Developing Draft Emission Rates for MOVES2006:

Criteria pollutant emissions from Light-Duty Vehicles

FACA Modeling Workgroup Meeting
August 8, 2006

Assessment and Standards Division
EPA Office of Transportation & Air Quality

MOVES

Outline

- **Recap MOVES design**
- **Review Data Sources**
  - MOVES 2004
  - Additional for MOVES2006
    - I/M program data
    - Remote Sensing
- **Options to represent I/M**
- **Addressing Measurement Error**
  - Example: preconditioning
  - Approach and example: time-series alignment
- **Model-year by Age Coverage**
- **Next Steps**

MOVES

Emissions Source
(*SourceBin*)

- **MOVES classifies light-duty vehicles on basis of**
  - Fuel type
    - Gasoline, diesel, CNG, LPG, Ethanol, Hydrogen
  - Engine Technology
    - "conventional" internal combustion
  - Regulatory Class
    - LDV, LDT
  - Model Year Group
    - 36 groups covering 1960 - 2050

ModelYearGroups (for Analysis…)

Pre Tier 1 … ***And Post Tier 1 …***

MY	LDV & LDT
1979	
1980	
1981	3-way catalysts
1982	
1983	I/M programs
1984	
1985	
1986	
1987	
1988	
1989	
1990	CAA Amendments
1991	
1992	
1993	

MY	LDV	LDT			
		LDT1	LDT2	LDT3	LDT4
1994	FTP Tier 1 Certification				
1995					
1996	FTP Tier 1 Final In-Use			FTP Tier 1 Certification & Interim In-Use	
1997					
1998				FTP Tier 1 Final In-use	
1999	FTP NLEV begins				
2000					
2001	SFTP NLEV begins				
2002					
2003					
2004	Tier 2 Requirements take effect				
2005					
2006					
2007					

Vehicle Age
(*Deterioration*)

- **MOVES estimates deterioration by providing different rates for seven age groups**
  - 0-3 years,
  - 4-5 years,
  - 6-7 years,
  - 8-9 years,
  - 10-14 years,
  - 15-19 years,
  - 20 + years

Operating Mode

- **MOVES uses distinct emission rates for different "modes" of driving**
- **Modes are defined in terms of vehicle speed, acceleration**
- **Rates estimated for 23 modes:**
  - Deceleration/braking (1 mode)
  - Idle (1 mode)
  - Coast (2 modes)
  - Cruise/acceleration (19 modes)

Operating Modes for Running Exhaust Emissions

VSP Class (kW/tonne)	Speed Class (mph) 1-25	25-50	50 +
30 +	16	30	40
27-30			
24-27		29	39
21-24		28	38
18-21			
15-18			37
12-15		27	
9-12	15	25	
6-9	14	24	35
3-6	13	23	
0-3	12	22	33
< 0	11	21	

***For coast and cruise*,**

13 modes retained from MOVES2004, *plus*

8 modes added for MOVES2006.

PLUS

One mode each for
idle, and
decel/braking

Gives a total of
23 opModes

Overview: Data and Methods

		Gasoline	Diesel	CNG	LPG	EtOH	H_2
"Conventional"	MY 1960 to 2001	SBS dynamometer fill holes by regression I/M	no SBS data ratio to gasoline rates				
	MY 2002 to 2050	Ratio to Applicable Standards		Ratio to Gasoline Rates			

Topic of this presentation

MOVES

Scope

- **Goal: develop emission rates**
  - Process: *running exhaust*
  - Pollutants: *CO, THC, NOx*
  - Emissions Sources:
    - Fuel: *gasoline*
    - Regulatory Classes: *LDV, LDT*
    - Model years: 1960 – 2001
  - Temperature range (68-86 deg. F)
  - Geographic scope: national default averages
  - representing I/M and non I/M areas
  - Calendar Years: 1990, 2002 +

- **Questions:**
  - What data is available and suitable?
  - How combine dynamometer and RSD measurements?

Data Sources: MOVES2004

Program	Sponsor	#vehicles
New York Instrumentation/Protocol Assessment	NYSDEC	6,661
Basic Emission Rates on Multiple Schedules	NVFEL	62
Basic Emission Rates on Multiple Schedules (with AC on/off)	NVFEL	35
Opacity and Exhaust Emissions in Gasoline Vehicles during an IM240	NVFEL	104
Inventory Cycles/LA92 Exhaust Emissions	NVFEL	20

(*plus additional programs in OTAQ's Mobile-Source Observation Database*)

Issues with reliance on these data for MOVES2006:
-Limited coverage of model-year and age combinations
-heavy reliance on single large program in one region (NE)?
-Presence of vehicles certified to CA standards?
-Mixture of vehicles representing I/M and non I/M conditions?

Additional Data for MOVES2006?

Dynamometer

I/M
(~36 million tests

AZ (Phoenix)
IL (Chicago)
MO (St. Louis)
British Columbia
CO (Denver)
Indiana
Ohio
Wisconsin

Non I/M

Kansas City (?)
Other MSOD

Remote Sensing (RSD)

I/M
(~15 million points)

AZ (Phoenix)
IL (Chicago)/ N. IN
MO (St. Louis)

Maryland/ N. Va
Los Angeles
Houston
Atlanta

Non I/M

VA (Richmond)

Augusta/Macon
Omaha
Tulsa

Pros and Cons

Lab vs. I/M vs. RSD

	Lab	I/M	RSD
High Emitters		☺	☺ ☺
Controlled Conditions	☺		
Sample Size		☺	☺
Operating Range	☺		
Data Points per Vehicle	☺ ☺	☺	
MY X Age coverage		☺ ☺	☺
Covers non I/M areas	☺		☺ ☺

I/M Program Data
MO (St. Louis)

- Type: **centralized on biennial cycle**
- Exemption: **2 most recent MY**
- Tests:
  - MY 1980 and earlier: curb idle
  - MY 1981 to 1995 IM240
  - MY 1996 and later OBDII
- Screening?: **Clean screen (RSD)**
- Fast Pass: **at 31 sec or later**
- Fast Fail: **NO**
- We have: **CY 2000 to 2005**
  - Appr. 2 million tests
- Random sample: **NO**
- Back to Back sample: **NO**

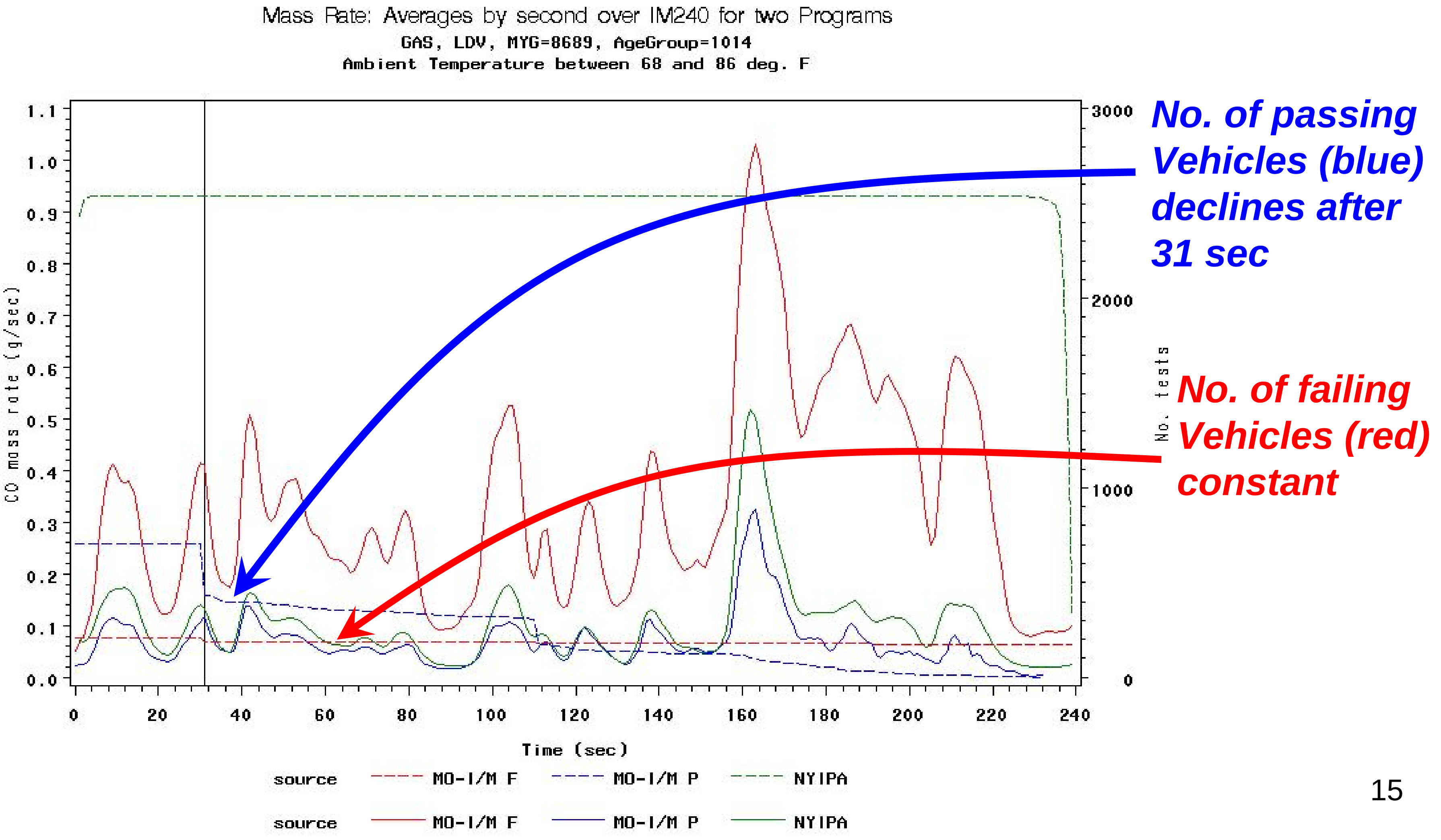
I/M Program Data
MO (St. Louis)

- **Issues:**
  - Clean Screen Bias
  - OBD Screen Bias
  - Fast-pass bias
  - Lack of preconditioning (?)
  - Misalignment of time series

- **CONCLUSION:**
  - Exclude for MOVES2006
  - REASON: lack of preconditioning compounded by fast-pass bias

Example: Fast-Pass Bias (MO I/M) Vehicle Representation over I/M240

Gasoline, LDV, MY 1986-89, 10-14 years of age



MOVES

I/M Program Data
AZ (Phoenix)

- Type: **centralized on biennial cycle**
- Exemption: **4 most recent MY**
- Tests:
  - MY 1980 and earlier: loaded idle
  - MY 1981 to 1995 IM240 (through 1999)
  - IM147 (since 2000)
  - MY 1996 and later OBDII
- Screening?: **NO**
- Fast Pass: **at 31 sec or later**
- Fast Fail: **YES**
- We have: **CY 1994 to 2004**

- Random sample: **thru 1999, “2% of fleet”**
- **since 2002 “stratified random in triplicate”**
- Back to Back sample: **NO**

I/M Program Data
AZ (Phoenix)

