

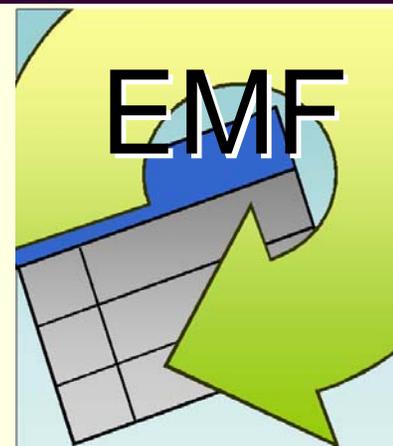
Emissions Modeling Framework: Status Update and EPA's Applications

Marc Houyoux, [US EPA](#)

Madeleine Strum, [US EPA](#)

Rich Mason, [NOAA](#)

Alison Eyth, [UNC Institute for the Environment](#)



Overview

- Background – what is the EMF?
- Part 1: Status of EMF Development
- Part 2: How EPA is using the EMF and benefits
- How to access, install, and learn
- To get on mailing list for EMF
 - Send me email at houyoux.marc@epa.gov
 - Fill in sign-up sheet today

What is the EMF?

Emissions Modeling Framework (EMF): Marc Houyoux (n)

File Manage Window Help

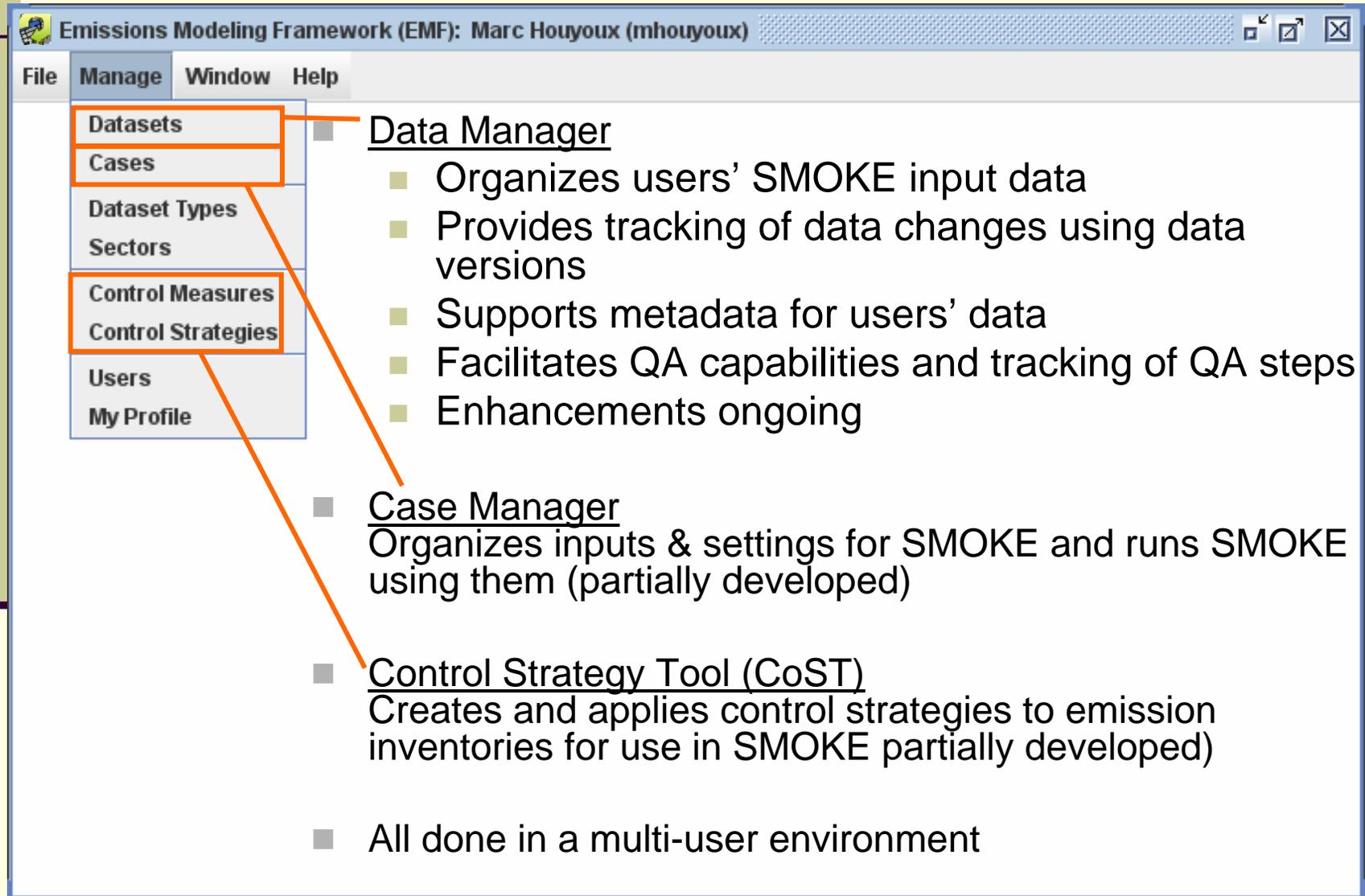
- Software application for Emissions Modelers
- Client-server (multiple users getting to the same data) OR single-machine
- Shares emissions modeling info within an organization, not over the internet

