allowed; 3 - Interruptions in operational services are a secondary concern.
---------------------------------------	---

B.3.3 Engineering Support ADS Information

Table B12 below contains the parameters needed by the CET for the Engineering Support functional area. The CET selection box for the engineering support parameter for your new activity's ADS is shown in Figure 18, Section 3.3.8 above.

Engineering support includes a range of items: systems engineering, test engineering, configuration management, system administration, network administration, database administration, logistics support, facility planning, etc.

Table B12 - Engineering Support Functional Area ADS Parameters

Engineering Support LOS (Advisory – No Effect on Estimate)	Level of service: 1 - No or very infrequent interruptions in operational services are allowed; 2 - Occasional interruptions in operational services are allowed; 3 - Interruptions in operational services are a secondary concern.
---	--

B.3.4 Technical Coordination ADS Information

Table B13 below contains the parameters needed by the CET for the Technical Coordination functional area. The CET selection box for the technical coordination parameters for your new activity's ADS is shown in Figure 19, Section 3.3.8 above.

Table B13 - Technical Coordination Functional Area ADS Parameters

Metrics	Coordinating on definition of standard metrics, metrics reporting requirements, tools, etc., perhaps by membership on the Metrics and Performance Analysis and Reporting Working Group.
Technology Infusion	Overall data systems and services architecture and information technology and processes for bring new technology, architectural evolution, into new data activities, perhaps by membership on Technology Infusion Working Group.
Data Life Cycle	Data stewardship best practices, standards, sharing of responsibilities (e.g. mutual backup, planning for transition to long term archive, etc.)
Re-Use	Planning processes for facilitating or accomplishing technology re-use, perhaps by membership on Re-Use Working Group.
Standards	Participation in NASA processes for adopting and maintaining standards, interfaces between NASA program elements, etc., perhaps by membership on the Standards Working Group.
Levels of Service	Coordination of user support standards and practices, referrals, exchange of information, to provide best overall service (e.g., the DAAC User Services Working Group).
User Services	Coordination of user support standards and practices, referrals, exchange of information, to provide best overall service, perhaps by membership on the DAAC User Services Working Group.

B.3.5 Miscellaneous Non-Staff Cost Items

Table B14 below contains the Miscellaneous Non-Staff Cost parameters needed by the CET. The CET selection box for miscellaneous non-staff cost items for your new activity's ADS is shown in Figure 18, Section 3.3.8 above.

Table B14 – Miscellaneous Non-Staff Cost ADS Parameters

Backup Archive Facility LOS (Advisory – No Effect on Estimate)	Indicate your plans for a backup storage facility: 1 - Secure, off-site backup archive / storage facility; 2 - On-site but separate secure backup archive / storage facility (e.g. in same or near by building). 3 - On-site co-location of backup storage with primary archive (e.g. in same or adjoining rooms). 0 - No backup archive facility.
Annual Budget for Data Purchase	If you plan to purchase data for your activity, enter an approximate annual budget for data purchases.
Annual Budget for Computer Services	If you plan to obtain computer services from a provider, enter an approximate annual budget for computer services.
Annual Budget for Travel	Provide the approximate annual budget required to support travel (U.S., international as appropriate) by the staff of your data activity.

Appendix C - “Moxie”, an Example of an Activity Dataset

This appendix provides a set of data entry information for a sample Activity Dataset named “Moxie”. See Appendix B above for an explanation of the items, with options (e.g. for valid levels of service). The Activity Dataset “Moxie” describes a data activity that performs every function and is intended to serve as an example or model for a user building a new Activity Dataset. Note that a new Activity Dataset might not perform all of the functions that Moxie includes. The Moxie Activity Dataset information is presented in the sequence and form in which it would be (and was) entered into the CET. Following the information for “Moxie” are the specific changes made to “Moxie” to produce the “What-If” example shown above in Table 4 above.

C.1 Activity Level Information for Moxie

Activity Dataset Name: Moxie

Functions Performed: Ingest, Processing, Documentation, Archive, Access and Distribution, User Support.

Mission / Project Start Calendar Year: 2010

Period of Initial Implementation Period, Years: 2

Calendar Year of Operations Start: 20012

Duration of Operations, Years: 6

Management Staff Labor Rate: 175

Admin Support Staff Labor Rate: 100

Development / Engineering Staff Labor Rate: 150

Technical / Science Staff Labor Rate: 150

Operations Staff Labor Rate: 110

Inflation Rate: 4%

C.2 Ingest Information for Moxie

Moxie includes three ingest streams:

Parameter	Stream 1	Stream 2
Name of Ingest Stream	Stream_A	Stream_B
Number of Product types in Stream	2	4
Calendar Year - Ingest Start	0	0
Calendar Year - Ingest End	0	0
Source	Site_A	Center_A
How Received	1	1
Level of Service	1	2
Retention Period, Years	0	0
Automation LOS	2	2
Products of Stream Ingested / Day	2	24
Total Volume of Stream, GB/Day	1.5	3.0