- **Issues:**
  - OBD Screen Bias (if using program data)
  - Fast-pass/ Fast-fail bias (if using program data)
  - Lack of preconditioning (?)
  - Misalignment of time series
  - "Analyzer saturation," a.k.a. "plateaus"
- **CONCLUSION:**
  - To include for MOVES2006
  - APPROACH: use random sample
  - select duplicate or triplicate tests
  - differential weighting for pass/fail results

I/M Program Data
IL (Metro Chicago)

- Type: **centralized on biennial cycle**
- Exemption: **4 most recent MY**
- Tests:
  - MY 1980 and earlier: curb idle
  - MY 1981 to 1995 IM240
  - MY 1996 and later OBDII
- Screening?: **NO**
- Fast Pass: **at 31 sec or later**
- Fast Fail: **NO**
- We have:
- Random sample: **“2% of fleet”**
- Back to Back sample: **YES**

I/M Program Data
IL (Metro Chicago)

- **Issues:**
  - OBD Screen Bias (if using program data)
  - Fast-pass bias
  - Lack of preconditioning (?)
  - Misalignment of time series

- **CONCLUSION:**
  - To include for MOVES2006
  - APPROACH: combine random and B2B samples
  - select duplicate tests

MOVES

Program Data

New York Instrumentation/ Protocol Assessment (NYIPA)

- Type: **within NY metro area (decentralized)**
- Exemption: **N/A**
- Tests: **IM240**
- Screening?: **N/A**
- Fast Pass: **N/A**
- Fast Fail: **N/A**
- We have: **CY 2000 to 2004 (~5,500 tests)**
- Random sample: **???**
- Back to Back sample: **???**

Program Data
New York Instrumentation/ Protocol Assessment (NYIPA)

- **Issues:**
  - Random recruitment?
  - NY adopted CA stds in 1999
    - are NYIPA vehicles nationally representative?
  - Lack of preconditioning (?)
  - Misalignment of time series
- **CONCLUSION:**
  - Include for MOVES2006
  - APPROACH: ID replicate B2B tests?
  - evaluate recruitment
  - evaluate influence of CA stds?

Summary: Sources and Issues Dynamometer Data

Issue		Dataset			Approach
		NYIPA	AZ I/M	IL I/M	
Measurement Error					
	lack of pre-conditioning	???	R	R	use replicate B2B tests
	implausibly high fuel economy	???	R	???	outlier screening
	mis-alignment of time series	R	R	R	statistical alignment
	"analyzer saturation," "plateaus"	???	R?	???	exclude extreme cases
Non-representativeness					
	influence of CA LEV program?	???	n/a	n/a	
	non-random recruitment	???	R	R?	use random samples
	fast-pass/fast-fail bias	n/a	R	R	use random samples
	clean-screen bias	n/a	n/a	n/a	
	OBD -screen bias	R	R	R	use random samples

MOVES

Data Sources: Remote Sensing
(Approximate *no. measurements*)

Location	I/M	Non I/M
St Louis (2004)	2,000,000	
Virginia (2003-05)	210,000	79,000
Colorado (2004-05)	280,000	14,000
Georgia (2004)	170,000	
Illinois (2003-05)	165,000	4,800
Indiana (2003-05)	49,000	56,000
Tulsa, OK		19,000
Omaha, NE		18,000

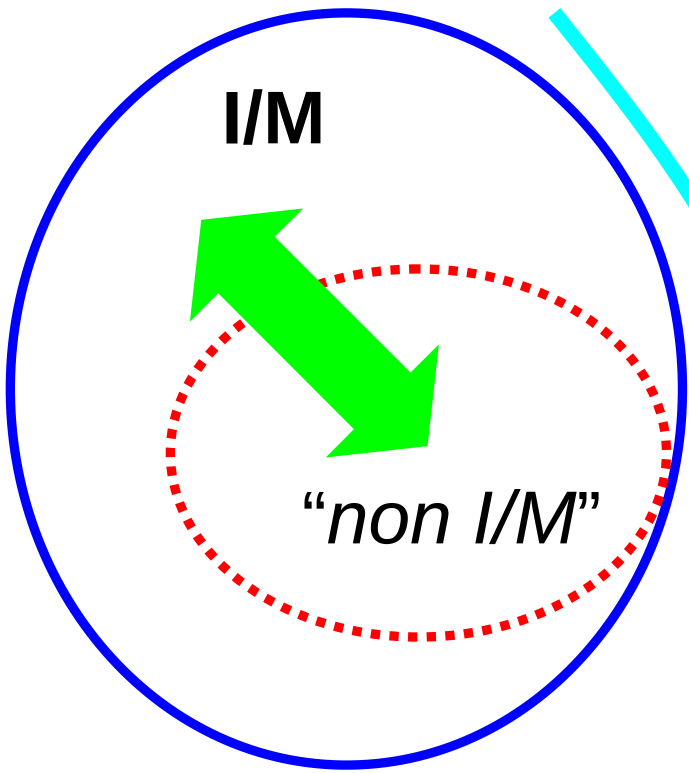
RSD: Processing

- **Convert concentration ratios to mass ratios**
  - Multiply by ratio of molecular weights
- **Apply aromatic adjustment factor (THC only)**
  - RSDs detect only alkanes
  - Adjustment accounts for missing aromatics
- **Restrict to temperature range (68-86 deg. F)**
- **Use CO_2 rates to estimate mass rates**
  - Can’t generate mass rates independently
- **Calculate VSP for each measurement**
  - Necessary to fit into MOVES opModes
  - Necessary to compare to dynamometer data

Options to Represent I/M Programs?

- **Adopt simplified approach ?**
  - Develop two sets of emission rates
    - I/M and non I/M
  - Modify by program fraction
    - To represent program differences
  - Default condition to be “I/M”
    - Adjust to “non-I/M”

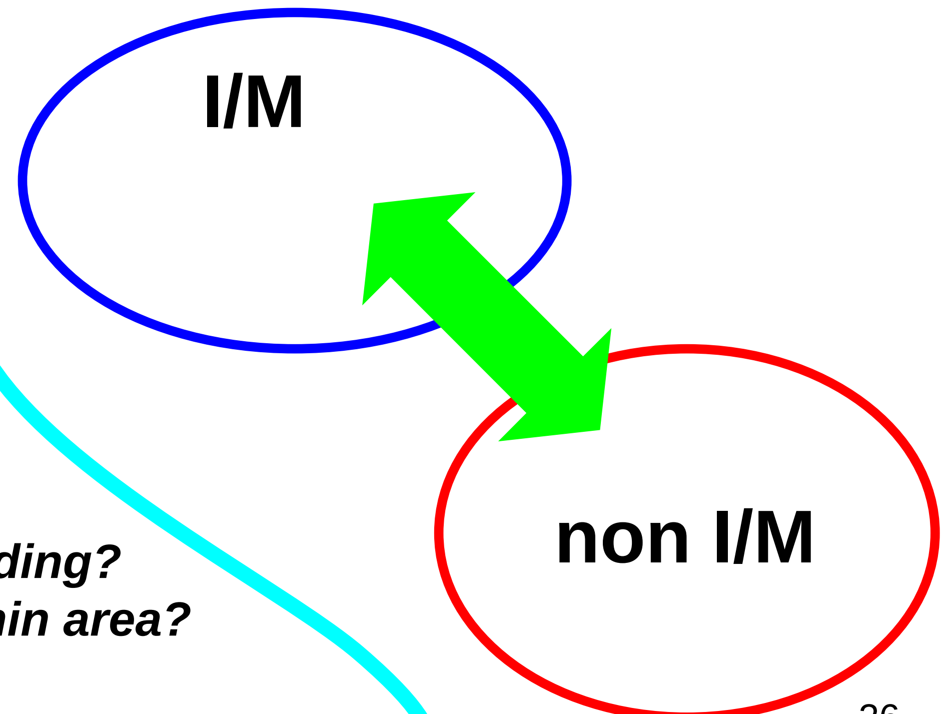
How represent I/M difference?



“Between Program Areas”

PRO: ***May be only option?***

CON: ***prone to confounding?***
difficult to match areas, interpret differences?



“Within Program Area(s)”

PRO: ***free from extraneous confounding?***

CON: ***Can ID “non I/M” vehicles within area?***

How represent I/M difference? Options for analysis

- **Within-Program**
  - "program evaluation" method
  - "out-of-program" vehicles method
    - Vehicles operating within I/M area that may not be influenced by the program?

- **Between Program**
  - Develop rates independently
    - Generate two sets of rates from separate data sources
  - Develop default rates and adjustments
    - With I/M as default, and
    - Adjust to non I/M
  - Remote-sensing is probable data source

“Program Evaluation” Method

- **Average initial and final Tests within each program cycle**
- **Calculate reduction from initial to final**
  - Usually expressed as %
  - Reduction(%) = (initial-final)/initial
- **ISSUES with respect to MOVES2006**
  - Requires use of program data
    - Need to estimate missing portions of partial tests
      - Sec-by-sec modal approach requires full time series
    - Initial and final tests not included in random samples
- **CONCLUSION: adaptations needed to use I/M data in modal framework probably preclude this approach**

Identifying Out-of-Program Vehicles? Example: MO (St. Louis)

- **Can we use license plate information?**
- **Two Sources:**
  - MO Dept. Revenue provides to MO DNR
  - Contractor collects in I/M Lane
- **Criteria (*vehicle may be O-of-P if*…)**
  - 1) If “Dept. Rev.” plate not from Program Area?
  - 2) If “I/M Lane” plate is out-of-state
  - 3) If “Dept. Rev.” is MISSING and “I/M” is present?

Identifying "Out-of-Program" Vehicles?
Example: MO (St. Louis)

Book-keeping for sources of information on license plates:

	Dept. Revenue	in I/M Lane
Within Area	134,015	211,020
outside of area	46	
out of State		35
MISSING	116,850	39,856
TOTAL	250,911	250,911

***1) 0.018% of plates*: TOO LOW**

***2) 0.014% of plates*: TOO LOW**

3) (Dept.Rev. – I/M) / Total ~ 30% of plates: TOO HIGH

Acknowledgment: Phil Heirigs, Sierra Research

Between-Program Comparisons?

Dynamometer

I/M
(~36 million tests

AZ (Phoenix)
IL (Chicago)
MO (St. Louis)
British Columbia
CO (Denver)
Indiana
Ohio
Wisconsin

Non I/M

Kansas City (?)
Other MSOD

Remote Sensing (RSD)

I/M
(~15 million points)

AZ (Phoenix)
IL (Chicago)/ N. IN
MO (St. Louis)

Maryland/ N. Va
Los Angeles
Houston
Atlanta

Non I/M

VA (Richmond)

Augusta/Macon
Omaha
Tulsa

Next Steps

- **General approach (open to debate)?**
  - I/M rates as default
    - Based on dynamometer data
  - Non I/M rates by adjustment
    - Based on RSD
    - With perhaps a role for Kansas City results?

$$E_{\mathrm{non-I/M}} = r\,E_{\mathrm{I/M}}\,?$$

- **Issues**
  - Should I/M differences vary by opMode?
  - Should I/M differences vary
    - by MYG,
    - by ageGroup,
    - or both?
    - RSD has poor coverage of age ranges in MY

High Emitters?
Or High Emissions?

