Total Hg in Fish Tissue vs. Hydrology Data

August 28, 2006

John W. Green

Overview

- **There are significant relationships between**
  - **days of high discharge rates (probably storm events) and high fish tissue Hg content**
  - **between periods of low discharge rates and low fish tissue Hg content**
- **Should not be over-interpreted**
  - **Some plots suggest the regressions are driven by three large storm events and a few periods of unusually low flow rates**

Years Fish Sampled

	Years Species Sampled			
	LMB	SMB	Sucker	SunFish
Year				
1977	X	X	X	X
1978	X	X	X	X
1979	X	X	X	X
1980	X	X	X	X
1981	X	X	X	X
1983	X	X	X	X
1984	X	X	X	X
1985	X	X	X	X
1986	X	X	X	X
1987	X	X	X	X
1992	X	X	X	X
1994	X	X	X	X
1996	X	X	X	X
1999	X	X	X	X
2001		X		
2002	X	X	X	X
2005	X	X	X	X

Fish were not sampled every year.

There are 1, 2, 3 and 5 years between samples.

Relationship between discharge rates (or storm events) and total Hg in fish tissue might be confounded by delay in sampling.

Analysis should allow for up to 3-year time lag between storm event and effect observed in fish.

Regression of Fish Tissue Hg on Discharge Data

- **Total Hg in fish tissue is adjusted for fish size through ANCOVA of log(THg) on log(Length), with factors Year, Station, and slope adjustments for each factor**
  - **Separately for each species**
- **Log(Adjusted total Hg) then regressed on maximum daily discharge rate at 0, 1, 2, and 3-year time lags**
  - **Separately for each species, and station**

Regression

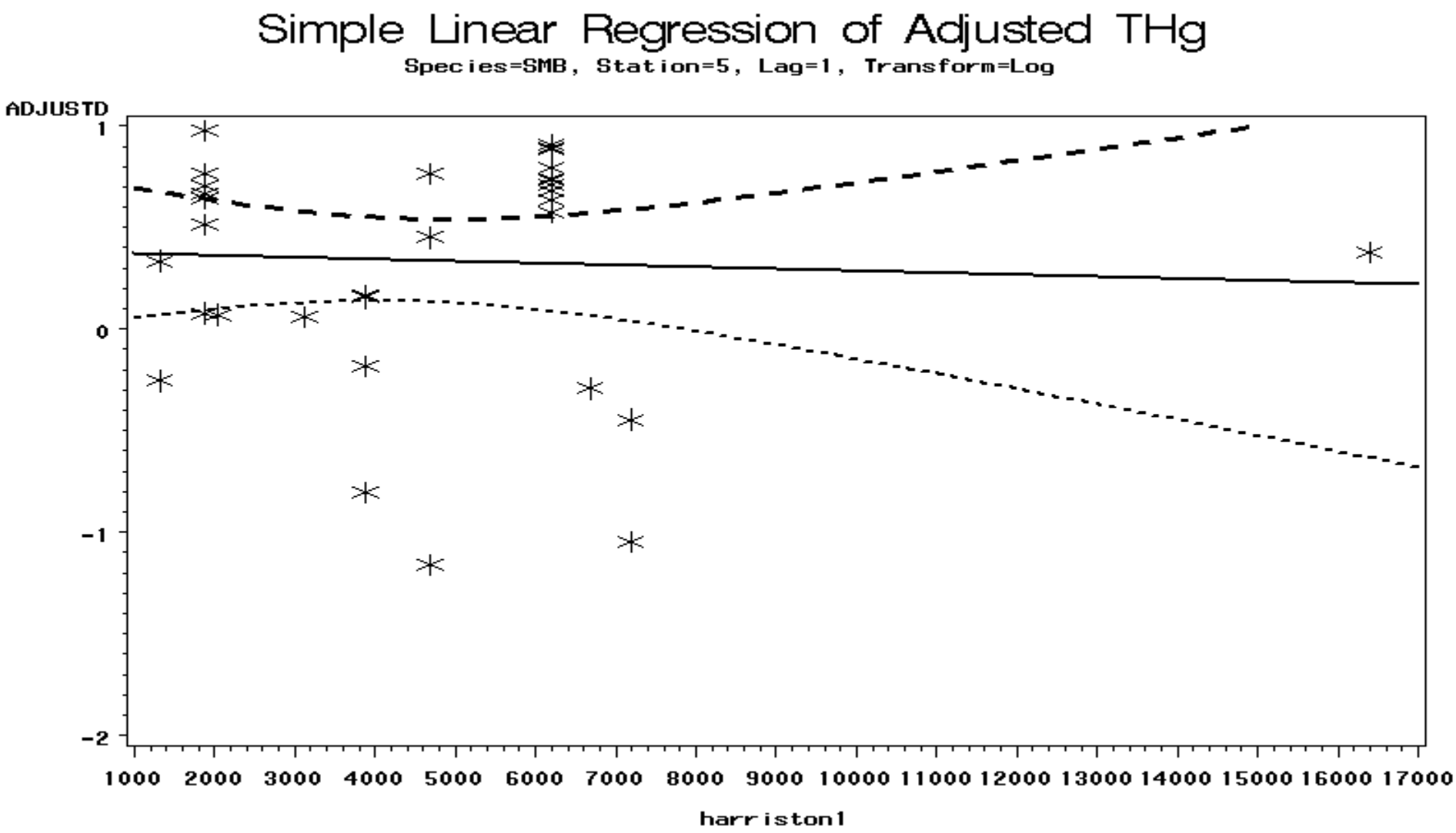
- **Visual and formal analysis showing relationships between total fish tissue Hg and maximum daily discharge 0, 1, 2, and 3 years previous to fish sample**

