Evaluation of EPA's Modeled Attainment Test Software (MATS) for Modeled Future Design Value (DV_F) Calculation in HGB

Jim Smith
Southeast Texas Photochemical Modeling Technical Committee
February 18, 2009

CAMx Ozone Modeling in SIP Development

The Big Picture

Base Case	Day-specific meteorology and emissions; replicate what actually happened
Baseline Case	Day-specific meteorology and Typical emissions; used in RRF to predict future design values
Future Base Case	Apply future growth + on-the-books controls to estimate future ozone
Control Strategy Testing	Determine control strategies that will effectively reduce ozone
SIP	Document modeling procedures

Background

- EPA developed Model Attainment Test Software (MATS) for calculation of future design values for attainment demonstrations.
- Prior to the release of the software, TCEQ modelers developed in-house procedures to implement the EPA guidance for performing this calculation.
- Some stakeholders and modeling researchers contend TCEQ should be using MATS because
  - EPA has devoted considerable resources in providing the software for states to use
  - Other states are using MATS
  - There appear to be some minor inconsistencies between MATS' and TCEQ's calculations.

Background (Continued)

- The question - Is TCEQ justified in continuing to use its own PERL script after evaluation of MATS?

- To address this question, we used MATS to process output from the 1c baseline and the CS02 2018 future case model runs, and compared MATS' DV_F values with our own.

DV_F and RRF Primer

- **Regulatory Design Value (DV_R)** :
  - Average of three consecutive years' fourth highest 8-hour ozone concentrations measured at an individual monitor
  - Highest DV_R in an area determines attainment status, classification
- **Baseline Design Value (DV_B):**
  - Average of three years' DV_R values, as per EPA Guidance
  - Basis for modeled attainment test
- **Relative Response Factor (RRF):**
  - A ratio estimating the model's response at a monitoring site
  - Based on modeled baseline and future ozone concentrations in a neighborhood near the monitor
- **Future Design Value (DV_F):**
  - The product of the RRF and the DV_B: $DV_F = RRF * DV_B$
  - Used to demonstrate attainment of the ozone NAAQS

Baseline Design Value Example

- To calculate the DV_B for 2005 at Deer Park, we need the three DV_R values which include 2005:
  - The 2005 DV_R is the average 4th highest ozone concentration for 2003, 2004, and 2005. For Deer Park,

    The 2005 DV_R is (113 + 97 + **92**) / 3 = 100 (truncated)

    The 2006 DV_R is (97 + **92** + 101) / 3 = 96 (truncated)

    The 2007 DV_R is (**92** + 101 + 86) / 3 = 93

    2005 4th high

- The DV_B for 2005 is then the average of the three DV_R values: (100 + 96 + 93) / 3 = 96.3

2005 Base Year

2005 Design Value	4th high 2003	4th high 2004	4th high 2005	
2006 Design Value		4th high 2004	4th high 2005	4th high 2006
2007 Design Value			4th high 2005	4th high 2006	4th high 2007

Average of 2005 DV, 2006 DV, and 2007 DV - weights the 2005 4th high 8-hour ozone value as most influential

How RRF is calculated at a monitor

- Select a suitable area surrounding the monitor, usually 3X3, 5X5, or 7X7 grid cells, depending on grid cell size.
- For each day modeled, find the maximum modeled baseline 8-hour ozone concentration in the selected area.
- Select days to use in the RRF calculation. EPA recommended method is:
  1. Select days with max modeled baseline 8-hour ozone concentration ≥ a threshold value T_1 (default 85 ppb)
  2. If < 10 days selected in Step 1, then reduce threshold progressively until either:
     a) Ten days are selected, or
     b) a lower threshold T_2 (default 70 ppb) is reached.
     c) If T_2 is reached before 10 days are selected for a monitor, Guidance recommends states discuss with regional office.

How RRF is calculated at a monitor (cont.)

- After days are selected, the baseline and future case modeled ozone concentrations* are averaged across days for each monitor.