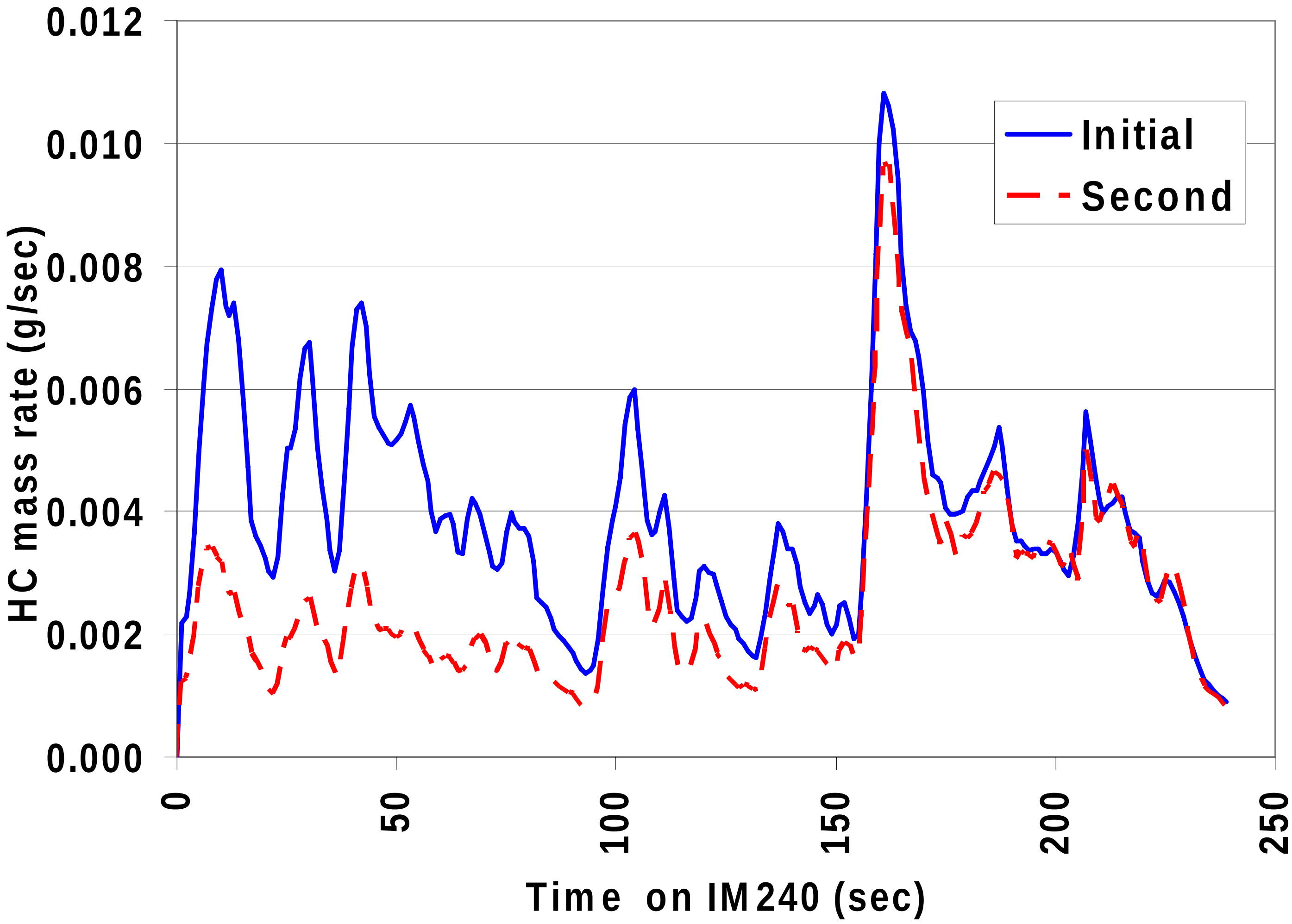
- **Presumption: MOVES will not model a separate class of "high emitters"**
  - Issue still under discussion
- **No available definition that is not arbitrary**
  - distinct population of 'high emitters' not obvious
  - High on all pollutants, every day?
  - High in all operating modes?
- **More rates require more book-keeping in MOVES**
  - Separate set of "hi emitter rates would double number of rates needed
  - Could reduce model performance
- **Current direction: Attempt to derive "representative" rates**
  - Assume emissions represent single distribution with long tail
  - "clean and "dirty" vehicles in "correct" proportions

Pre-conditioning

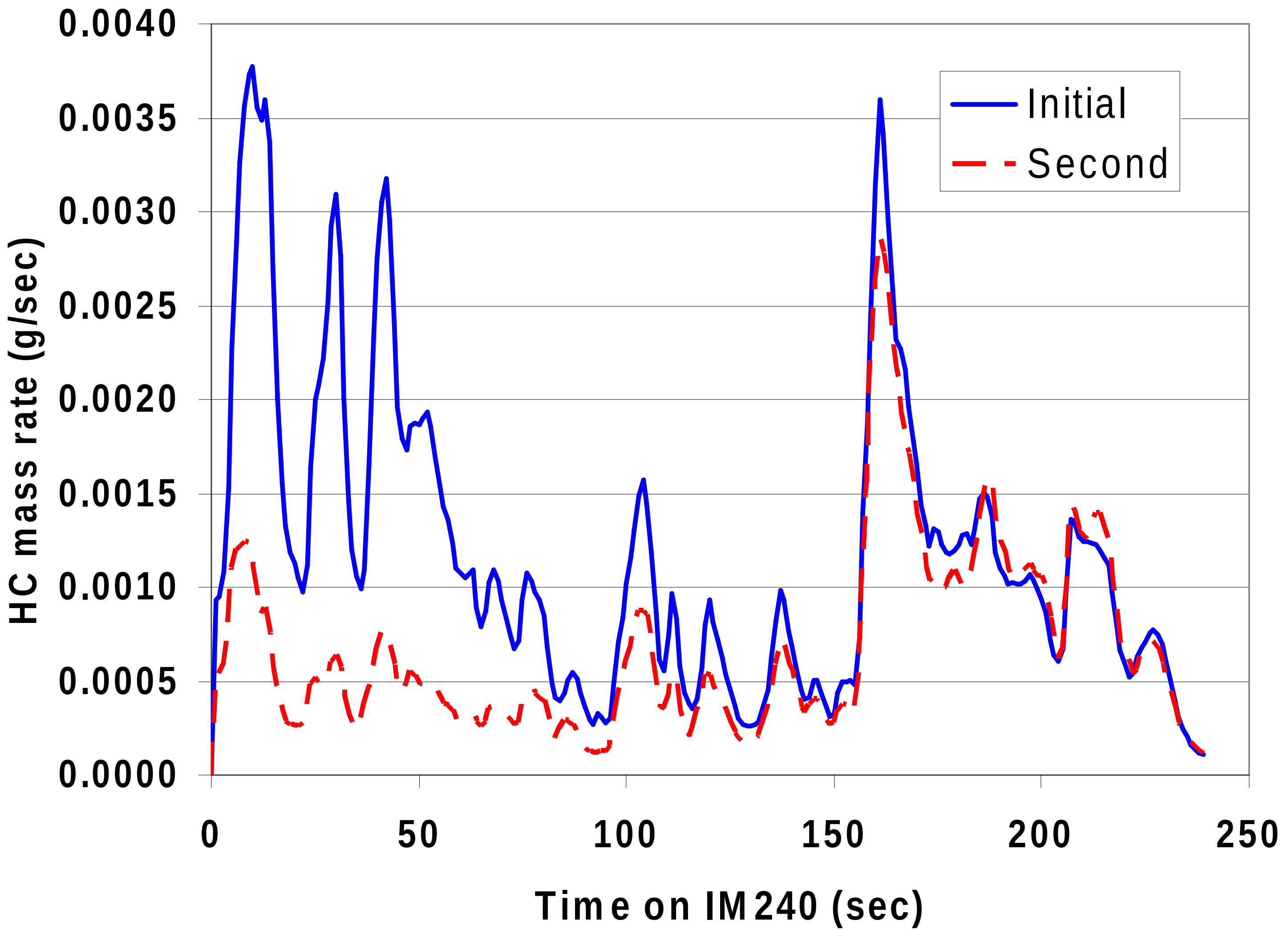
- **Example: data from Metro Chicago**
- **To assess level of conditioning:**
  - Isolate sample of back-to-back tests
  - Average all tests by second over IM240
  - Plot “first” and “second” time series
- **Examine three model years**
  - 1994, 1998, 2000
  - Use THC as example

Acknowledgement: Phil Heirigs, Sierra Research

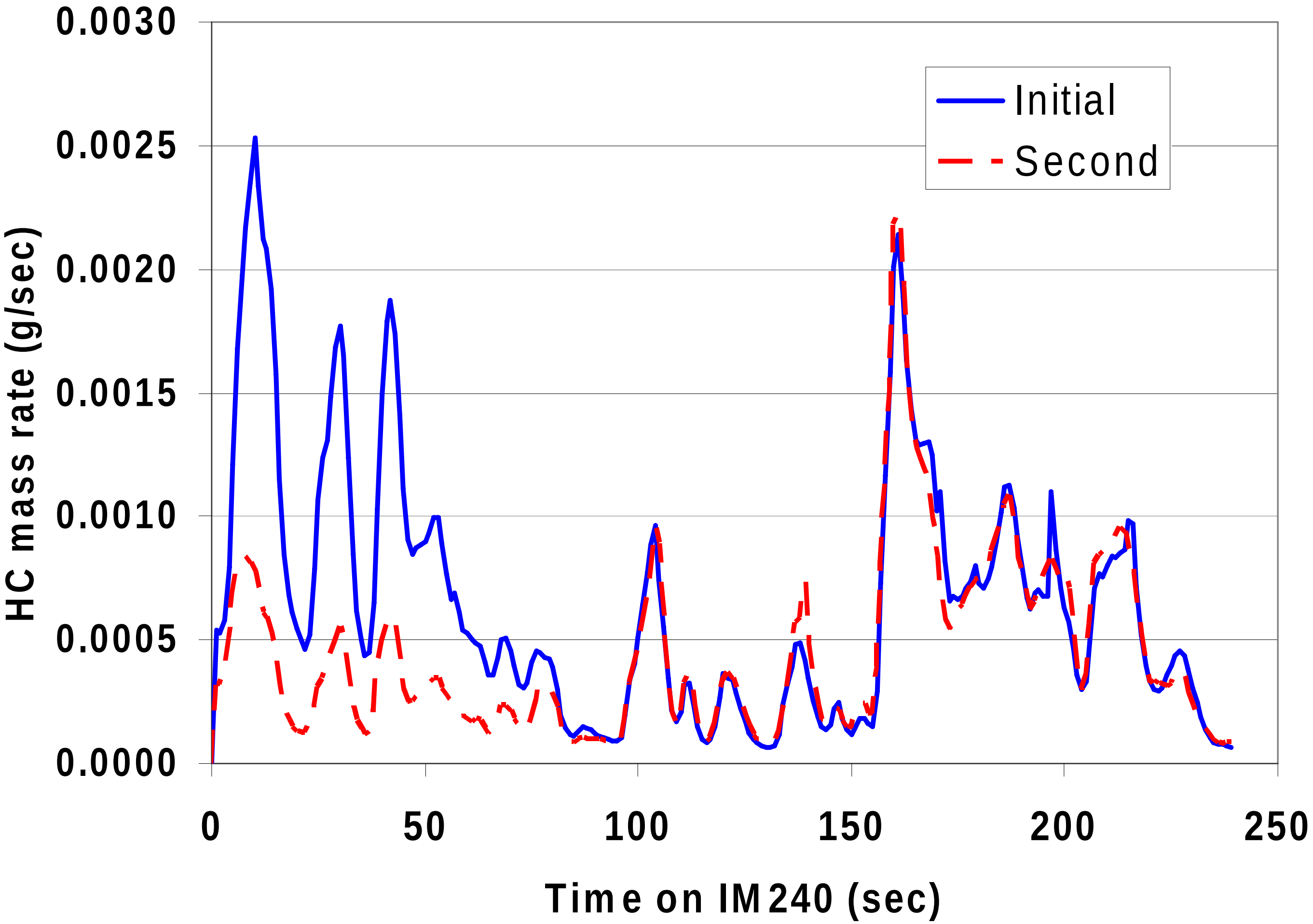
THC time series on IM240 MY1994



THC time series on IM240 MY 1998



THC time series on IM240
MY 2000



Time Series Alignment: Process Example: AZ I/M

- **Steps**
  - Calculate Vehicle-specific power (VSP)
  - Set negative VSP to 0.0
  - Smooth VSP
  - Calculate correlation coefficients at each offset over window of +/- 6 sec
  - Get offsets associated with maximum correlation
  - Apply offset to emissions time-series