- **Similar results obtained regressing on maximum monthly or annual discharge**

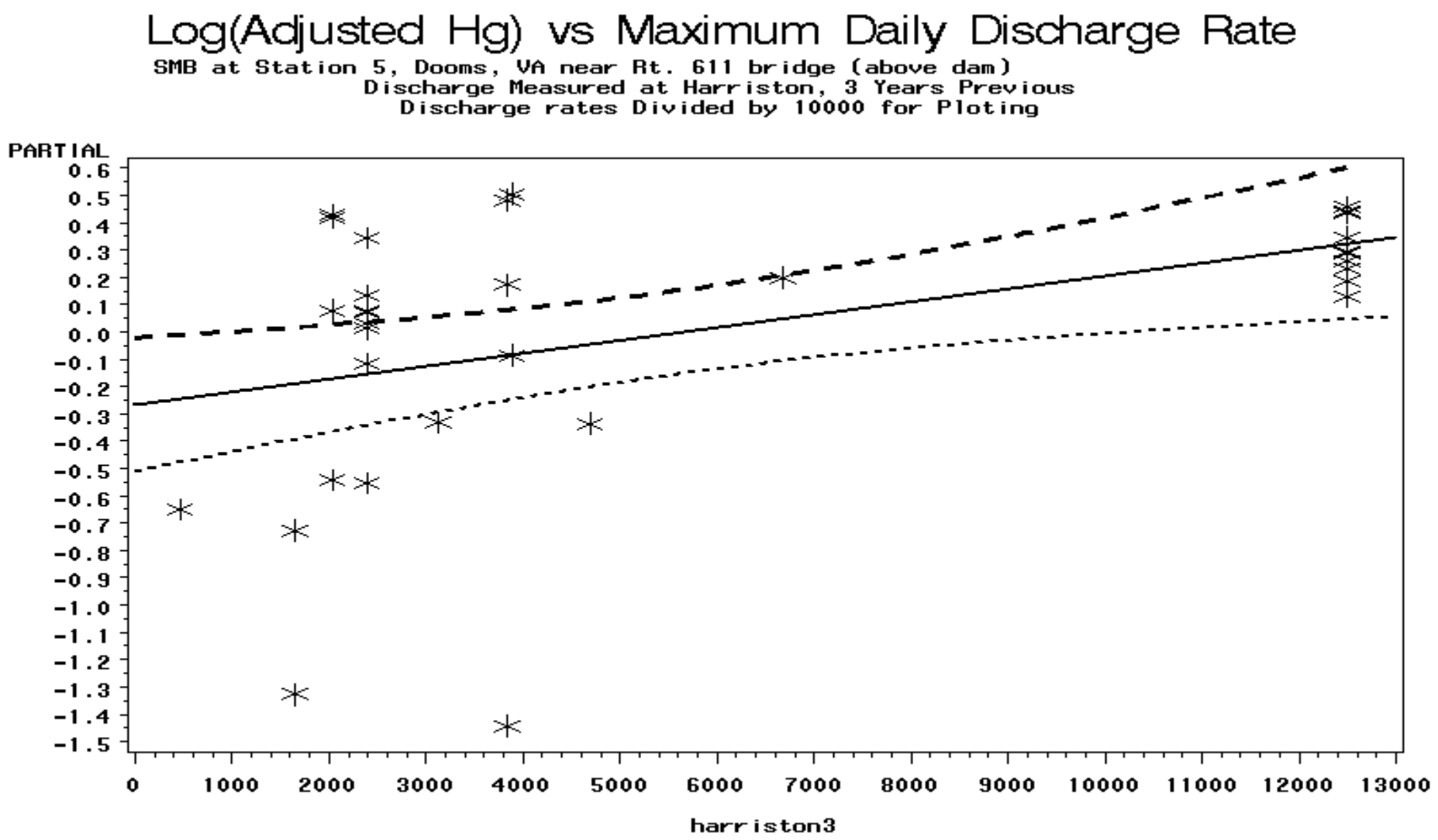
Regression of Log(Adjusted Hg) vs Year
SMB at Station 5, Dooms, VA near Rt. 611 bridge (above dam)

model	rsquare	ratio/rsqr	Source	DF	FValue	ProbF
Year	.	.	Model	9	5.03	0.0007
Year	.	.	Error	24	_	_
Year	.	.	Corrected Total	33	_	_
Year	0.65372	.	R-Square	.	.	.
Hydro	.	.	Model	4	6.51	0.0007
Hydro	.	.	Error	29	_	_
Hydro	.	.	Corrected Total	33	_	_
Hydro	0.473198	72	R-Square	.	.	.
Hydro	.	.	harriston0	1	2.46	0.1278
Hydro	.	.	harriston1	1	4.89	0.0350
Hydro	.	.	harriston2	1	1.72	0.2004
Hydro	.	.	harriston3	1	11.01	0.0024