Status

Last Update : 2007/05/10 16:23

Message Type	Message	Timestamp
--------------	---------	-----------

What can the EMF do?



Data Manager

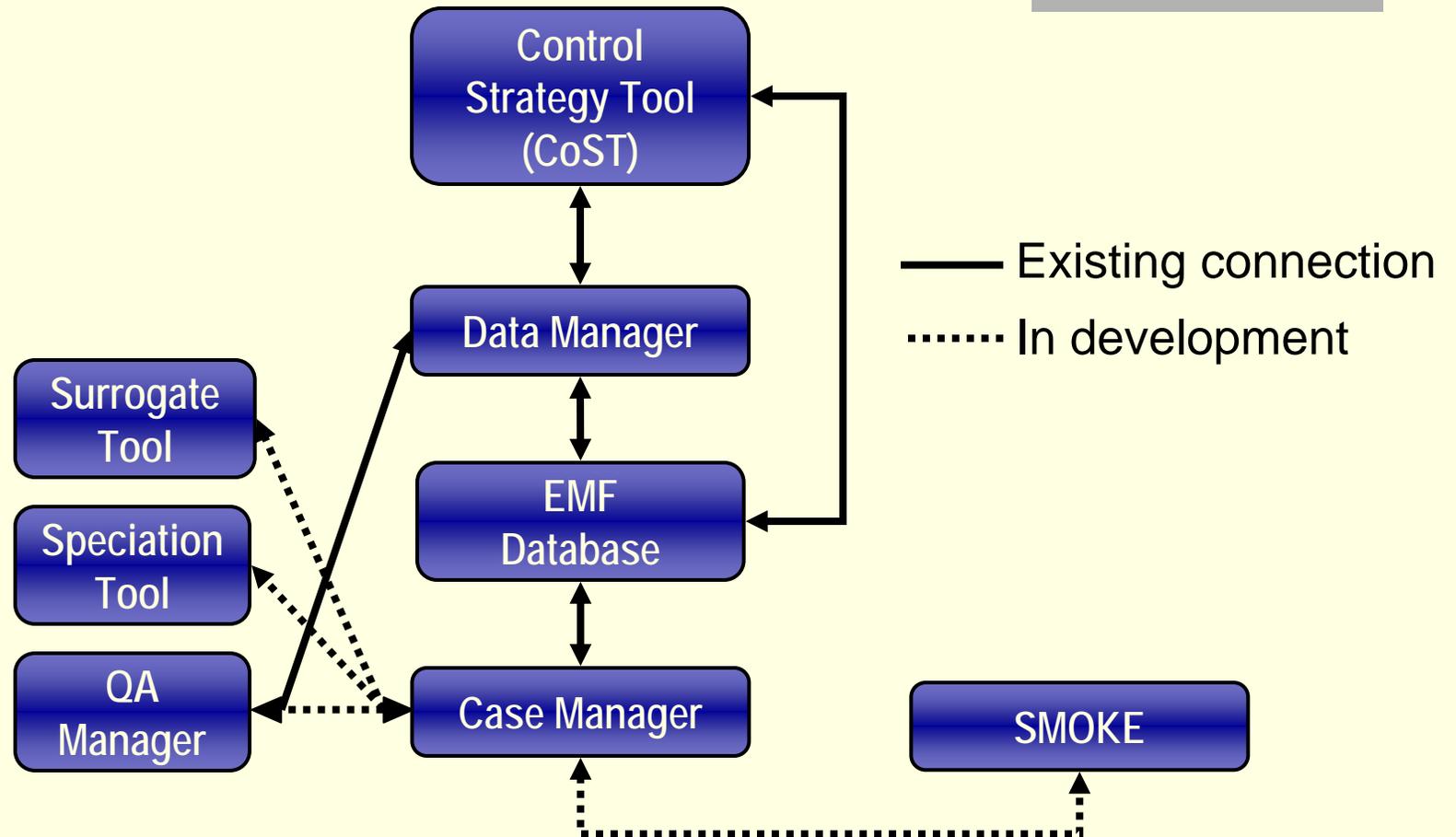
- Organizes users' SMOKE input data
- Provides tracking of data changes using data versions
- Supports metadata for users' data
- Facilitates QA capabilities and tracking of QA steps
- Enhancements ongoing