Calculate and Smooth VSP…

VSP represents the vehicle's tractive power normalized to its weight, and calculated is a function of velocity, acceleration, weight and the Vehicles road-load coefficients

$$\mathrm{VSP}_t = \frac{Av_t + Bv_t^2 + Cv_t^3 + v_t\, a_t}{m}$$

v = velocity, m/sec
a = acceleration m/sec2
m = weight (tonne)
A = rolling resistance (kW-sec/m)
B = rotating resistance (kW-sec^2/m^2)
C = aerodynamic drag (kW-sec^3/m^3)

Then smooth VSP, using a weighted-centered moving average…

$$\mathrm{VSP}_{wcma,t} = \frac{1}{4}\mathrm{VSP}_{t-1} + \frac{1}{2}\mathrm{VSP}_t + \frac{1}{4}\mathrm{VSP}_{t+1}\ ,\ \mathrm{VSP} \neq 0$$

Offsetting the Time-Series

Offset Emissions against VSP over 6-sec window

-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6									
												1	2	3	4	5	6	7	8	9	10
											1	2	3	4	5	6	7	8	9	10	
										1	2	3	4	5	6	7	8	9	10		
									1	2	3	4	5	6	7	8	9	10			
								1	2	3	4	5	6	7	8	9	10				
							1	2	3	4	5	6	7	8	9	10					
						1	2	3	4	5	6	7	8	9	10						
					1	2	3	4	5	6	7	8	9	10							
				1	2	3	4	5	6	7	8	9	10								
			1	2	3	4	5	6	7	8	9	10									
		1	2	3	4	5	6	7	8	9	10										
	1	2	3	4	5	6	7	8	9	10											
1	2	3	4	5	6	7	8	9	10												

Emissions time-series "slides on VSP"

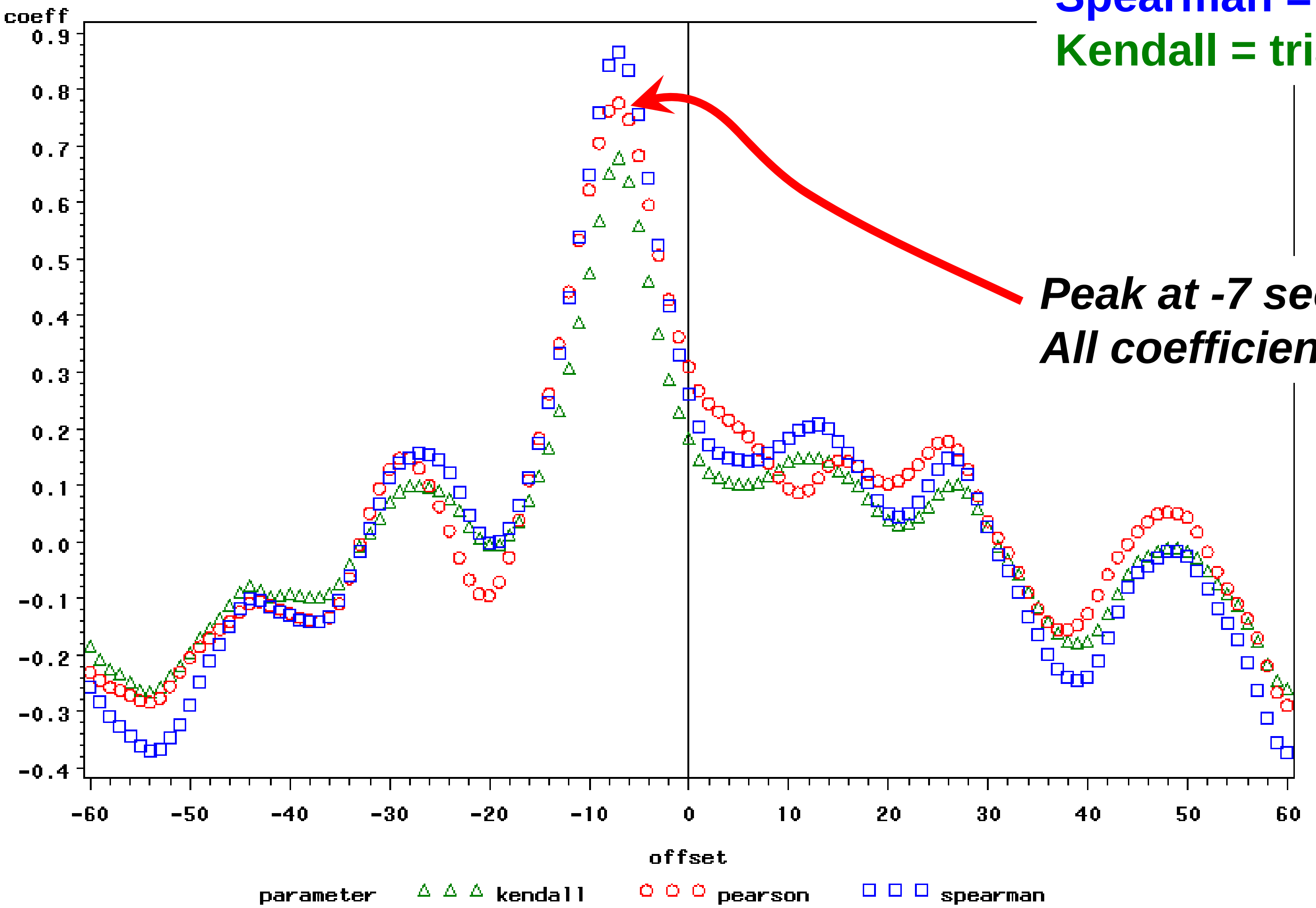
VSP time-series "stationary"

Get correlation coefficient at each offset (R_{max}), Select offset giving peak in correlation