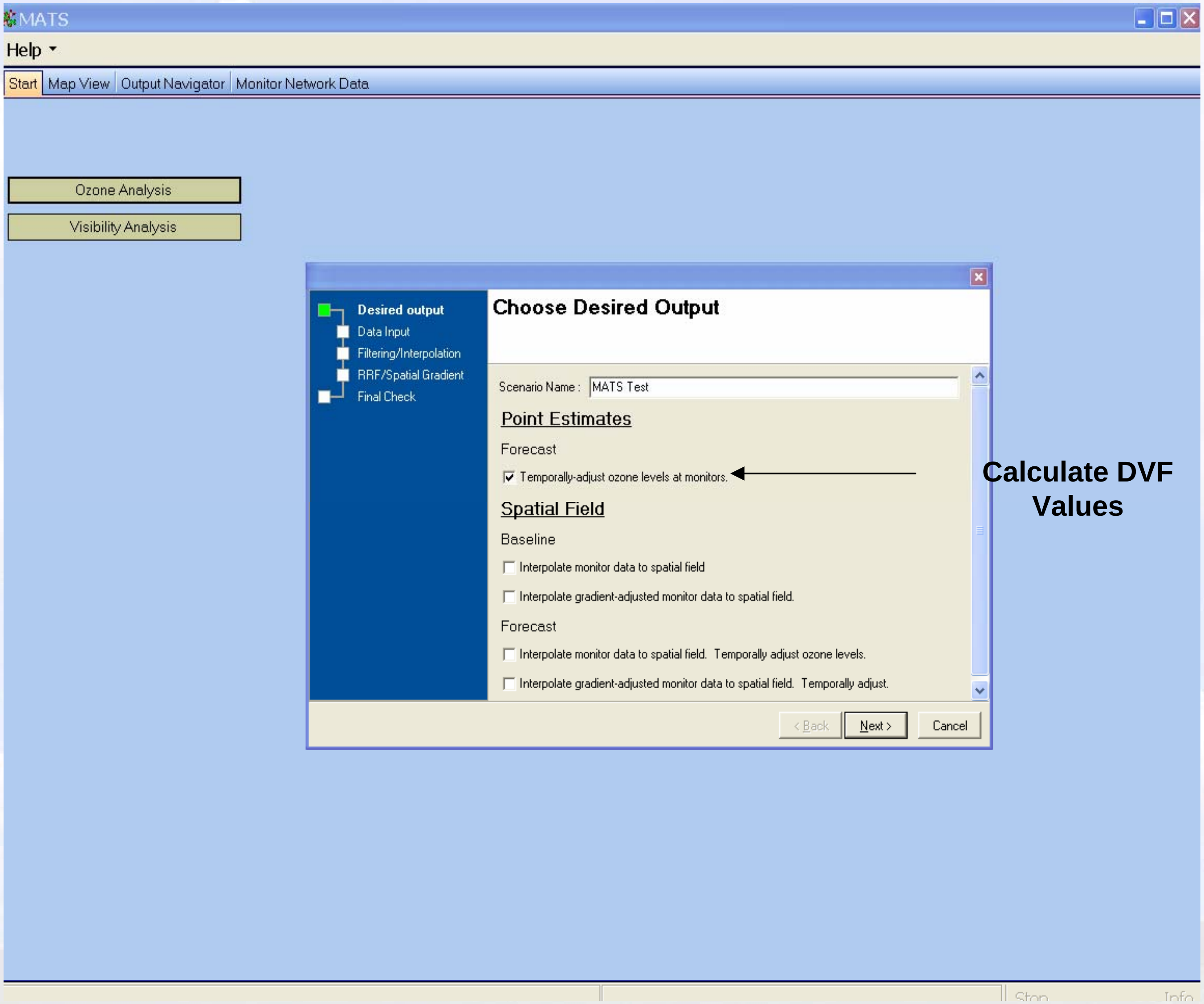
- $$RRF = \frac{\text{Average modeled future case concentration}}{\text{Average modeled baseline concentration}}$$

* maximum baseline and future case modeled concentrations within nearby grid cells

Modeled Attainment Test Software (MATS)