■ **Case Manager**
Organizes inputs & settings for SMOKE and runs SMOKE using them (partially developed)

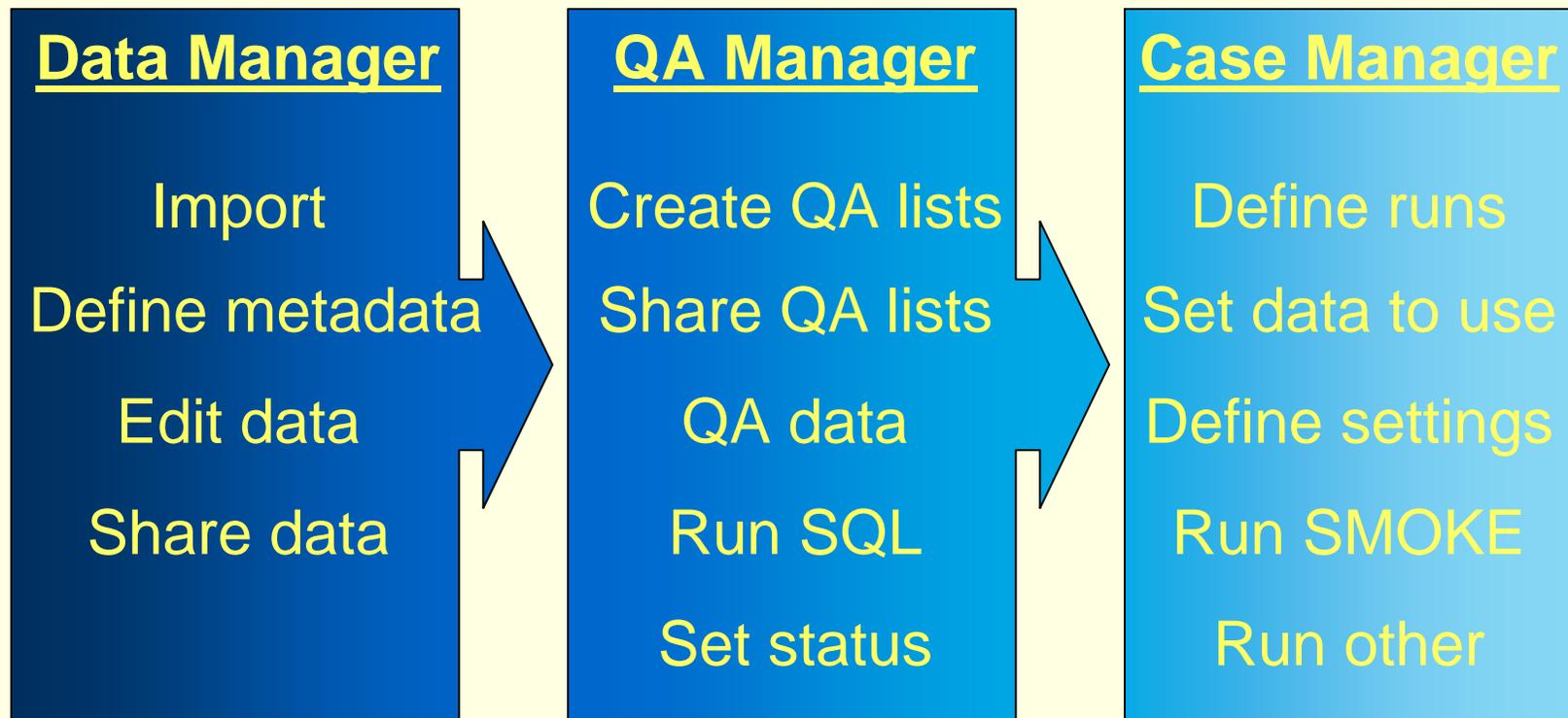
■ **Control Strategy Tool (CoST)**
Creates and applies control strategies to emission inventories for use in SMOKE (partially developed)