Parameter	Estimate	StdErr	tValue	Probt
Intercept	0.4918267901	0.18569322	2.65	0.0129
harriston0	-.0000281926	0.00001798	-1.57	0.1278
harriston1	-.0000654337	0.00002958	-2.21	0.0350
harriston2	-.0000296161	0.00002260	-1.31	0.2004
harriston3	0.0000647686	0.00001952	3.32	0.0024



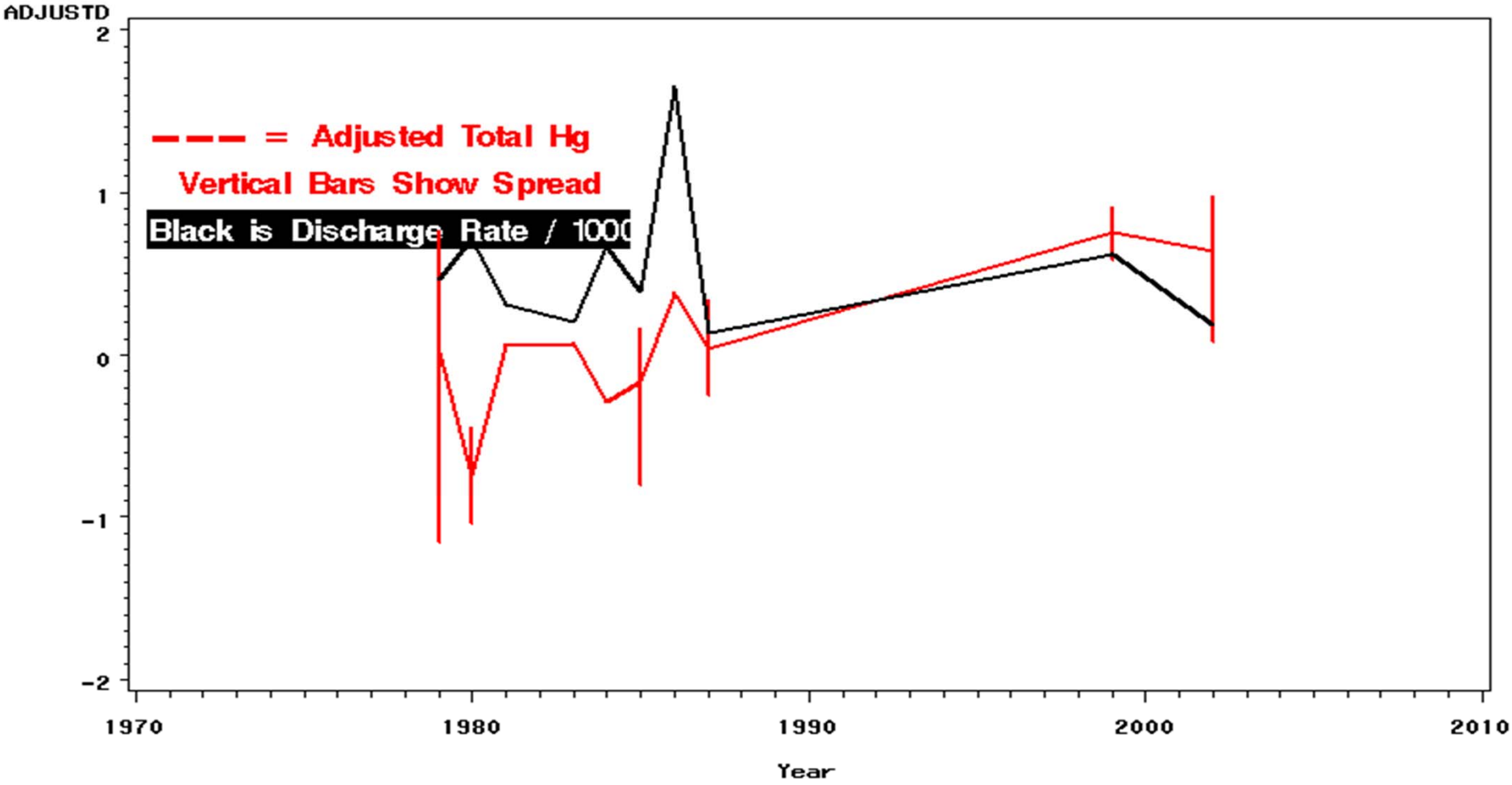
Slight negative slope evident, due to high discharge rate in 1985