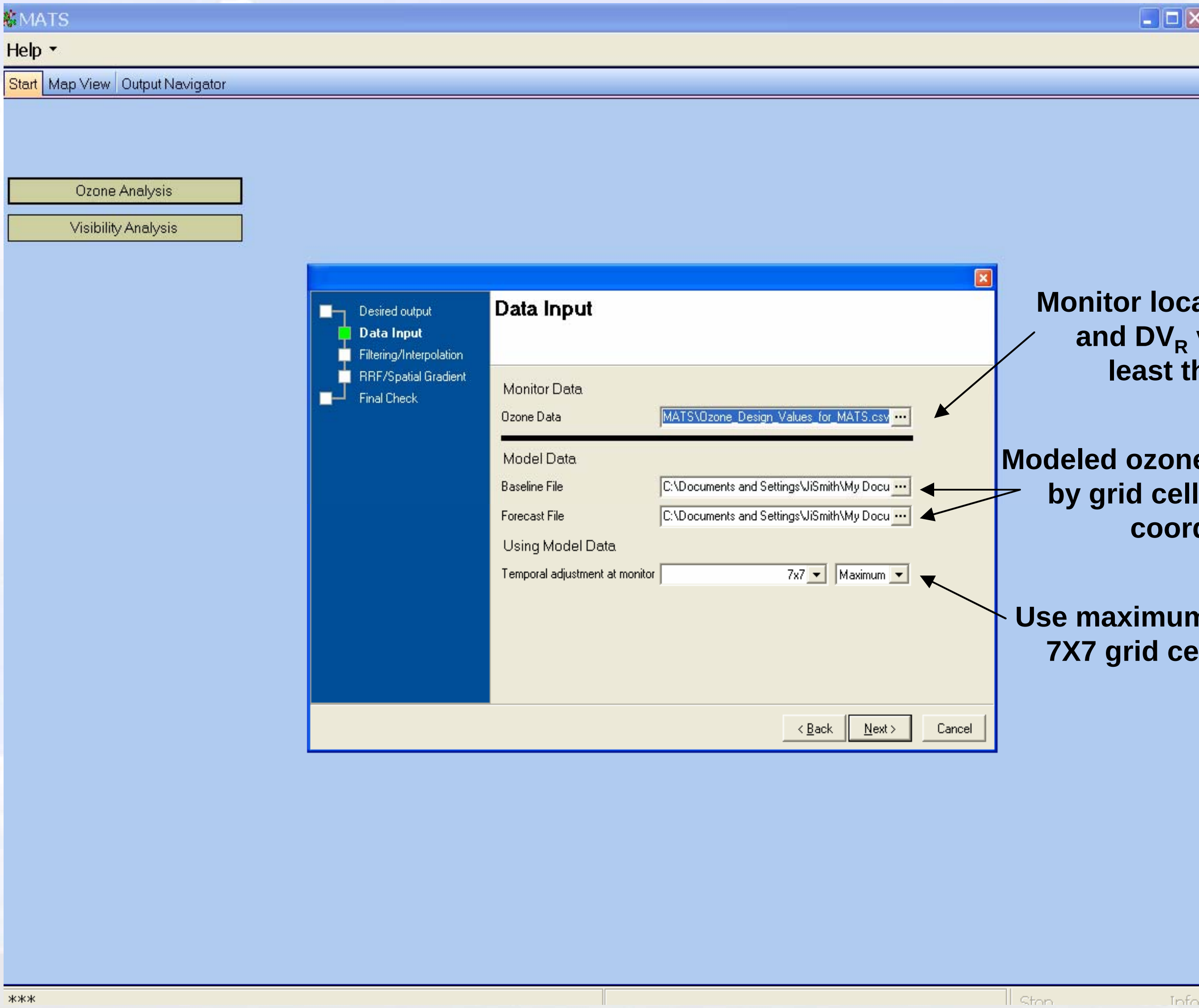
- MATS is provided by EPA to help states use model output in their attainment demonstrations
  - Performs RRF and DV_F calculations
  - Performs an "unmonitored area analysis" (see Dave Westenbarger's SIM presentation from March 20, 2008).
- MATS is a Windows-based interactive program which supports a limited number of choices in performing RRF/DV_F calculations:
  - Lets you choose the size of the area around the monitor from which to pick 8-hour ozone maximum concentration
  - Lets you try different thresholds for selecting days (T_1 and T_2).