■ All done in a multi-user environment

Components of the EMF



Steps for Using the EMF



Part 1: Status of EMF Development



Overall status

- Data Manager
 - In use by EPA for its current 2002 modeling
 - Ongoing improvements this year
- QA Manager
 - In use by EPA, and being phased in gradually
 - Ongoing improvements this year
- Case Management
 - Partially completed, not yet running SMOKE
 - This feature is expected to be in trial at EPA summer 2007
- Control Strategy Tool (CoST)
 - Have a working version, but development continues
 - Database in development
- Spatial Surrogate Tool: final version available, but not yet linked to EMF
- Speciation Tool: beta version available, but not yet linked to EMF

Data Manager

– View metadata

- History of changes

– View data values

Export SMOKE-ready data

Add, remove, or revise metadata

- Summary info
- User-defined keywords
- Notes
- QA steps, status, and tables

Import new datasets

– Create new versions of data

– Change data

– Record who, when, what, why

#	Select	Modified D...	Creator
1	<input type="checkbox"/>	7/02/23 ...	mls
2	<input type="checkbox"/>	2007/04/17 ...	cy
3	<input checked="" type="checkbox"/>		mls
4	<input type="checkbox"/>		mls
5	<input type="checkbox"/>		mls
6	<input type="checkbox"/>		mhouyoux
7	<input type="checkbox"/>		cy
8	<input type="checkbox"/>		mhouyoux
9	<input type="checkbox"/>		mhouyoux
10	<input type="checkbox"/>		mhouyoux
11	<input type="checkbox"/>	2006/11/27 ...	mhouyoux
12	<input type="checkbox"/>		ozl
13	<input type="checkbox"/>		mls
14	<input type="checkbox"/>		mls
15	<input type="checkbox"/>		mls
16	<input type="checkbox"/>		mls

41 rows : 10 columns [Filter: Type does not contain PTDAY, Type does not contain Specific, Type does not contain Report, Type does not contain CMAQ, Sort: Typ...

View **Edit Properties** **Edit Data** **Remove** **Import** **Export** **Close**

QA Manager

The screenshot shows the 'Dataset Properties Editor' window for 'gspro_cb05_notoxics_cmaq_poc'. The 'QA' tab is active, displaying a table of QA steps. The table has columns for #, Select, Version, Name, Required, Program, Order, Status, When, Who, and Comments. Two rows are visible: Row 1 (Check that all profiles sum to 1) and Row 2 (Check that all pollutants have profile entries). Below the table, there are four buttons: 'Add from Template', 'Add Custom', 'Edit', and 'Set Status'. Four yellow callout boxes with orange borders provide instructions for these buttons.

#	Select	Version	Name	Required	Program	Order	Status	When	Who	Comments
1	<input type="checkbox"/>	0	Check that all profiles sum to 1	<input type="checkbox"/>	SQL	0.0	Not Started	N/A	mls	
2	<input checked="" type="checkbox"/>	0	Check that all pollutants have profile entries	<input type="checkbox"/>	Manual (No program)	0.0	Not Started	N/A	mls	

41 rows : 10

2 rows : 12 columns [Filter: None, Sort Version (-), Order(-)]

Buttons: Add from Template, Add Custom, Edit, Set Status, View, Save, Close

Callout 1: Add QA steps to a version of this dataset from defined steps for all datasets of this type

Callout 2: Add QA steps to a version of this dataset only

Callout 3: Edit, set Status, and/or run QA step

Callout 4: Set status of multiple steps at same time

QA M

Dataset Properties Editor

Summary Data Ke...

#	Select	#	Select	Version	Ch...
1	<input type="checkbox"/>	1	<input type="checkbox"/>	0	Ch...
2	<input type="checkbox"/>	2	<input checked="" type="checkbox"/>	0	Ch...
3	<input checked="" type="checkbox"/>				
4	<input type="checkbox"/>				
5	<input type="checkbox"/>				
6	<input type="checkbox"/>				
7	<input type="checkbox"/>				
8	<input type="checkbox"/>				
9	<input type="checkbox"/>				
10	<input type="checkbox"/>				
11	<input type="checkbox"/>				
12	<input type="checkbox"/>				
13	<input type="checkbox"/>				
14	<input type="checkbox"/>				
15	<input type="checkbox"/>				
16	<input type="checkbox"/>				

41 rows : 10

Add from Template Add

View

Edit QA Step: Check that all pollutants have profile entries - gspro_cb05_notoxics_cmaq_poc (...)