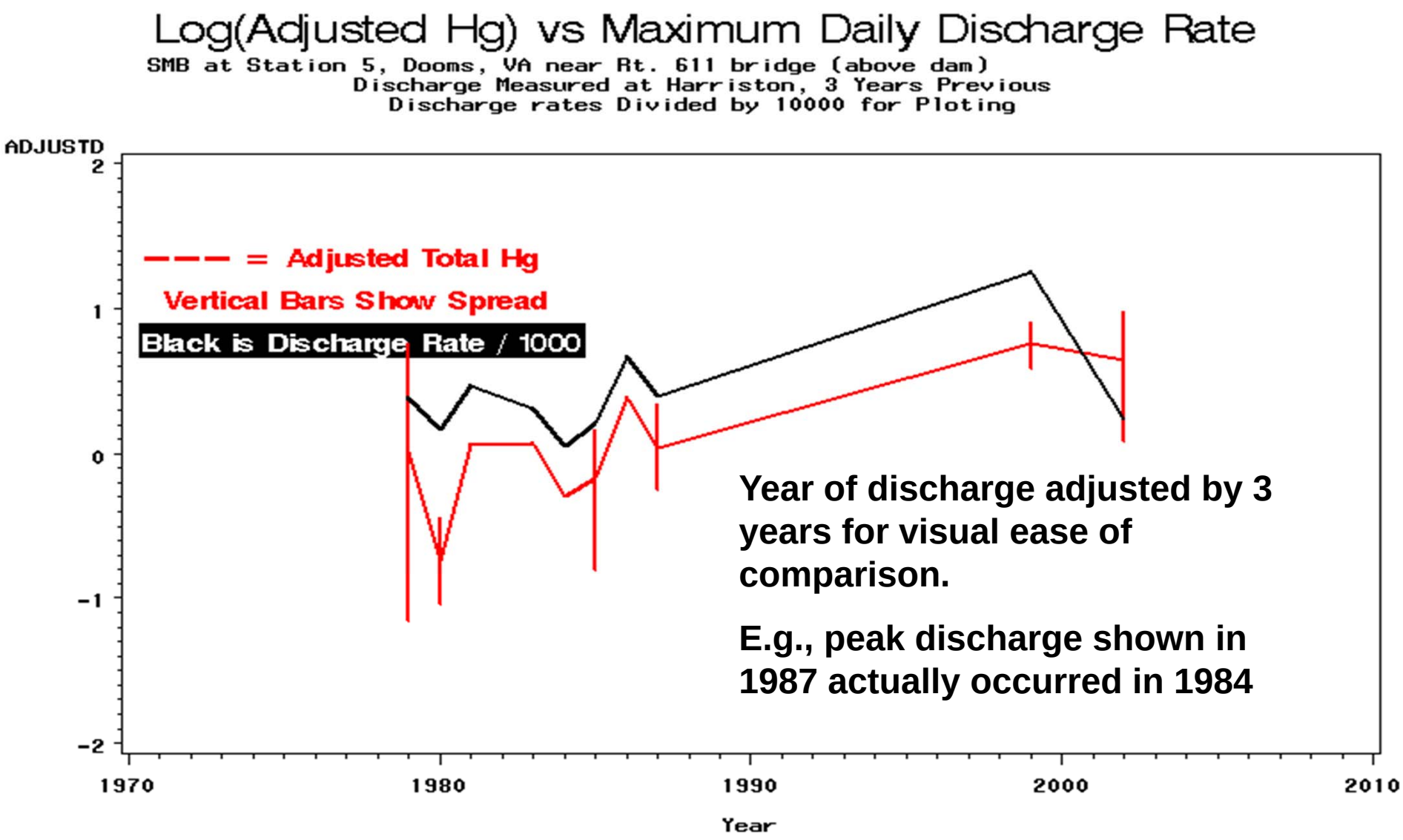
Partial regression plot showing relationship of THg vs Lag3 discharge rate after correcting for lags 0, 1, and 2. This corresponds to ANOVA table on previous slide. Regression driven by high rate (occurring in 1996)

Log(Adjusted Hg) vs Maximum Daily Discharge Rate

SMB at Station 5, Dooms, VA near Rt. 611 bridge (above dam)
Discharge Measured at Harriston, 1 Years Previous
Discharge rates Divided by 10000 for Ploting



The inverse relationship is evident prior to 1985. Effect of major 1985 storm is associated with increase in 1986 Hg levels.