MATS



Calculate DVF Values

MATS

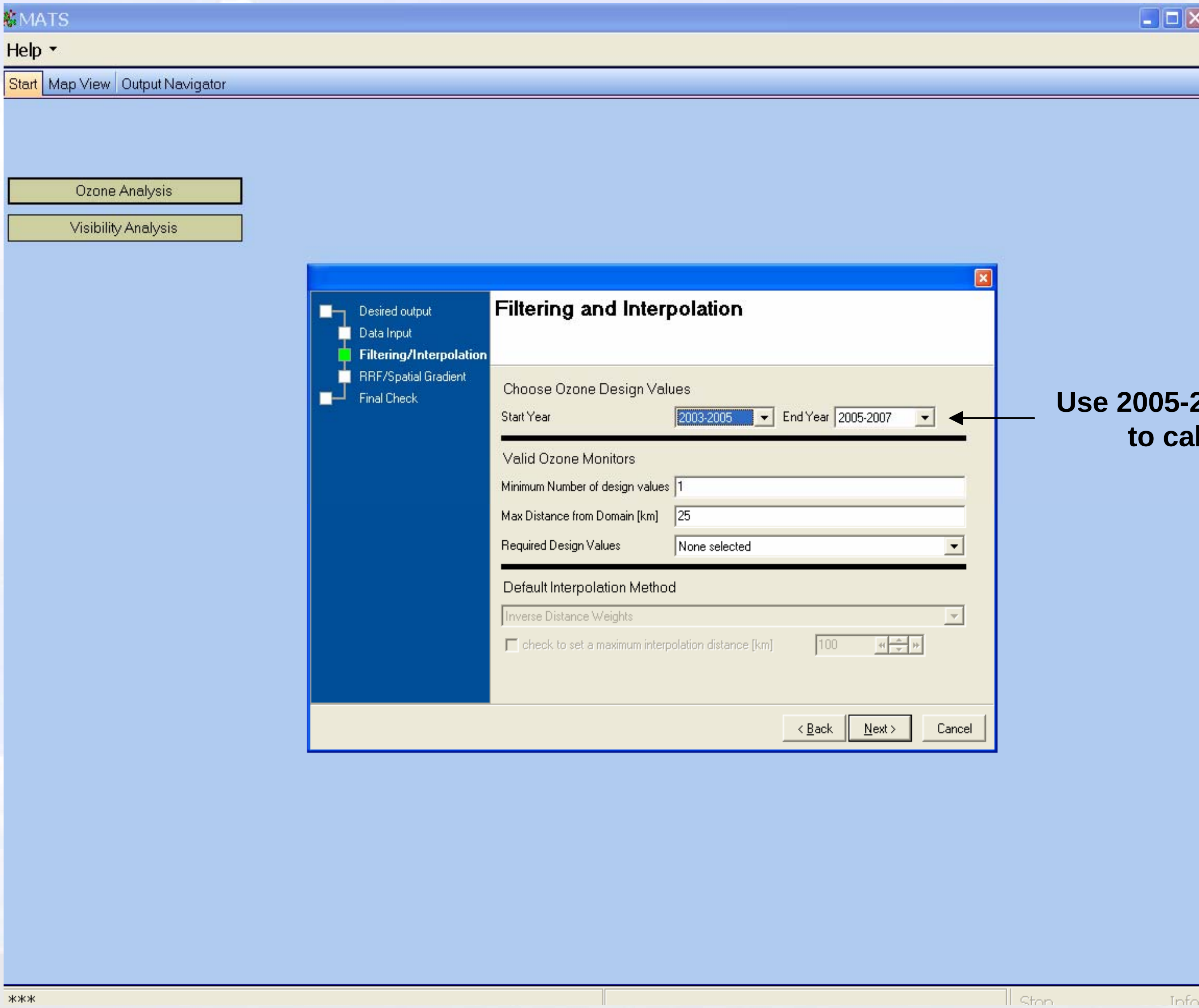


Monitor locations (lat/long) and DV_R values for at least three years