Name: Check that all pollutants have profile entries

Version: Initial Version (0)

Program: Manual (No program)

Arguments: If SQL step, include syntax here

Order: 0.0

Required?

Description: Space to include more information on how to do the step

QA Status: Not Started

QA User: Not Started

QA Date: Skipped

Configuration: In Progress

Generated

Output Name:

Run Status:

Run Date:

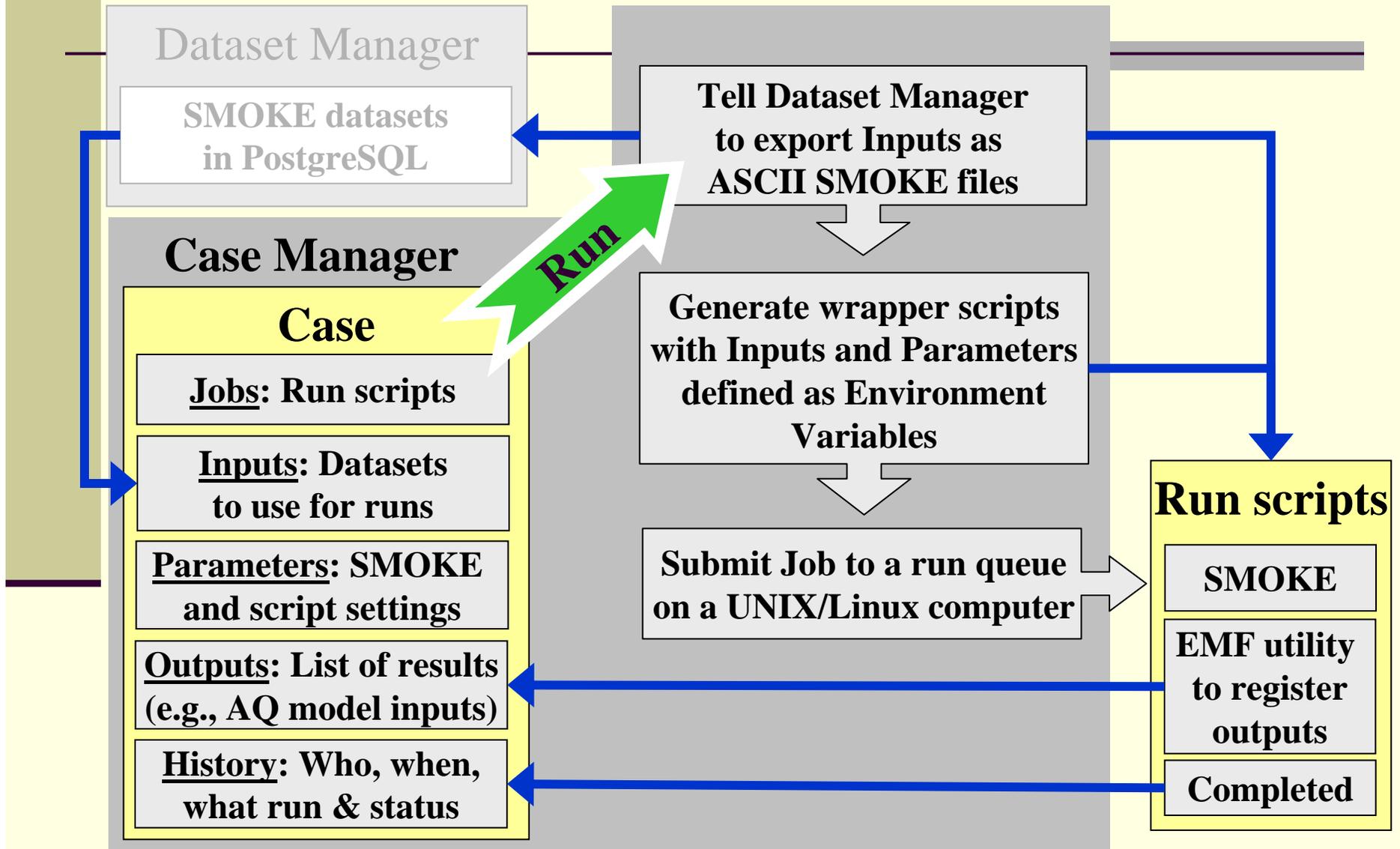
Current Output?