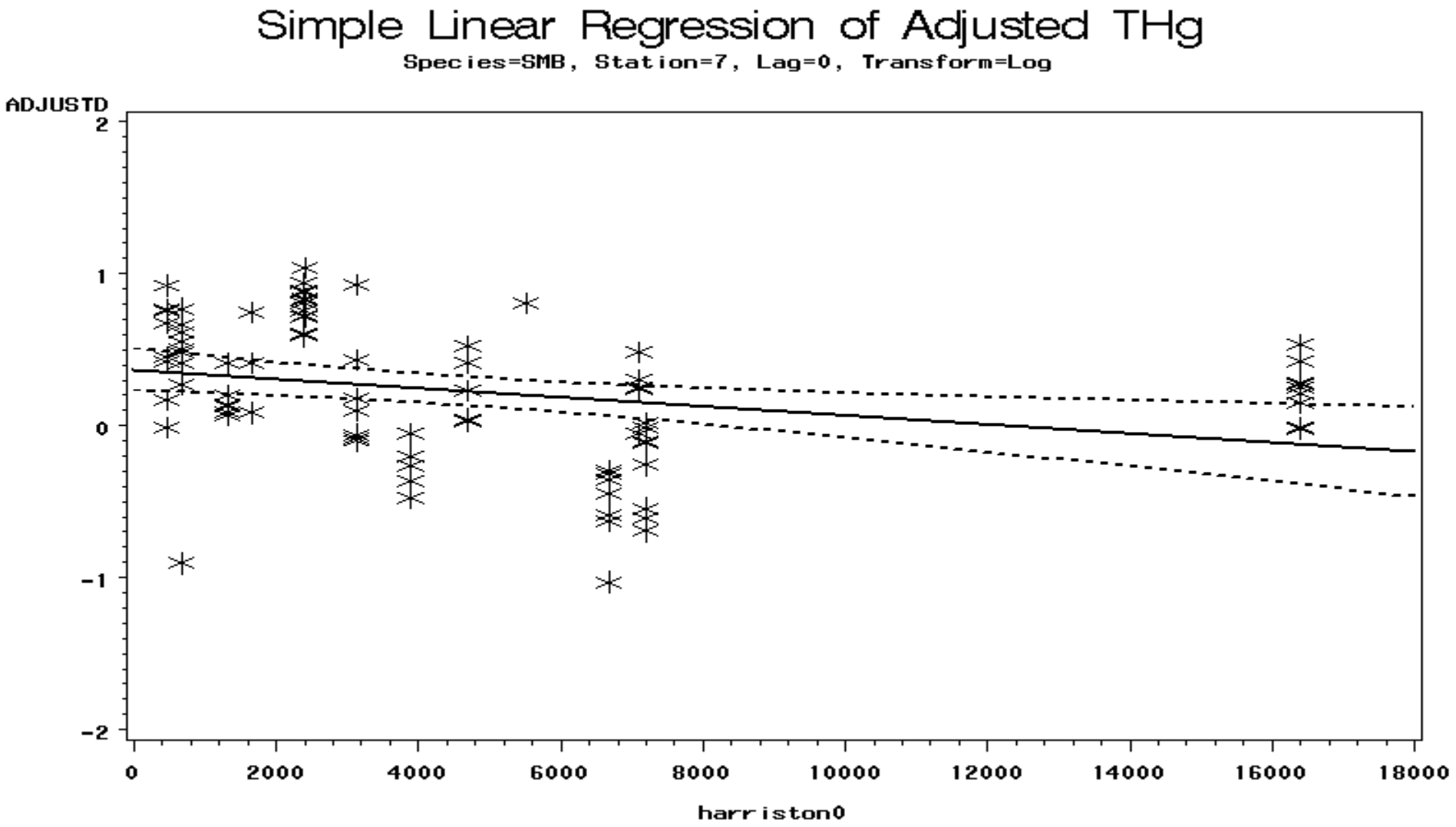


Good correspondence between 3-year lag discharge rate and Hg levels. Effect of 1985 major storm not seen because no fish were sampled in 1988.

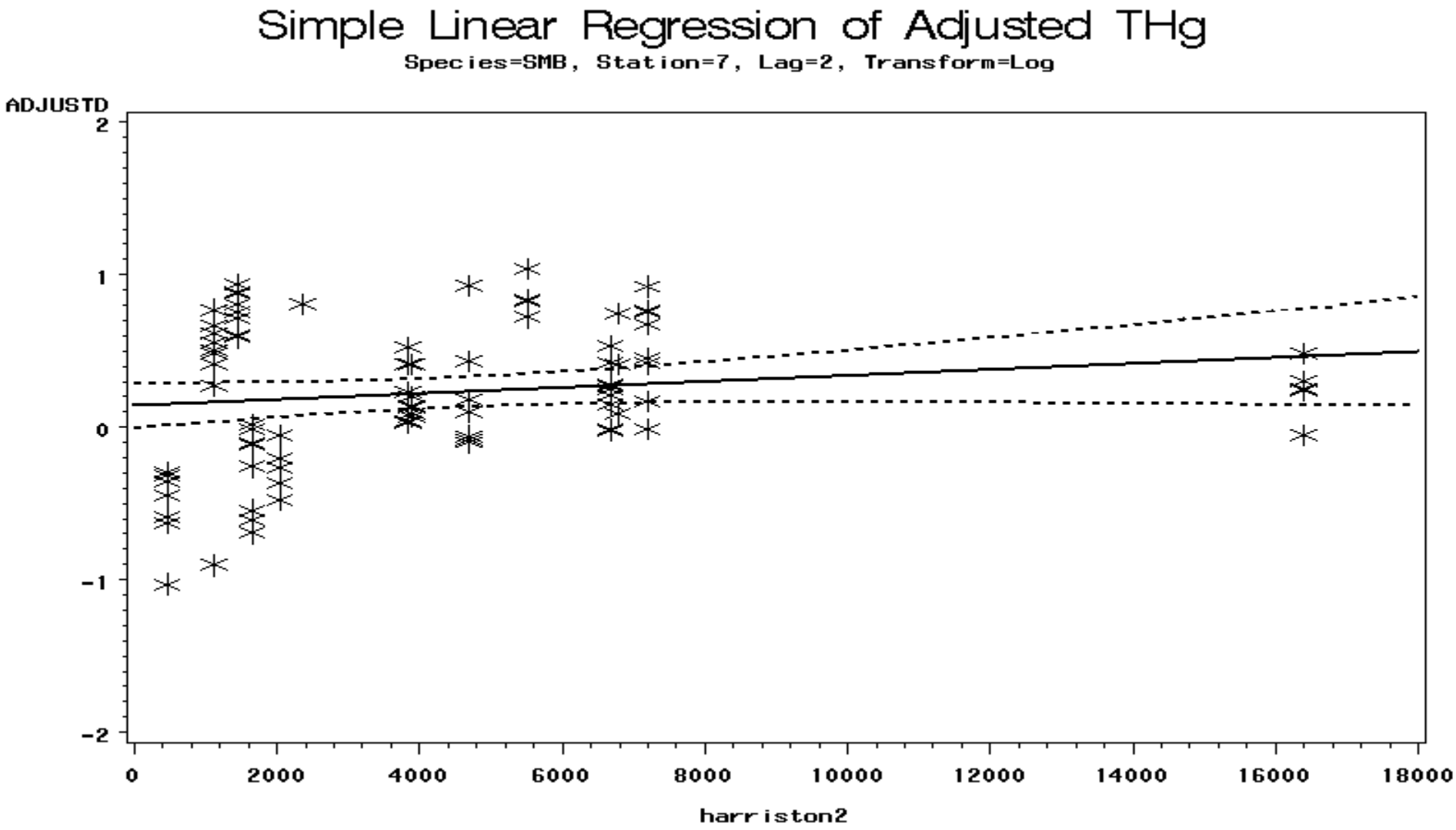
Regression of Log(Adjusted Hg) vs Year
SMB at Station 7, Grottoes, VA near Grand Caverns bridge