Modeled ozone concentrations by grid cell, with lat/long coordinates

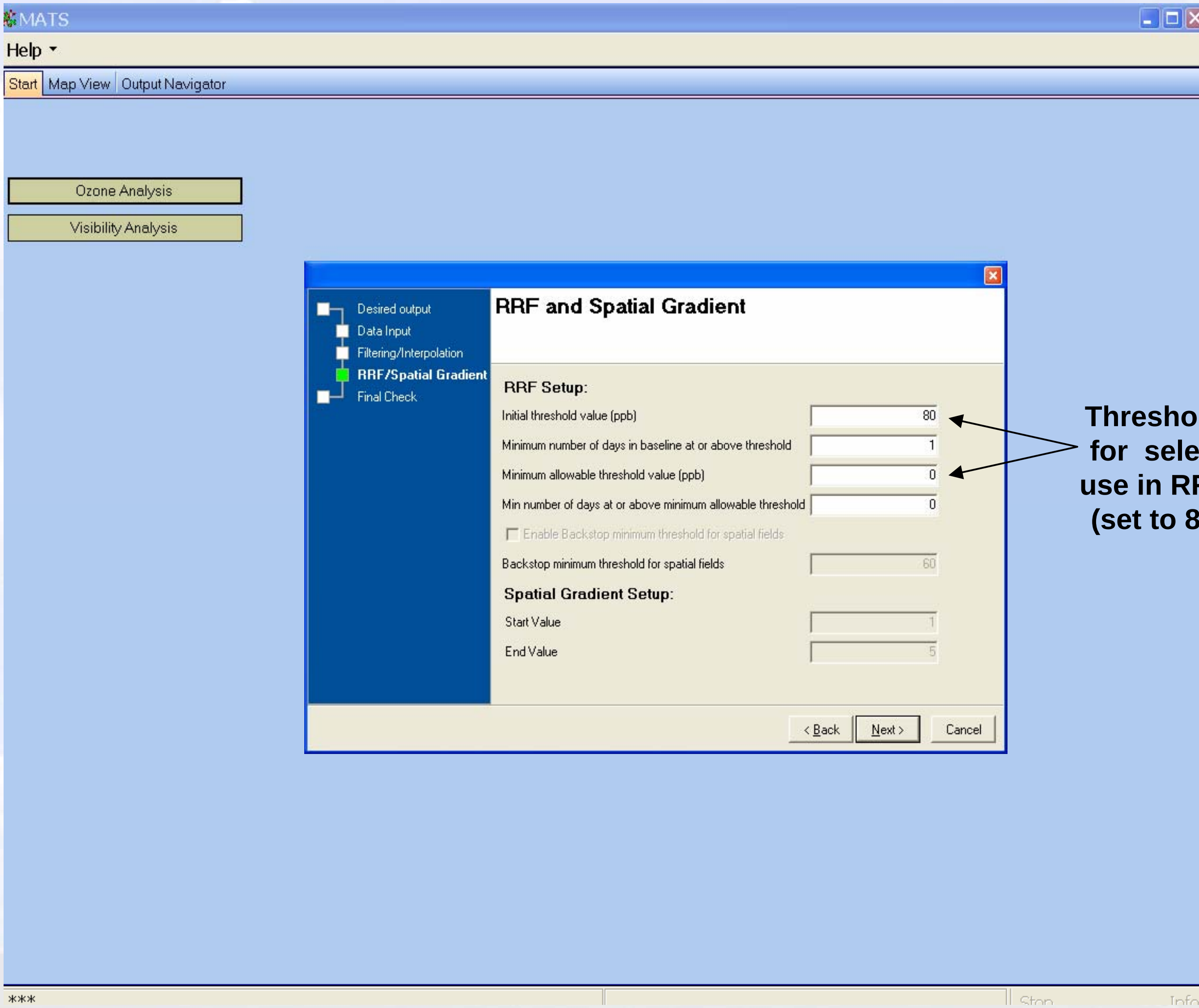
Use maximum concentration in 7X7 grid cell neighborhood