Comments: Space to document what you found

Export Folder: Browse

Save Close Run View Results Export

Case Management Overview



Case Manager

Please select one or more Cases

Refresh

#	Select	Project	Model to Run	Modeling Regn.	Last Modifie...	
1	<input type="checkbox"/>	based platform	SMOKE 2.3	National	Madeleine ...	
2	<input checked="" type="checkbox"/>	based platform	SMOKE 2.3	National	Chris Allen	
3	<input type="checkbox"/>	2002cc_V8_CARBONLY_36	2002 based platform	SMOKE 2.3	National	Madeleine ...
4	<input type="checkbox"/>	2002_P	d platform	SMOKE 2.3	National	Madeleine ...
5	<input type="checkbox"/>	2002_P	d platform	SMOKE 2.3	National	Madeleine ...

5 rows : 21 columns [Filter: Name starts with 2002, Name does not contain test, Sort: None]

View Edit New Remove Copy Close

Opens Case in View-only mode (not yet implemented)

Opens Case in "Edit" mode. Only way to open it now.

Case Manager

Summary Tab

Case Editor: 2002ac_V3_CAPONLY_36

Case Manager Summary Tab

Summary Jobs Inputs Parameters Outputs History

Name: 2002ac_V3_CAPONLY_36 Abbreviation: 2002ac

Category: run SMOKE for CMAQ Is Final: Is Template:

Description: Version 3 Platform (Spring) CAP only using NEI 2002 v3,CB05, SPECIATE4.0 for EVALUATION Sectors: ptfire alm nonat

Project: 2002-based platform Copied From: 2002_Platformv2_CAPONLY_36

Run Status: Not Started Last Modified Date: 2007/05/10 12:50 Last Modified By: Marc Houyoux

Model to Run: SMOKE 2.3.2 Air Quality Model: CMAQ

Modeling Region: National Speciation: cmaq_cb05pm

Control Region: National Meteorological Year: 2002

I/O API Grid Name: US36KM_148X112 Base Year: 2002

Grid Resolution: 36km Future Year: 2002

Start Date & Time: 2002/01/01 00:00 End Date & Time: 2002/12/31 23:59

5 rows : 21

View

Refresh

Modifie...

eleine ...

s Allen

eleine ...

eleine ...

eleine ...

Close

Save Close

Case Manager

Jobs Tab

Case Editor: 2002ac_V3_CAPONLY_36

Summary **Jobs** Inputs Parameters Outputs History

#	Select	Name	Job Num.	Sector	Executable	Arguments	Exec. Version
1	<input type="checkbox"/>	Annual afdust	0.00	afdust	smk_afdust_2002ac.csh	-q "1 2 3 4"	0
2	<input type="checkbox"/>	Annual ag	0.00	ag	smk_ptipm_2002ac.csh	-q "1 2 3 4"	0
3	<input type="checkbox"/>	Annual alm	0.00	alm	smk_alm_2002ac.csh	-q "1 2 3 4"	0
4	<input type="checkbox"/>	Annual nonpoint	0.00	nonpt	smk_nonpt_2002ac.csh	-q "1 2 3 4"	0
5	<input type="checkbox"/>	Annual nonptfire	0.00	nonptfire	smk_nonptfire_2002ac.csh	-q "1 2 3 4"	0
6	<input type="checkbox"/>	January nonroad	0.00	nonroad	smk_nonroad_2002ac.csh	-m 1 Y	0
7	<input type="checkbox"/>	July nonroad	0.00	nonroad	smk_nonroad_2002ac.csh	-m 7 Y	0
8	<input type="checkbox"/>	January onroad	0.00	onroad	smk_onroad_2002ac.csh	-m 1 Y	0
9	<input type="checkbox"/>	July onroad	0.00	onroad	smk_onroad_2002ac.csh	-m 7 Y	0
10	<input type="checkbox"/>	January ptfire	0.00	ptfire	smk_ptfire_2002ac.csh	-m 1 N	0
11	<input type="checkbox"/>	July ptfire	0.00	ptfire	smk_ptfire_2002ac.csh	-m 7 Y	0
12	<input type="checkbox"/>	January ptipm	0.00	ptipm	smk_ptipm_2002ac.csh	-m 1 Y	0
13	<input type="checkbox"/>	July ptipm	0.00	ptipm	smk_ptipm_2002ac.csh	-m 7 Y	0
14	<input type="checkbox"/>	Annual ptnonipm	0.00	ptnonipm	smk_ptnonipm_2002ac.csh	-q "1 2 3 4" Y	0

5 rows : 21
14 rows : 18 columns [Filter: None, Sort: Sector(+), Name(+), Executable(+)]

View Add Remove Edit Copy **Run**

Save Close

Refresh

Modifie...

eleine ...