Correlation by Offset for NO_x

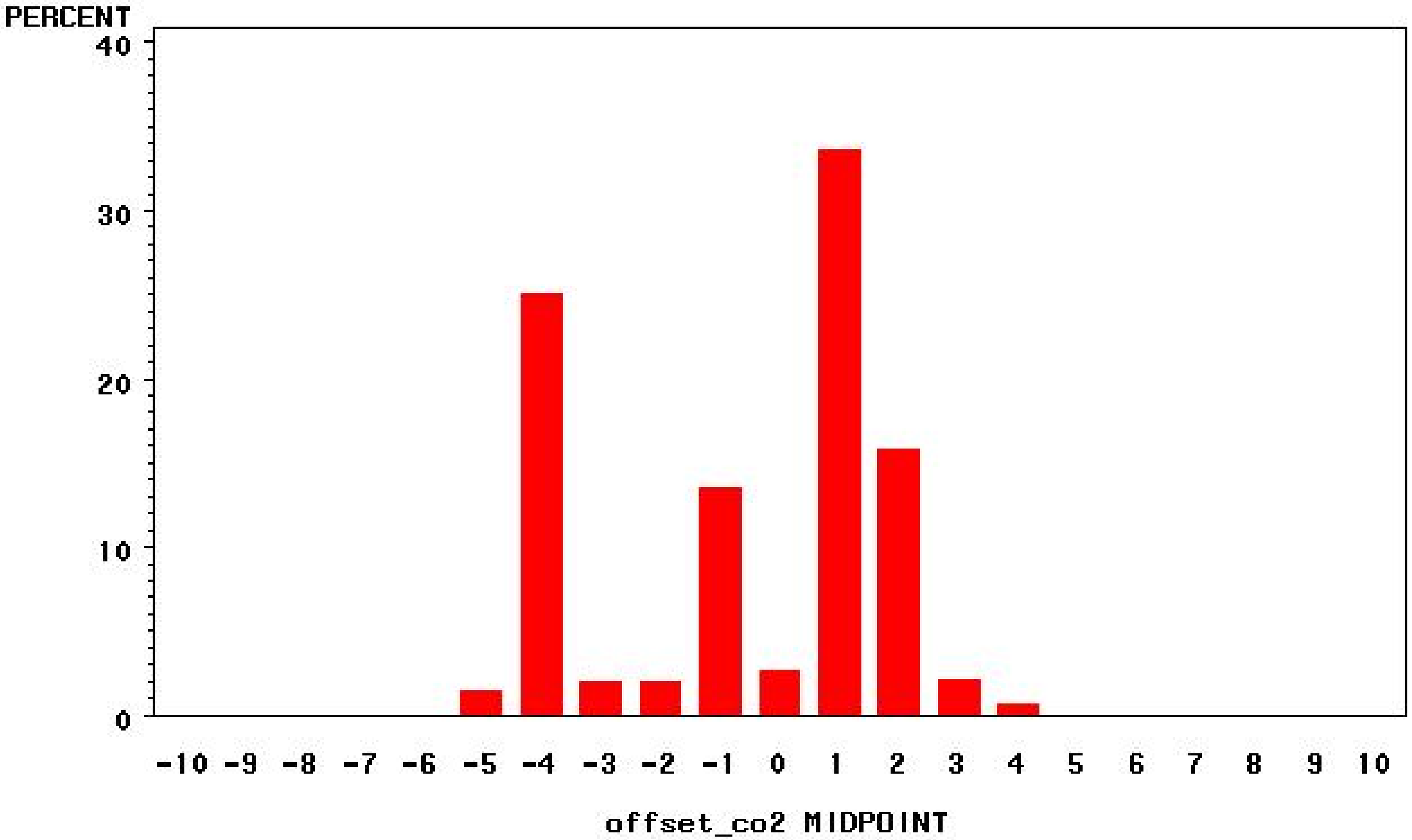
Example: test from NYIPA

Coefficients:
Pearson = circles
Spearman = squares
Kendall = triangles

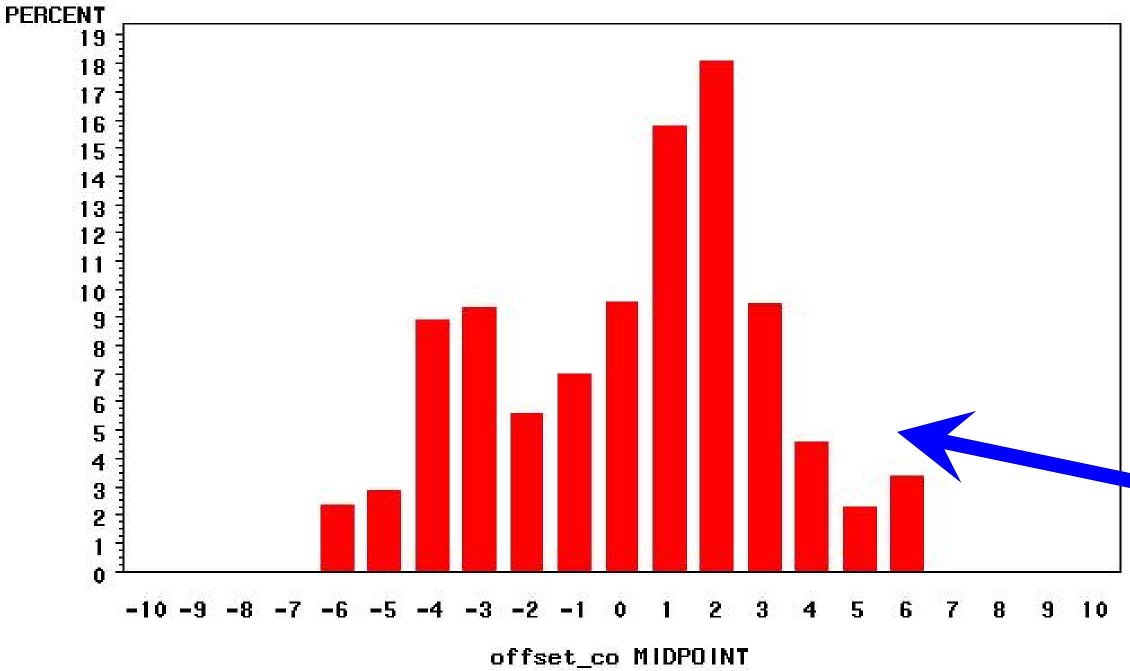


Peak at -7 seconds
All coefficients agree

Distribution of Offsets: CO_2

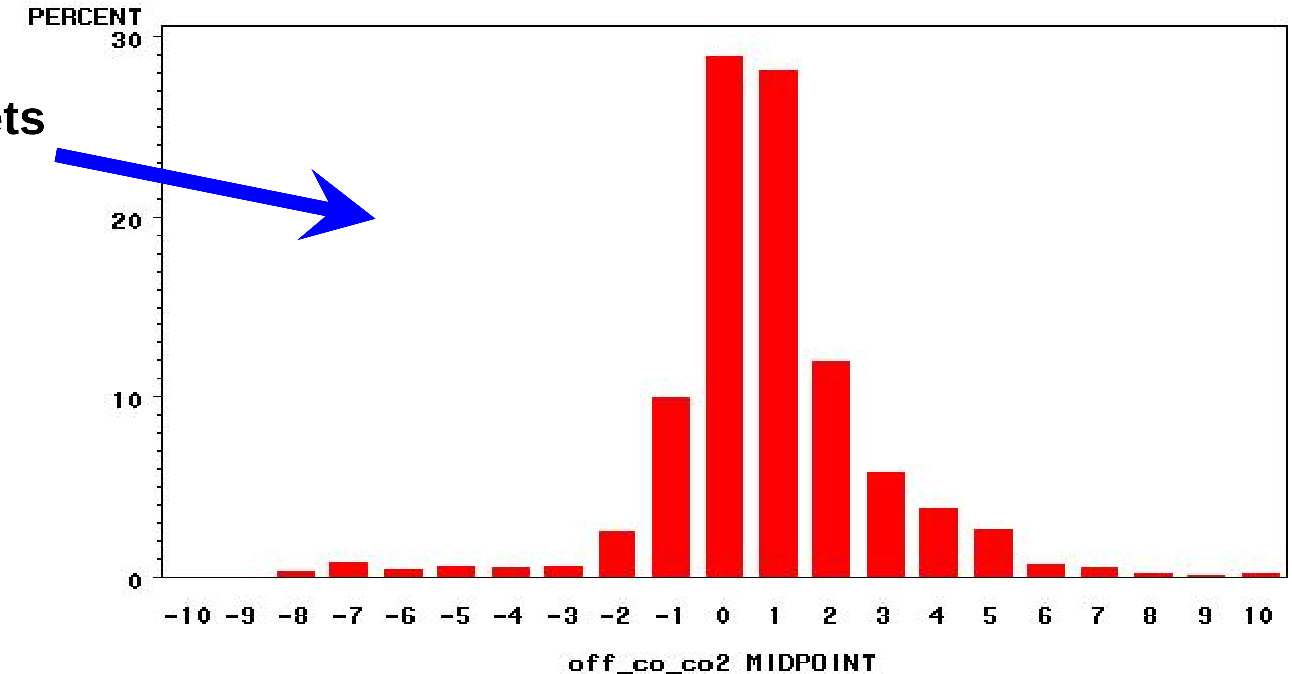


Distribution of Offsets: CO

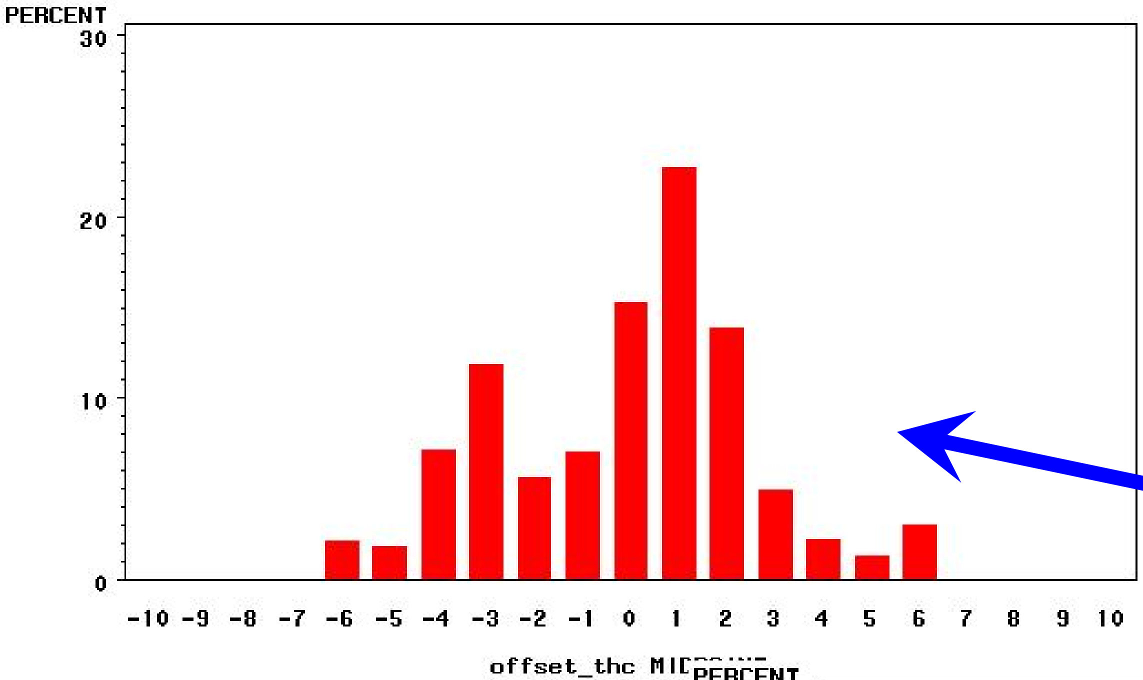


Distribution of offsets

Distribution of offsets With respect to CO_2

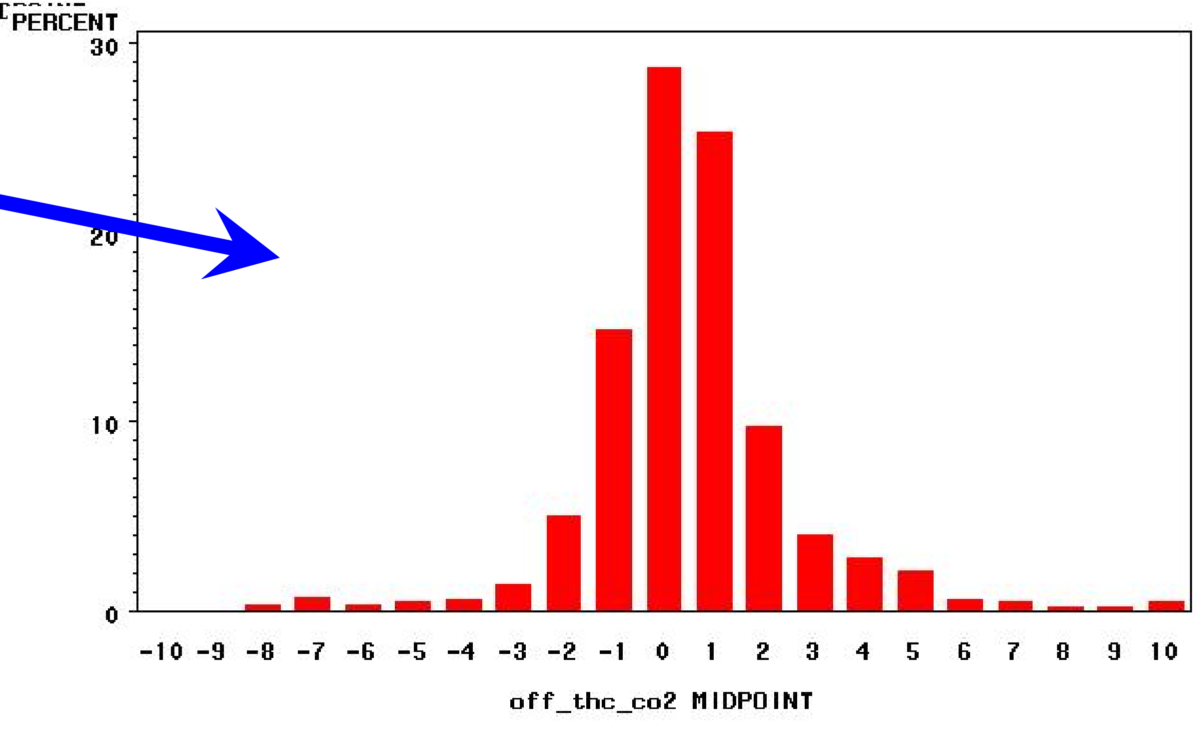


Distribution of Offsets: THC

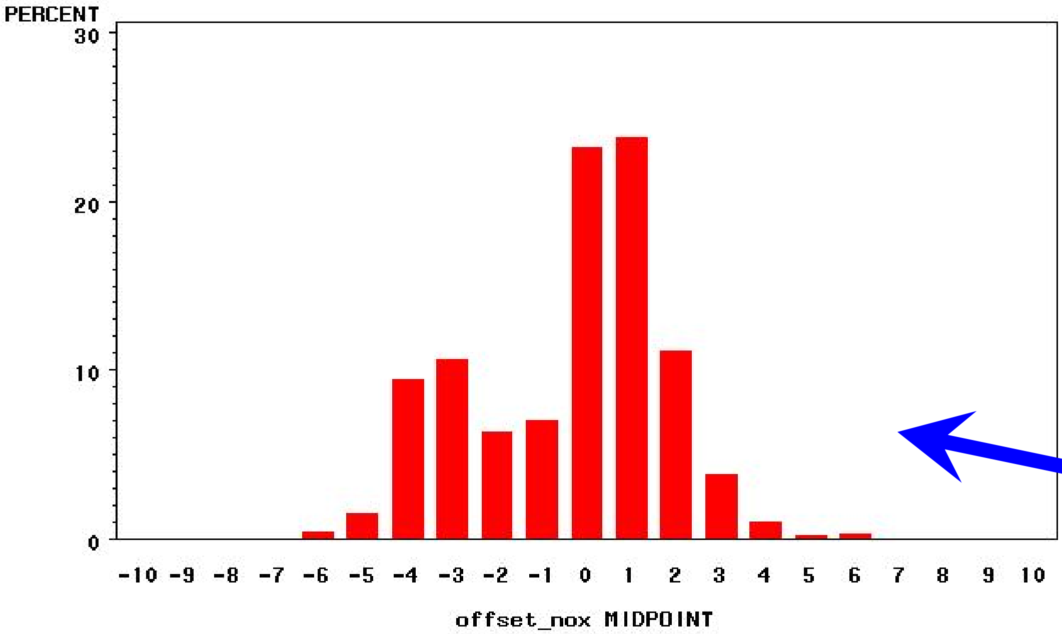


Distribution of offsets

Distribution of offsets With respect to CO_2

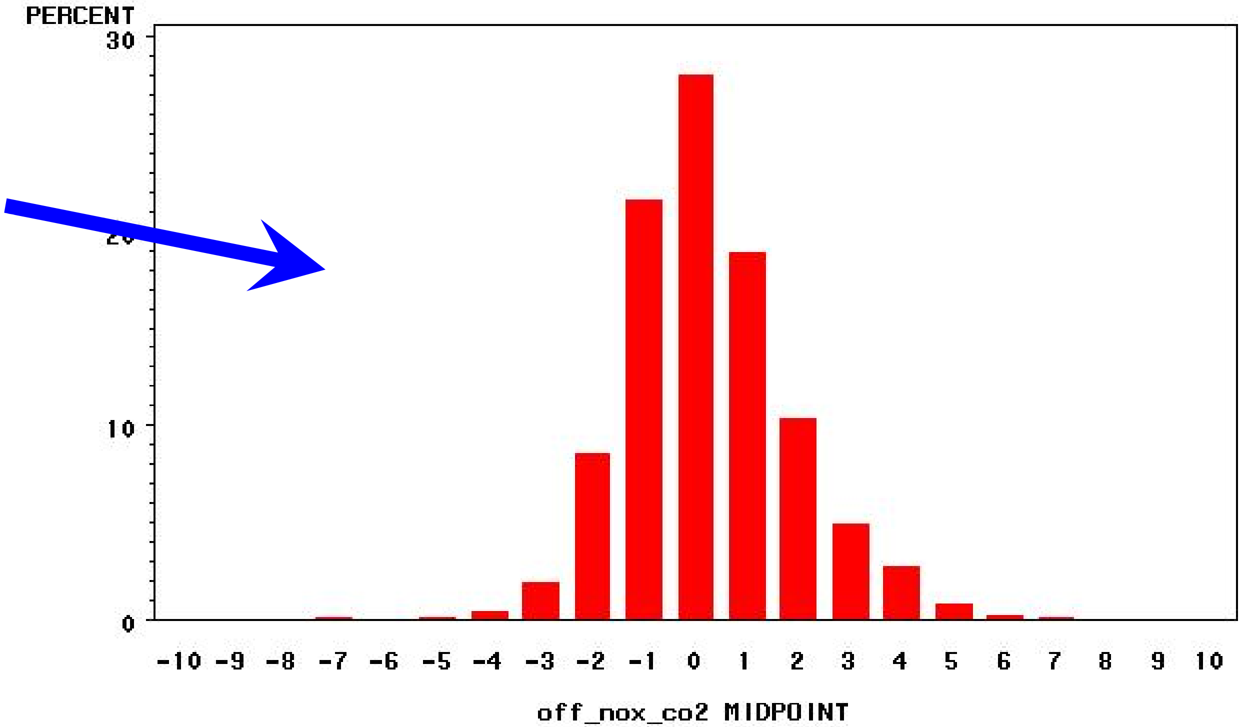