C.3 Processing Information for Moxie

Activity Level Processing Information:
Science Software Service Option (Advisory): 3
Calibration - Validation LOS: 1

Moxie includes three operational Processing Streams:

Parameter	Stream 1	Stream 2	Stream 3
Name of Processing Stream	Level_1	Level_2	Level_3
Number of Product Types in Stream	12	6	2
Calendar Year - Production Start	0	0	0
Calendar Year - Production End	0	0	0
Science Algorithm Software Source	1	1	1
Product Science Quality Assurance	1	1	1
Processing Automation LOS	2	2	2
Production Mode	1	1	1
Operational Production Mode, if Applicable	1	1	1
Operational Processing LOS, if Applicable	1	2	2
Non-Operational Processing LOS	0	0	0
Product Retention Period (each version)	0	0	0
Products in this Stream generated per day	104	24	48
Volume of this Stream generated, GB/day	6.000	3.0	2.000
Reprocessing LOS for this Stream	4	4	4
Reprocessing Interval for this Stream	0	0	0

C.4 Documentation Information for Moxie

Documentation LOS: 1
User Comment LOS: 2

C.5 Archive Information for Moxie

Archive Purpose: 2
Archive Media Unit Capacity: 50
Archive Media Unit Price: \$100
Archive Backup Fraction, Pct.: 75
Archive Media Re-Use Plan: 4
Archive Monitoring LOS: 2
Archive Entry/Exit Screening LOS (Advisory) : 1
Archive Initial Inactive Product Types: 0
Archive Initial Inactive Products: 0
Archive Initial Inactive Volume: 0
Archive Initial Active Product Types: 0
Archive Initial Active Products: 0
Archive Initial Active Volume: 0

C.6 Access and Distribution for Moxie

a. Overall:

Access and Distribution Scope: 1
Distribution Service Modes: 1
Distribution Means LOS: 2
Supporting Data Services LOS (Advisory): 2

b. Operational Distribution:

Moxie includes two Operational Distribution Streams:

Parameter	Stream 1	Stream 2
Name of Operational Distribution Stream	Level_2X	Level_3X
Number of Product Types in Dist. Stream	4	2
Destination for Distribution Stream	Center_A	Center_B
Distribution Mode	1	1
Distribution Stream Start Calendar Yr	0	0
Distribution Stream End Calendar Year	0	0
Dist. Stream Products / Day, Network	24	18
Dist. Stream Products / Day, Media	0	0
Dist. Stream Volume / Day, Network, GB	3.0	1.0
Dist. Stream Volume / Day, Media, GB	0	0

c. By Request Distribution:

By-Request Distribution Profile: 1
Expected number of users when services are well established: 3500
Estimated average number of requests per user, per year: 1.25
Estimated average number of products per request: 20
Out of every ten products distributed, estimated number by network: 8
Out of 10 units of volume distributed, estimated number by network: 8

C.7 User Support Information for Moxie

Outreach Activity LOS (Advisory): 2
Average contacts per user per year: 0.60
User Multiplier: 0.80

C.8 Sustaining Engineering, Engineering Support & Implementation Information for Moxie

Sustaining Engineering LOS: 1
Engineering Support LOS (Advisory): 1
Applications Software Development LOS (Advisory): 2

C.9 Technical Coordination for Moxie

Technical Coordination is performed in the following areas:

Metrics,
Data Life-Cycle,
User Services.

C.10 Miscellaneous Non-Staff Cost Information for Moxie

Backup Archive Facility LOS (Advisory): 2
Annual Budget for Data Purchase: 2
Annual Budget for Computing Services: 3
Annual Budget for Travel: 25

C.11 “What If” Changes to the Moxie Activity Dataset

The following changes were made to the “Moxie” activity dataset to produce the “What-If” example shown in Table 4 above.

Technical / Science Staff Labor Rate: changed to 170 from 150
Operations Staff Labor Rate: changed to 90 from 110

A new ingest stream was added:

Parameter	
Name of Ingest Stream	Stream_C
Number of Product types in Stream	4
Calendar Year - Ingest Start	0
Calendar Year - Ingest End	0
Source	Center_B
How Received	1
Level of Service	2
Retention Period, Years	0
Automation LOS	2
Products of Stream Ingested / Day	12
Total Volume of Stream, GB/Day	2.0

A new processing stream was added:

Parameter	
Name of Processing Stream	Level_2B
Number of Product Types in Stream	6
Calendar Year - Production Start	0
Calendar Year - Production End	0
Science Algorithm Software Source	1
Product Science Quality Assurance	1
Processing Automation LOS	2
Production Mode	1
Operational Production Mode, if Applicable	1
Operational Processing LOS, if Applicable	2
Non-Operational Processing LOS	0
Product Retention Period (each version)	0
Products in this Stream Generated per day	24
Volume of this Stream Generated, GB/day	2.5
Reprocessing LOS for this Stream	4
Reprocessing Interval for this Stream	0