& Allen

eleine ...

eleine ...

eleine ...

Close

Case Manager

Inputs Tab

Case Editor: 2002ac_V3_CAPONLY_36

Summary Jobs **Inputs** Parameters Outputs History

Input Folder: /bud2/2002/smoke/inventory/v3/2002ac Browse

Icons: Home, Signal, Eye, \$000, Left Arrow, Right Arrow, Refresh

#	Select	Input	Env. Var.	Sector	Job	Pr
1	<input type="checkbox"/>	Emission inventory, "afdust" annual	ARINV	afdust	All jobs for sector	Smk
2	<input type="checkbox"/>	Emission inventory, "ag" annual	ARINV	ag	All jobs for sector	Smk
3	<input type="checkbox"/>	GRIDDESC, Grid Description Information	GRIDDESC	All sectors	All jobs for sector	Grdr
4	<input type="checkbox"/>	SRGDESC, Surrogate descriptions	SRGDESC	All sectors	All jobs for sector	Grdr
5	<input type="checkbox"/>	Surrogate cross-reference, all nonpoint and nonroad	AGREF	All sectors	All jobs for sector	Grdr
6	<input type="checkbox"/>	Surrogate ratios USA 36km		All sectors	All jobs for sector	Grdr
7	<input type="checkbox"/>	PELVCONFIG, point elevated source selection	PELVCONFIG	All sectors	All jobs for sector	Layp
8	<input type="checkbox"/>	ARTOPNT, area-to-point data	ARTOPNT	All sectors	All jobs for sector	Smk
9	<input type="checkbox"/>	COSTCY, Country, State, County Information	COSTCY	All sectors	All jobs for sector	Smk
10	<input type="checkbox"/>	Emission inventory, "ptfire", firedata 07 July firstday	PTDAY_07C	All sectors	All jobs for sector	Smk
11	<input type="checkbox"/>	INVTABLE, Inventory Table	INVTABLE	All sectors	All jobs for sector	Smk
12	<input type="checkbox"/>	PSTK, Stack replacement	PSTK	All sectors	All jobs for sector	Smk
13	<input type="checkbox"/>	SCCDESC, SCC Descriptions	SCCDESC	All sectors	All jobs for sector	Smk
14	<input type="checkbox"/>	GSCNV, pollutant-to-pollutant conversions	GSCNV	All sectors	All jobs for sector	Spcr
15	<input type="checkbox"/>	GSPRO, Speciation profile (CAP only CB05 POC)	GSPRO	All sectors	All jobs for sector	Spcr
16	<input type="checkbox"/>	GSREF, Speciation xref (for CB05)	GSREF	All sectors	All jobs for sector	Spcr
17	<input type="checkbox"/>	HOLIDAYS	HOLIDAYS	All sectors	All jobs for sector	Term
18	<input type="checkbox"/>	Temporal cross-reference (ATREF) all nonpoint and non...	ATREF	All sectors	All jobs for sector	Term

5 rows : 21 50 rows : 13 columns [Filter: None, Sort: Sector(+), Program(+), Input(+)]

Buttons: View, Add, Remove, Edit, Copy, **View Dataset**, **Export Inputs**, Show All, Save, Close

Case Manager

Parameters Tab

Case Editor: 2002ac_V3_CAPONLY_36

Summary Jobs Inputs **Parameters** Outputs History

Refresh

#	Select	Parameter	Sector	Job	Program	Emt. Var.
1	<input type="checkbox"/>	SMOKE run type	All sectors	All jobs for sector	n/a	SMOKE_RUNTYPE
2	<input type="checkbox"/>	Case name	All sectors	All jobs for sector	n/a	CASE
3	<input type="checkbox"/>	I/O API sphere setting	All sectors	All jobs for sector	All Programs	IOAPI_ISPH
4	<input type="checkbox"/>	Spinup duration	All sectors	All jobs for sector	All Programs	SPINUP_DURATION
5	<input type="checkbox"/>	Check for duplicates	All sectors	All jobs for sector	All Programs	RAW_DUP_CHECK
6	<input type="checkbox"/>	Use pollutant conversion file?	All sectors	All jobs for sector	Spcmat	POLLUTANT_CONVERSION
7	<input type="checkbox"/>	Renormalize temporal profiles?	All sectors	All jobs for sector	Temporal	RENORM_TPROF
8	<input type="checkbox"/>	Output time zone	All sectors	All jobs for sector	All Programs	OUTZONE
9	<input type="checkbox"/>	Temporal type	onroad	All jobs for sector	All programs for sector	L_TYPE
10	<input type="checkbox"/>	Merge type	onroad	All jobs for sector	All programs for sector	M_TYPE
11	<input type="checkbox"/>	Smkinven formula	onroad	All jobs for sector	Smkinven	SMKINVEN_FORMULA