Distribution of Offsets: NO_x



Distribution of offsets

Distribution of offsets With respect to CO_2



Alignment: Offsets and Options

- **Question: should pollutant emissions be aligned independently? Or aligned to CO_2?**
- **Rule of thumb: align independently,**
  - If correlation at peak "*high enough*," AND
  - Offset not "*too large*"
- **Options**
  - CO and THC: R_{max} >= 0.20
  - CO2 and NOx: R_{max} >=0.30

Alignment: An Example

Make: Oldsmobile
Model: Delta 88
Model year: 1983

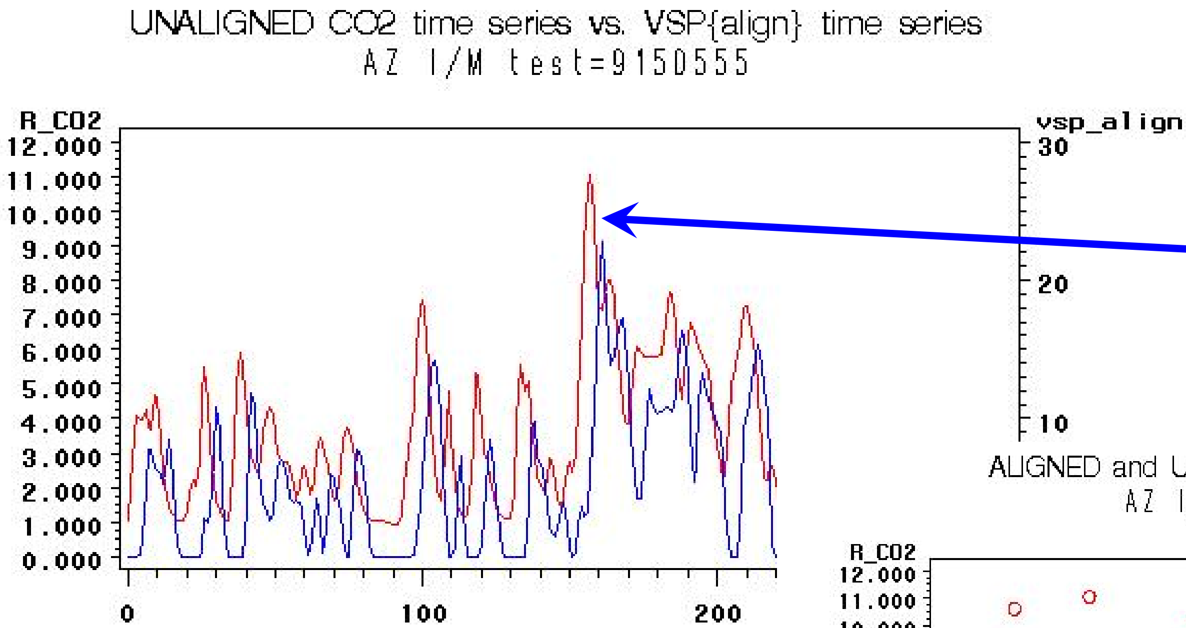
Test date: October, 1998
Cycle: IM240

Options:

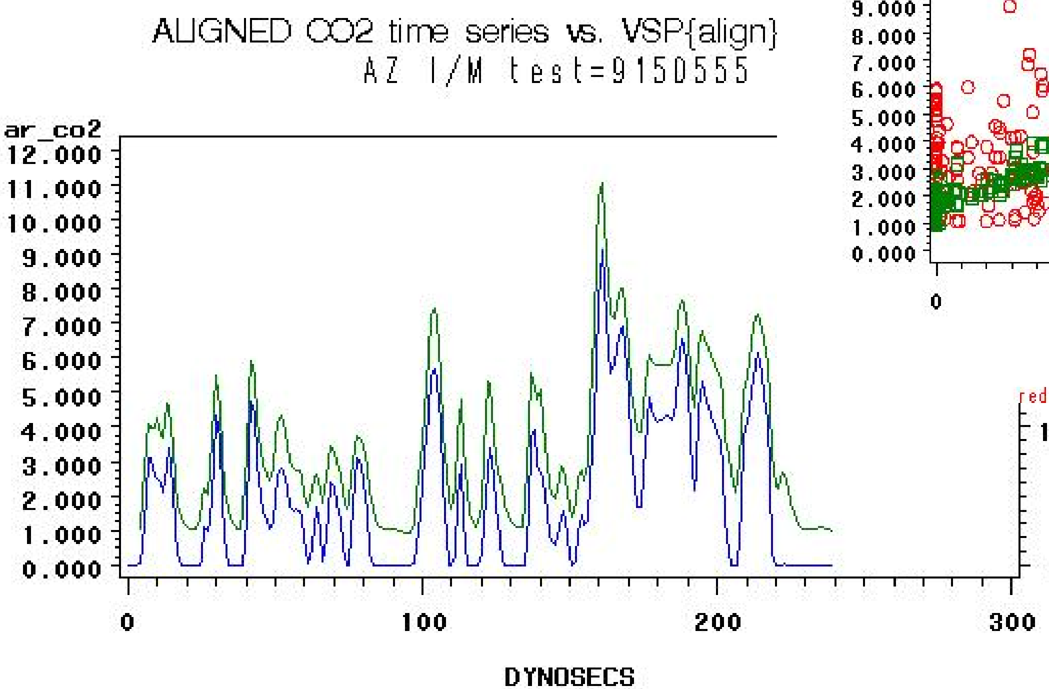
CO	independent
THC	dependent (align to CO_2)
NOx	independent