model	ratio/rsqr	rsqr	Source	DF	FValue	ProbF
Year	.	.	Model	13	13.93	<.0001
Year	.	.	Error	76	_	_
Year	.	.	Corrected Total	89	_	_
Year	0.704344	.	R-Square	.	.	.
Hydro	.	.	Model	4	7.61	<.0001
Hydro	.	.	Error	85	_	_
Hydro	.	.	Corrected Total	89	_	_
Hydro	0.263811	37	R-Square	.	.	.
Hydro	.	.	harriston0	1	7.46	0.0077
Hydro	.	.	harriston1	1	0.94	0.3354
Hydro	.	.	harriston2	1	7.99	0.0059
Hydro	.	.	harriston3	1	14.47	0.0003

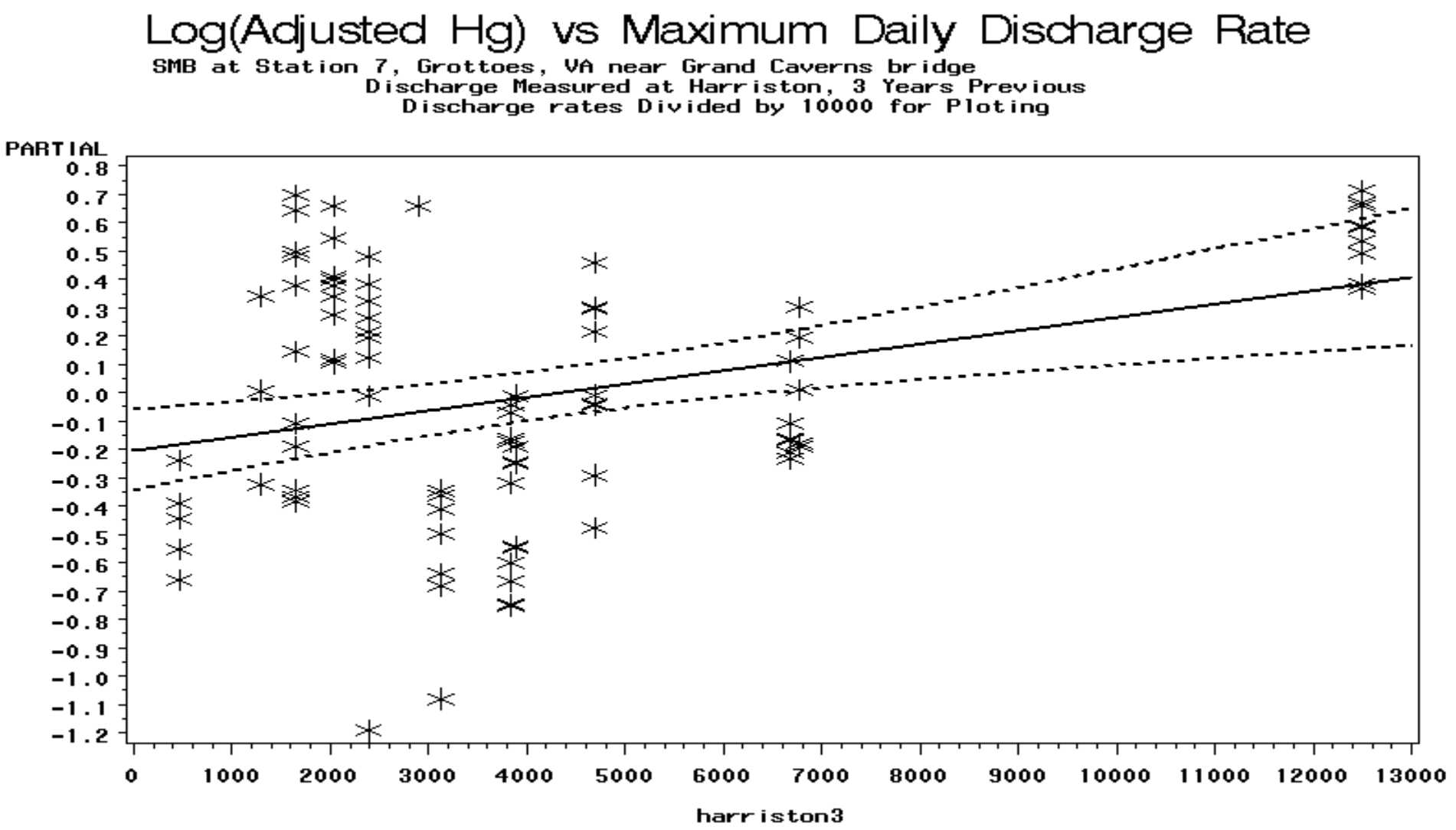
Parameter	Estimate	StdErr	tValue	Probt
Intercept	0.0445480645	0.12318603	0.36	0.7185
harriston0	-.0000281379	0.00001030	-2.73	0.0077
harriston1	-.0000123481	0.00001275	-0.97	0.3354
harriston2	0.0000328303	0.00001162	2.83	0.0059
harriston3	0.0000532205	0.00001399	3.80	0.0003