5 rows : 21

11 rows : 13 columns [Filter: None, Sort: None]

View Add Remove Edit Copy Show All

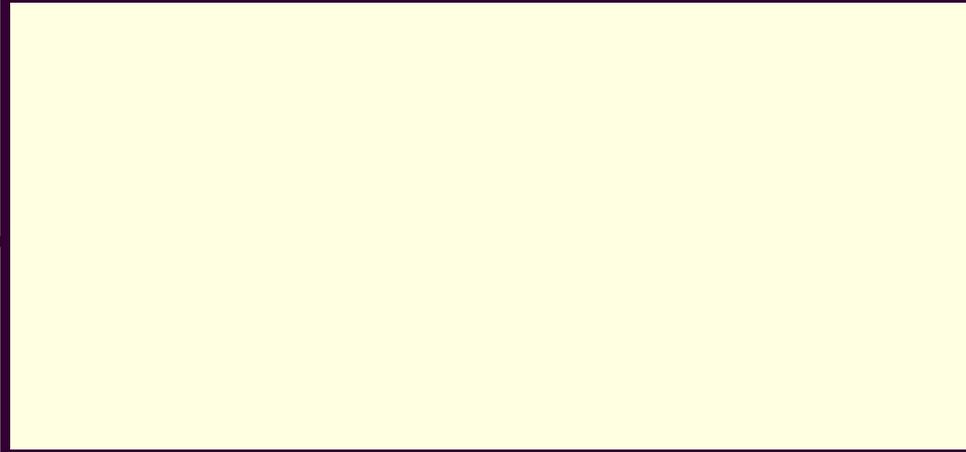
Save Close

Modifie...
eleine ...
s Allen
eleine ...
eleine ...
eleine ...

Close

Part 2:

How EPA is using the EMF and benefits



How Data Manager helps EPA

- Stores data in a PostgreSQL database accessible by all. All team members know where to find the latest data. Others' data changes shared.
- Supports multiple versions of datasets. We can access both old and current versions of data, for reproducibility and transparency.
- Users can edit data in a user interface to the database. Convenient.
- Automatically records metadata, such as who/when makes changes. Allows users to add what/why of changes. Eases our burden for keeping track of details manually. Adds transparency after work is done.
- Exports SMOKE-ready files, including documentation in file headers. Saves us time & prevents possible format errors.
- Prevents data problems - we can focus on analysis and modeling needs.

How QA Manager helps EPA

- Stores QA steps, who/when steps performed, & QA status along with the data. We can easily find QA status of data. We can see the overall QA status, what QA steps complete, who performed steps, and the results. Users can add steps as needed.
- Tracks QA steps by dataset version. We know what steps were performed on each version. Helps prevent skipping QA.
- Supports running SQL queries on the data to perform QA steps and summaries. Give us the flexibility and convenience of database queries. Runs within EMF and results stored for present and future use.
- Templates of QA steps can be created and reused. Helps ensure QA of data gets done, decreases learning curve for new emissions modelers.
- Tracks both automated (run by EMF) and external and/or manual QA steps. Gives us needed flexibility to get work done.

How Case Manager helps EPA

- Data in a “Case” (SMOKE run) linked to datasets in Data Manager. We (will) have one system to use for data handling and running SMOKE.
- Exports all SMOKE-ready files needed for a SMOKE run with a single mouse click. Saves time and prevents possible errors associated with manually writing out many data files & formats from many other data sources.
- User interface for SMOKE (in progress). We have a permanent record of SMOKE inputs, settings, and location of outputs. Reduces learning curve for new emissions modelers.
- Submits SMOKE jobs to a compute server and not run on local computer (in progress). Supports desktop access to data and SMOKE cases, while running elsewhere (e.g., Linux servers or clusters) to meet computational needs.

How to get access, install, learn

- EMF Website available now at:
<http://www.ie.unc.edu/cempd/projects/emf/install/>
- Beta public release in early June 2007
- More complete release by August 15th, 2007
- Basic and Advanced Training modules available
- EMF Mailing list – e-mail to houyoux.marc@epa.gov
or sign-up sheet today