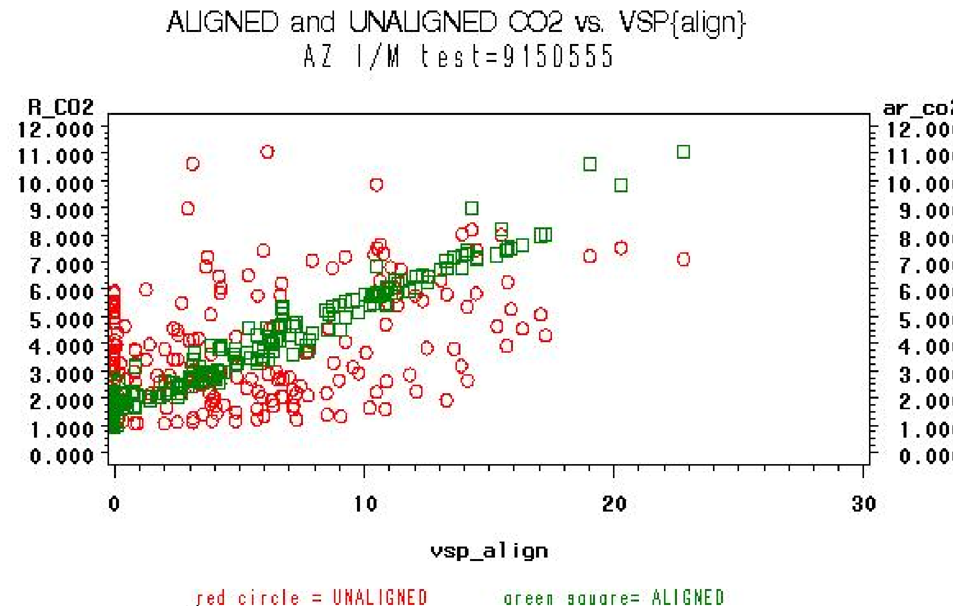
Ctr_tst_id =9150555

CO_2



Leads VSP by 4 sec



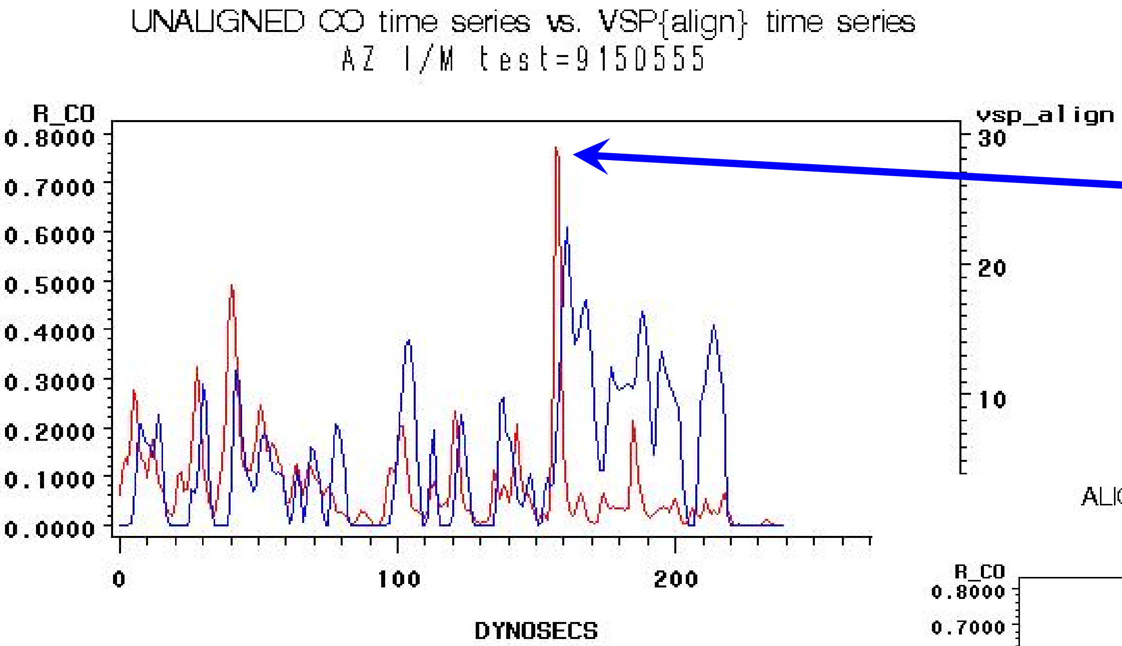


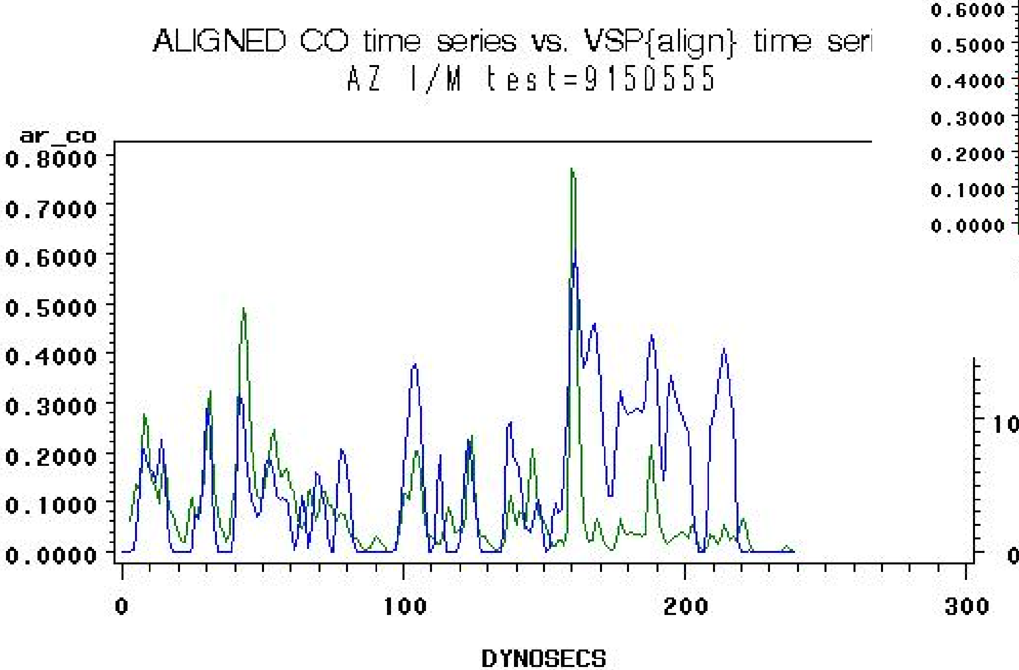
<u>Pearson CC</u>

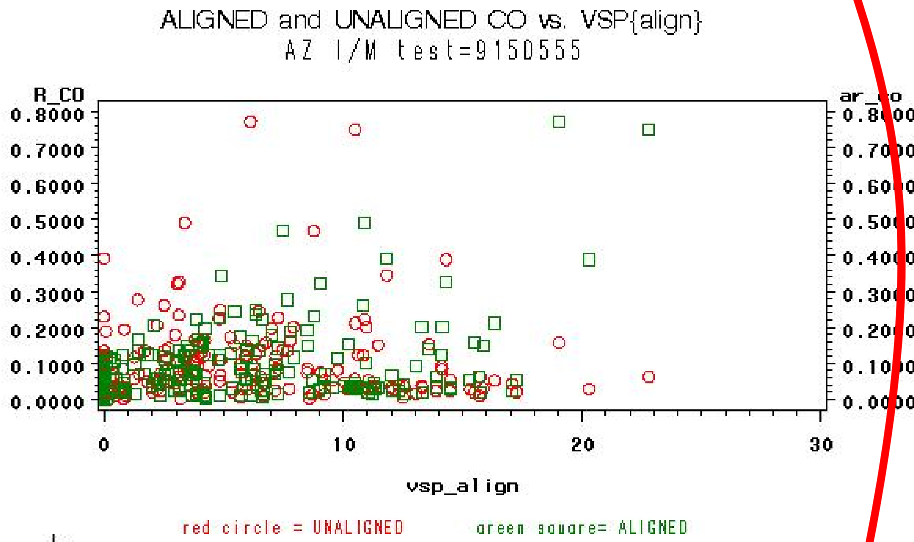
Unaligned	***R*=0.49**
Aligned:	***R*=0.97**

CO

Leads VSP by 2 sec (aligned independently)