MATS



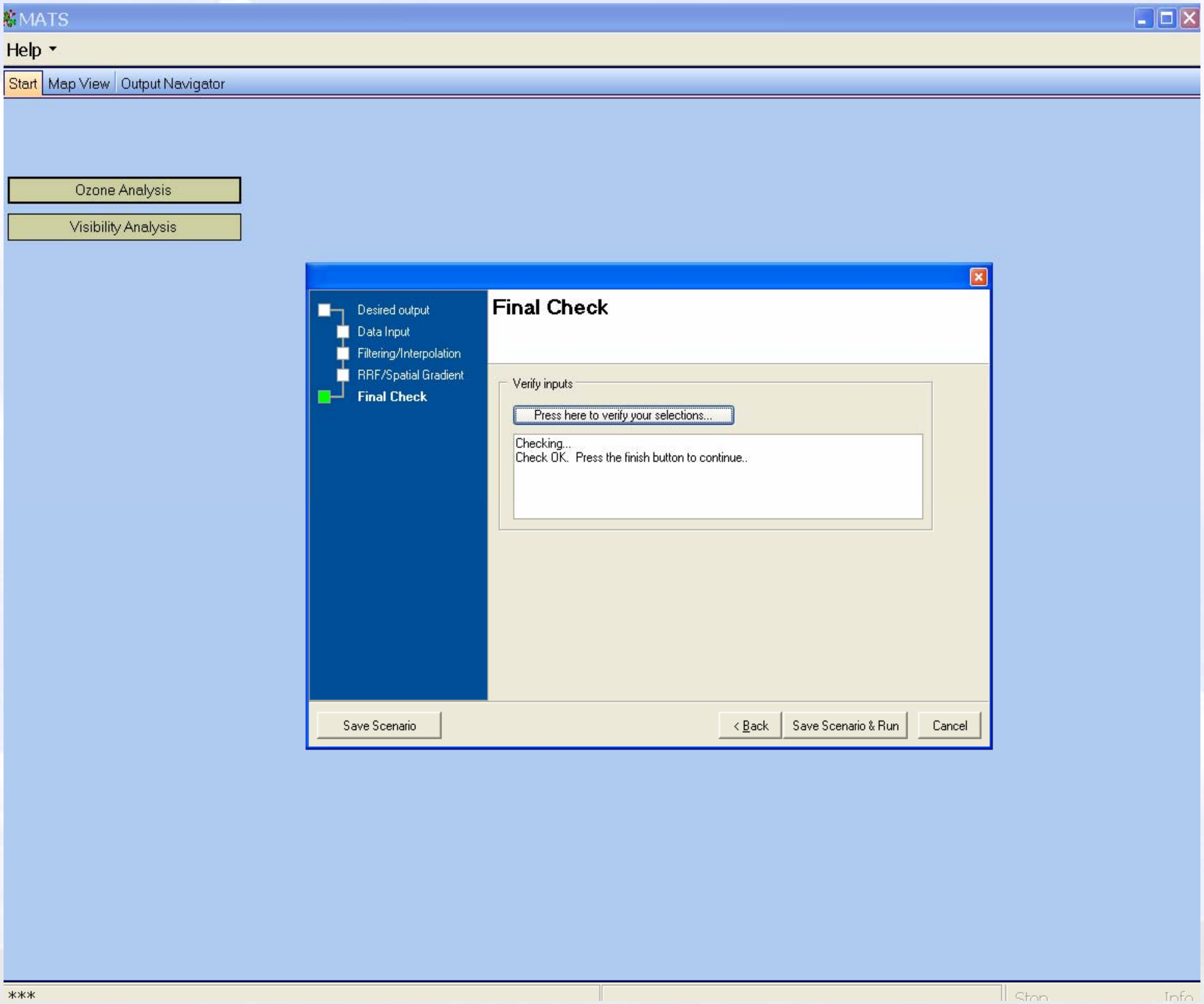
Use 2005-2007 DV_R values to calculate DV_B