Downward trend appears real, would be steeper without high discharge rate in 1985



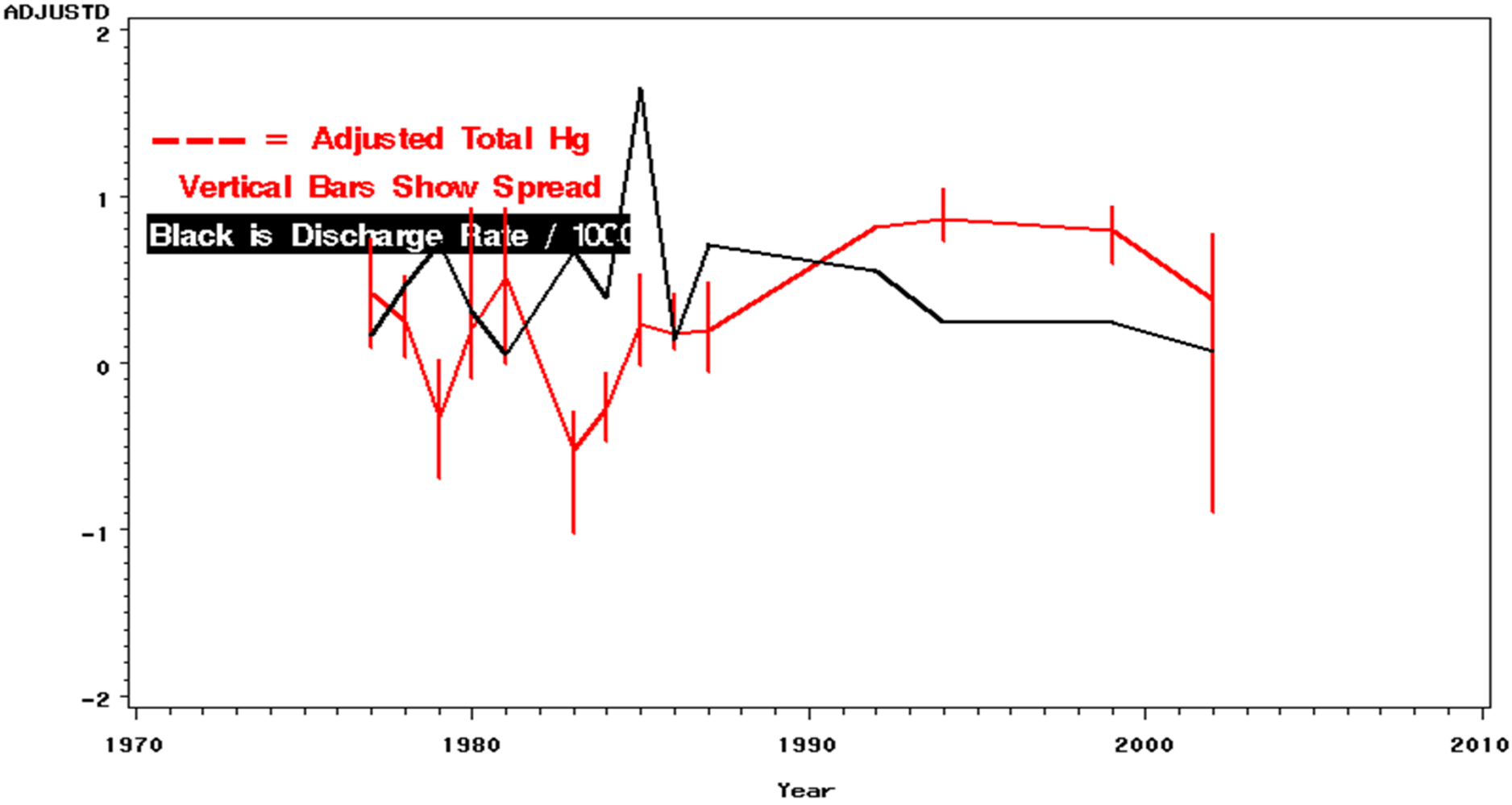
Positive slope due in part to high discharge rate in 1985, low rates in 1981, 2002