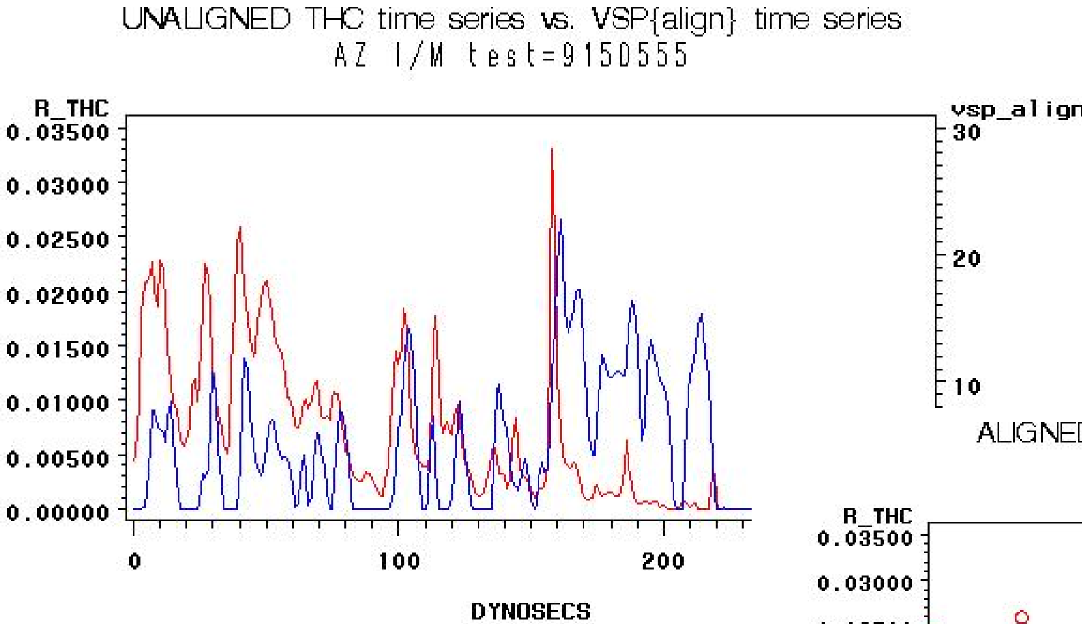






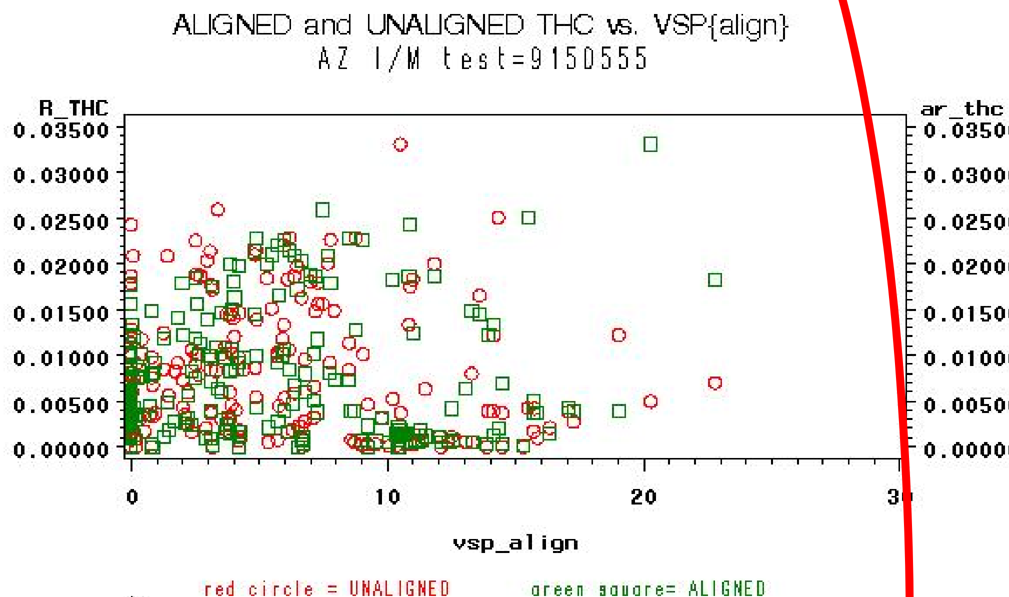
<u>Spearman CC</u>

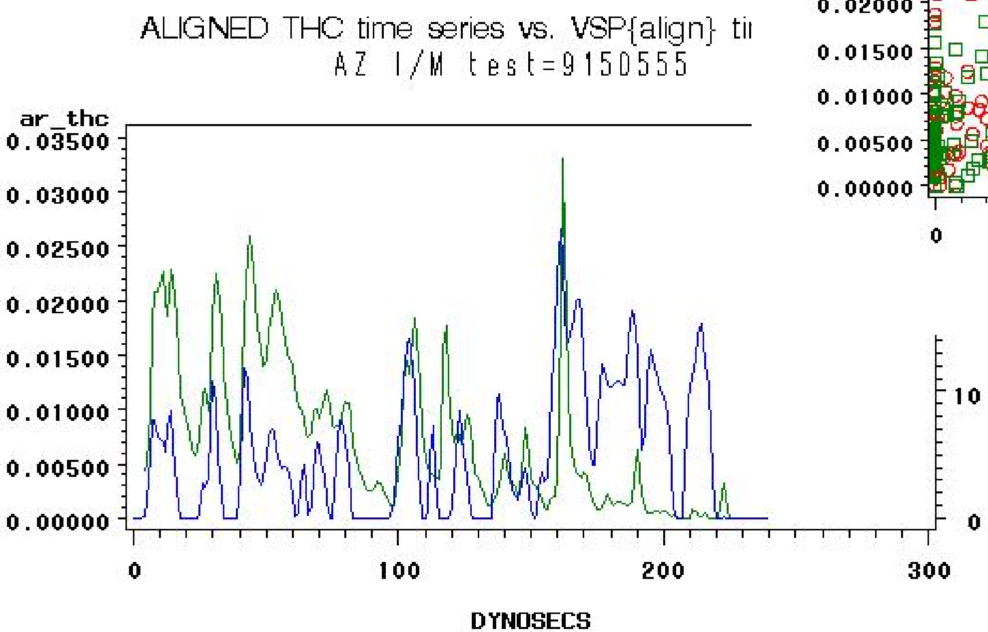
Unaligned	***R*=0.30**
Aligned:	***R*=0.44**



THC

Leads VSP by 3 sec (aligned to CO2)



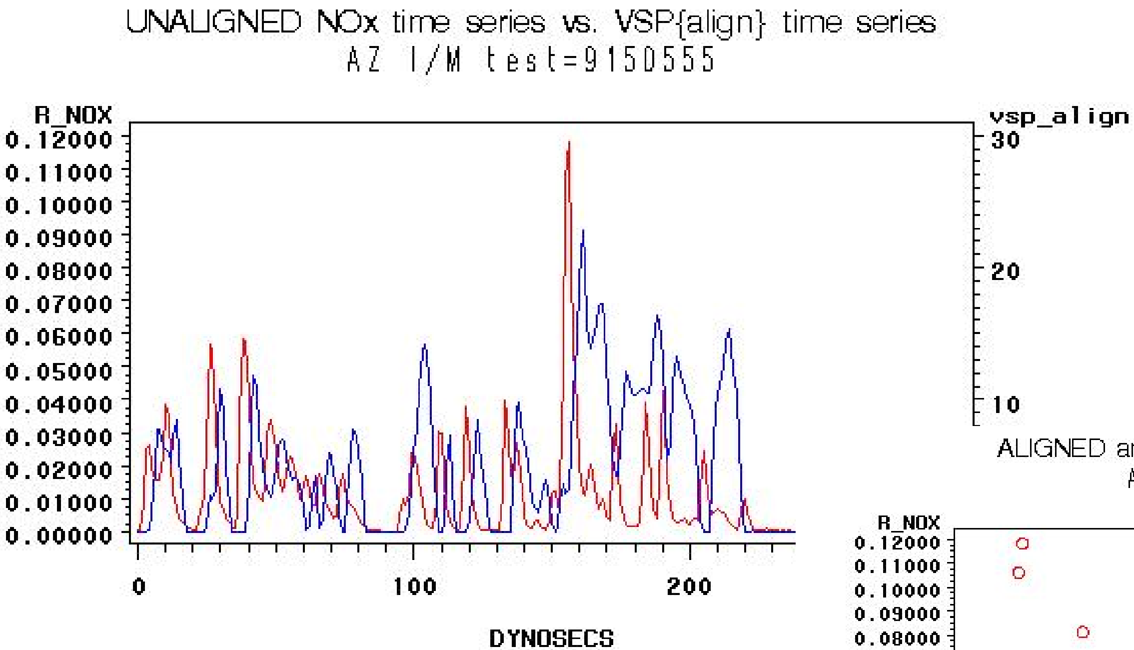


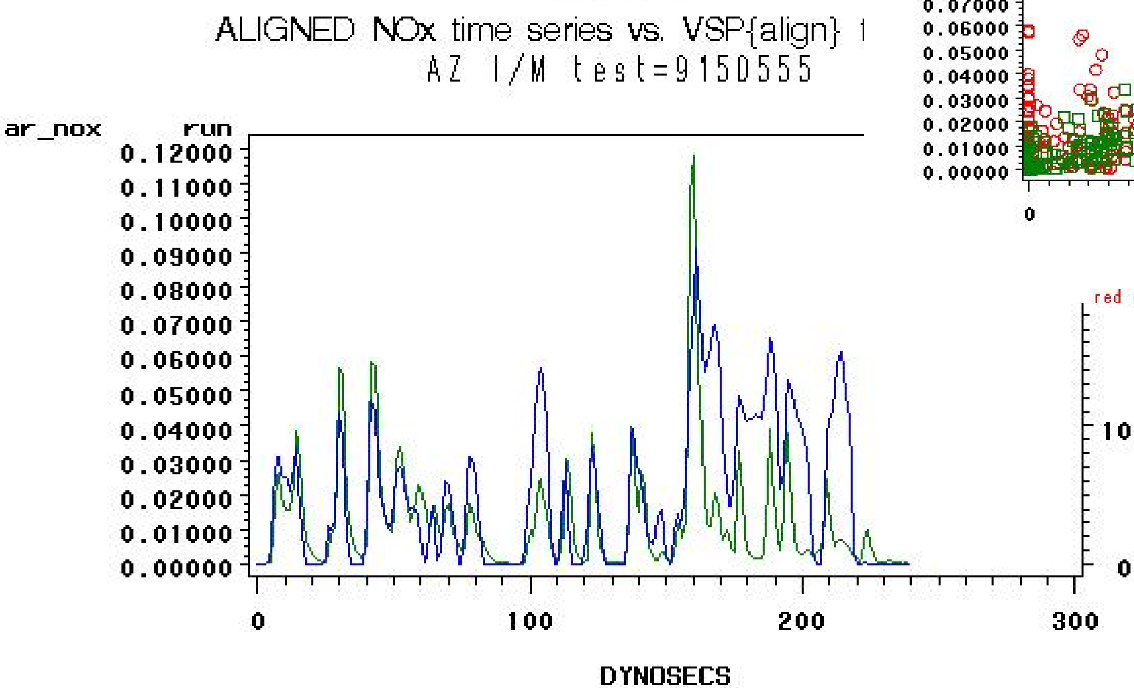
<u>Spearman CC</u>

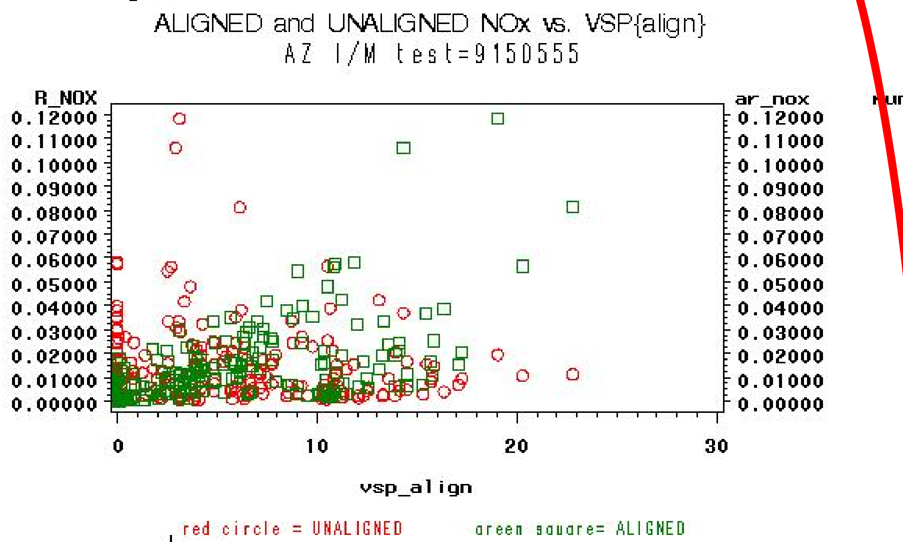
Unaligned	***R*=0.09**
Aligned:	***R*=0.17**

NO_x

Leads VSP by 4 sec (aligned independently)







Pearson CC	
Unaligned	R=0.02
Aligned:	R=0.60

Emissions by Source, Age, Mode

- MOVES uses a different rate for each combination of:
  - Source,
  - Age group, and
  - Operating mode

Gas-LDV-MY1998 / 8-9 years / "low-speed" coast

Gas-LDT-MY2002 / 4-5 years / "cruise/accel" (speed 25-50 mph, VSP 12-15 kW/tonne)

MY by Age Coverage: LDT

(*dynamometer, NYIPA + MSOD "I/M"*)

Model-Year Group by Age Group, LDT
(numbers of vehicles)

For area in green, we have data and Estimate rates by "binning"

MYG	**Age Group**						
	0 to 3	***4 to 5***	***6 to 7***	***8 to 9***	***10 to 14***	***15 to 19***	***20 plus***
pre 1981							
1981-82							
1983-85					**9**	**26**	
1986-89				**5**	**200**	**39**	
1990-93			**103**	**168**	**125**		
1994	**6**	**33**	**61**	**24**			
1995		**75**	**71**				
1996	**44**	**91**	**34**				
1997	**22**	**37**					
1998	**53**	**8**					
1999	**60**						
2000	**47**						
2001	**21**						

For area in yellow, we have no data; estimate rates using a regression model

MY by Age Coverage: LDV

(*dynamometer, NYIPA + MSOD "I/M"*)

Model-Year Group by Age Group,
(numbers of vehicles)

For area in blue, we have data and Estimate rates by "binning"

MYG	Age Group *0 to 3*	*4 to 5*	*6 to 7*	*8 to 9*	*10 to 14*	*15 to 19*	*20 plus*
pre 1981							**5**
1981-82						**43**	
1983-85					**53**	**207**	
1986-89				**15**	**1093**	**125**	
1990-93		**22**	**390**	**764**	**436**		
1994	**4**	**119**	**201**	**48**			
1995		**266**	**177**				
1996	**167**	**238**	**47**				
1997	**66**	**128**					
1998	**128**	**29**					
1999	**146**						
2000	**126**						
2001	**65**						

For area in yellow, we have no data; estimate rates using a regression model

MOVES

Next Steps

- **Current status: have generated draft set of rates**
  - Based on NYIPA and MSOD
- **Anticipated direction:**
  - Incorporate additional I/M datasets
    - AZ (Phoenix) and IL (Chicago)
  - Incorporate RSD datasets
  - Estimate set of I/M rates (as default)
    - Based on dynamometer data
    - Average available data within MOVES framework
    - Fill holes
  - Estimate set of non-I/M rates (relative to I/M)
    - Based on some combination of dynamometer and RSD data