MATS

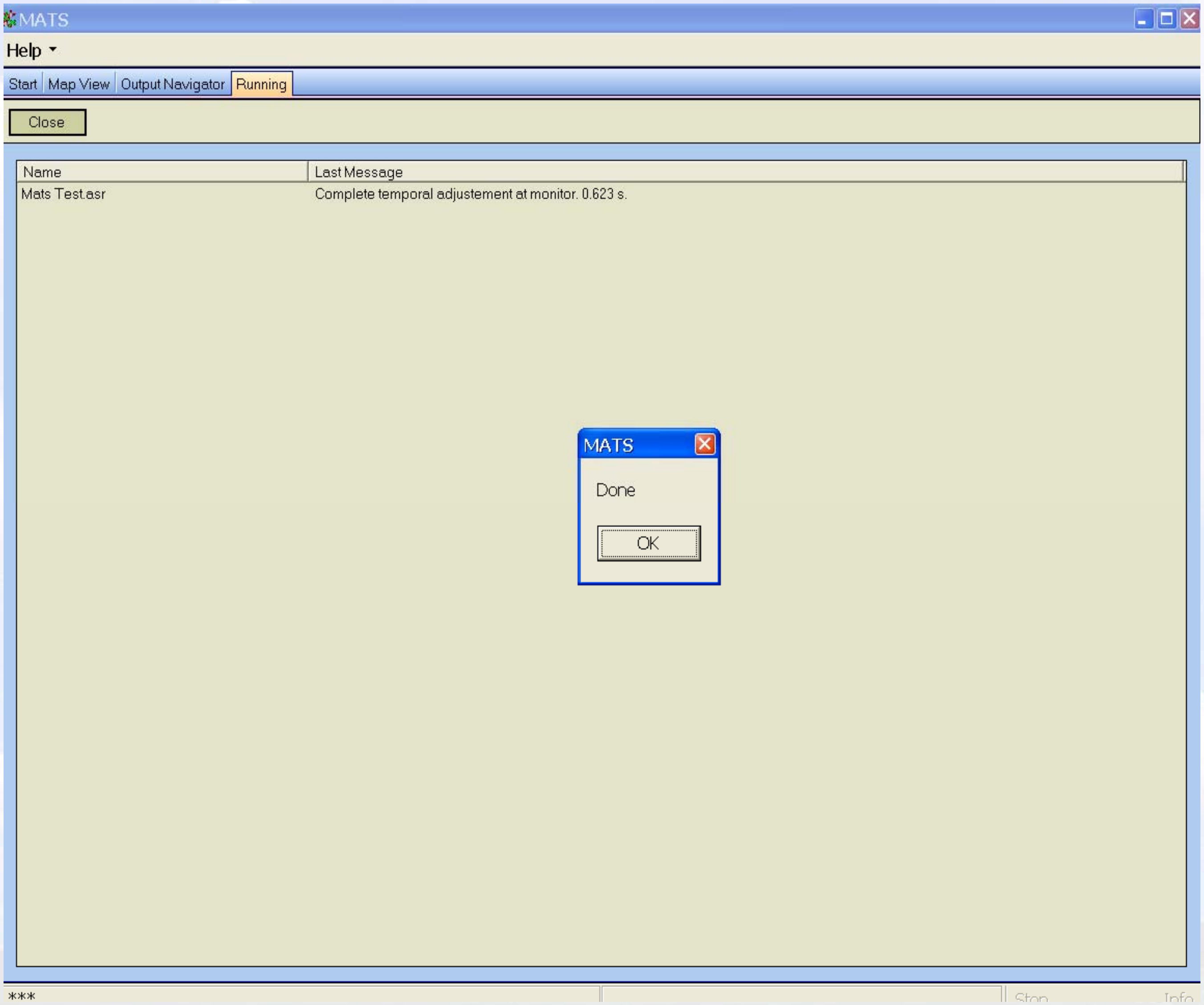


Thresholds T1 and T2 for selecting days to use in RRF calculation (set to 80 & 0 for test)

MATS


MATS
Help
Start
Map View
Output Navigator
Ozone Analysis
Visibility Analysis
Desired output
Data Input
Filtering/Interpolation
RRF/Spatial Gradient
Final Check
Final Check
Verify inputs
Press here to verify your selections...
Checking...
Check OK. Press the finish button to continue..
Save Scenario
< Back
Save Scenario & Run
Cancel
***

MATS


MATS
Help
Start
Map View
Output Navigator
Running
Close
Name
Last Message
Mats Test.asr
Complete temporal adjustement at monitor. 0.623 s.
MATS
Done
OK
***

MATS

MATS

Help

Start | Map View | Output Navigator | Monitor Network Data

Close

Show All or select a particular location to see data.

id	type	lat	long
480390618		29.1488889	-95.765
480390619		29.3136111	-95.2013889
480391004		29.5202778	-95.3925
480391016		29.0436111	-95.4727778
481670014		29.26306	-94.85639
481670056		29.4022222	-94.9463889
482010024		29.9011111	-95.3261111

Select Quantities that must be >= 0

- b_o3_dv
- f_o3_dv
- referencecell
- rrf
- ppb
- days

Export Export this data to CSV

Data

id	date	b_o3_dv	f_o3_dv	referencecell	rrf	ppb	days
480390618	2007	81.5	69.7	29026	0.8563	79.0	11.0
480390619	2007	88.7	77.7	56037	0.8771	80.0	14.0
480391004	2007	94.7	81.9	46048	0.8656	80.0	33.0
480391016	2007	77.5	67.6	43021	0.8731	80.0	12.0
481670014	2007	85.0	78.3	73035	0.9217	80.0	20.0
481670056	2007	87.7	80.4	68043	0.9179	80.0	21.0
482010024	2007	88.0	77.7	48069	0.8835	80.0	22.0
482010026	2007	85.7	80.4	58064	0.9391	80.0	21.0
482010029	2007	91.7	77.3	31076	0.844	80.0	22.0
482010046	2007	78.7	71.0	50065	0.9026	80.0	21.0
482010047	2007	78.7	67.5	40065	0.8581	80.0	30.0
482010051	2007	93.0	79.6	41053	0.8561	80.0	32.0
482010055	2007	100.7	86.7	40057	0.8619	80.0	33.0
482010062	2007	95.3	85.2	51054	0.8944	80.0	35.0
482010066	2007	92.3	77.5	33058	0.8406	80.0	30.0
482010070	2007	84.3	77.4	49060	0.9193	80.0	25.0
482010075	2007	83.3	76.0	47061	0.9134	80.0	24.0
482010416	2007	-7.00	-9.00	50057	0.9079	80.0	30.0
482010617	2007	94.0	86.9	64066	0.9254	80.0	15.0
482010803	2007	88.0	82.1	55062	0.9332	80.0	25.0
482011015	2007	89.0	84.3	60062	0.9482	80.0	20.0

MATS DV_F vs. TCEQ PERL Script DV_F

	MATS			TCEQ		
Monitor*	2005 DVb	RRf	DVf	2005 DVb	RRf	DVf
Danciger	81.5	0.8504	69.3	81.5	0.85	69.305
Houston Aldine	88	0.8835	77.7	88	0.884	77.751
Channelview	85.7	0.9391	80.4	85.67	0.939	80.45
Northwest Harris Co.	91.7	0.844	77.3	91.67	0.841	77.105
Houston Bayland Park	100.7	0.8619	86.7	100.67	0.862	86.768
Houston Monroe	95.3	0.8944	85.2	95.33	0.894	85.261
Wallisville Road	94	0.9254	86.9	94	0.925	86.983
HRM-3	88	0.9332	82.1	88	0.933	82.124
Lynchburg Ferry	89	0.9482	84.3	89	0.947	84.252
Houston East	82.7	0.9223	76.2	82.67	0.922	76.217
Clinton	86.3	0.9252	79.8	86.33	0.925	79.869
Hou.DeerPrk	96.3	0.921	88.6	96.33	0.921	88.716
Conroe Relocated	85	0.8554	72.7	85	0.855	72.706

* Selected HGB Monitors

Why the Difference?

- While MATS and PERL script results are very close, the difference of .003 in the RRF at NW Harris County is too big to attribute solely to round-off error.
- Further investigation showed MATS used 22 days in its RRF calculation, while PERL script used 21 (21 is correct). MATS actually showed other monitors in the wrong grid cells.

Calculation Differences

	MATS					TCEQ					
Monitor	2005 DVb	RRf	DVf	reference cell	days	2005 DVb	RRf	DVf	Grid X	Grid Y	N Days
Danciger	81.5	0.8504	69.3	29026	9	81.5	0.85	69.305	29	27	9
Houston Aldine	88	0.8835	77.7	48069	22	88	0.884	77.751	48	70	22
Channelview	85.7	0.9391	80.4	58064	21	85.67	0.939	80.45	58	65	21
Northwest Harris Co.	91.7	0.844	77.3	31076	22	91.67	0.841	77.105	31	76	21
Houston Bayland Park	100.7	0.8619	86.7	40057	33	100.67	0.862	86.768	41	58	33
Houston Monroe	95.3	0.8944	85.2	51054	35	95.33	0.894	85.261	52	55	35
Wallisville Road	94	0.9254	86.9	64066	15	94	0.925	86.983	65	66	15
HRM-3	88	0.9332	82.1	55062	25	88	0.933	82.124	56	63	26
Lynchburg Ferry	89	0.9482	84.3	60062	20	89	0.947	84.252	61	63	21
Houston East	82.7	0.9223	76.2	53062	24	82.67	0.922	76.217	54	63	27
Clinton	86.3	0.9252	79.8	52060	27	86.33	0.925	79.869	52	61	27
Hou.DeerPrk	96.3	0.921	88.6	58057	27	96.33	0.921	88.716	59	57	27
Conroe Relocated	85	0.8554	72.7	42094	10	85	0.855	72.706	42	94	10

Analysis and Possible Explanation

- Since MATS is a “Black Box”, it’s difficult to know for sure why the RRF/DV_F calculations differ subtly from TCEQ’s.
- Best guess is that discrepancy is due to using different map projections:
  - We conduct our modeling on a Lambert Conformal Grid, but
  - MATS requires input in Latitude/Longitude.
- Since the map projections are tilted a few degrees with respect to each other, it seems likely that some monitors might “move” when converting from LCP to Lat/Long, causing different sets of model output to be used in MATS’ calculations.

MATS Advantages

- MATS is easy to use and runs quickly.
- It is EPA's preferred tool for performing the unmonitored area analysis.
- MATS threshold values can be easily manipulated to test various combinations of values.

MATS Disadvantages

- MATS is Windows-based and cannot easily be incorporated into our LINUX-based runstream; it has to to be run separately on a Windows-based PC.
- MATS requires input in latitude and longitude which means extra processing of model output.
- MATS is not set up to handle baseline modeling for multiple years (2005 and 2006 in our case). It had to be "tricked" to work for us.
- MATS is a "black box" – it's impossible to know what it's really doing without some serious detective work.
- MATS gives close (but not exact) approximations of the true RRF and DV_F values.