Partial regression plot showing relationship of THg vs Lag3 discharge rate after correcting for lags 0, 1, and 2. This corresponds to last line of ANOVA table on previous slide. High discharge rate is from 1996.

Log(Adjusted Hg) vs Maximum Daily Discharge Rate

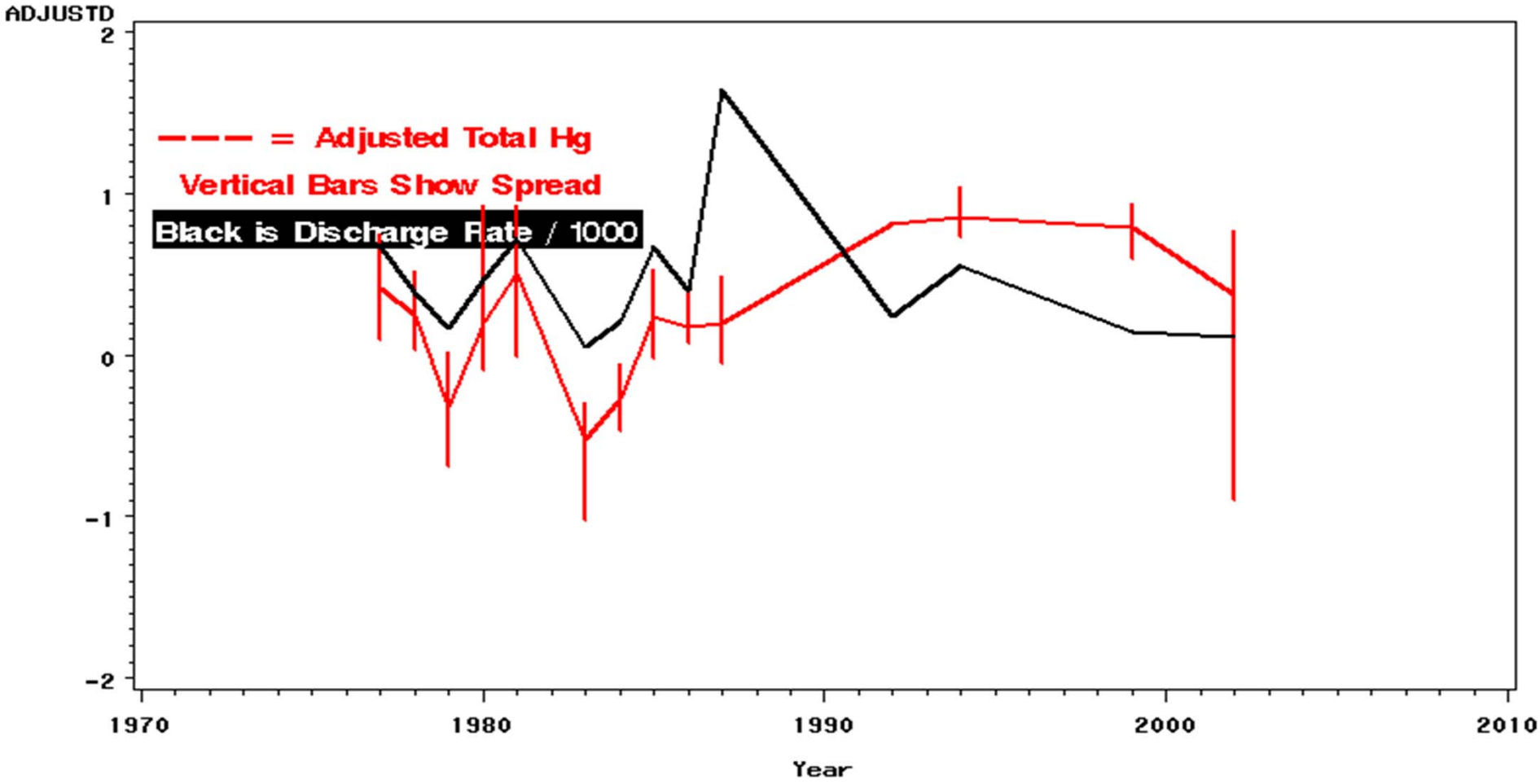
SMB at Station 7, Grottoes, VA near Grand Caverns bridge
Discharge Measured at Harriston, 0 Years Previous
Discharge rates Divided by 10000 for Ploting



Peak discharges match with decreases in Hg. Major storm in 1985 was after fish were sampled.

Log(Adjusted Hg) vs Maximum Daily Discharge Rate

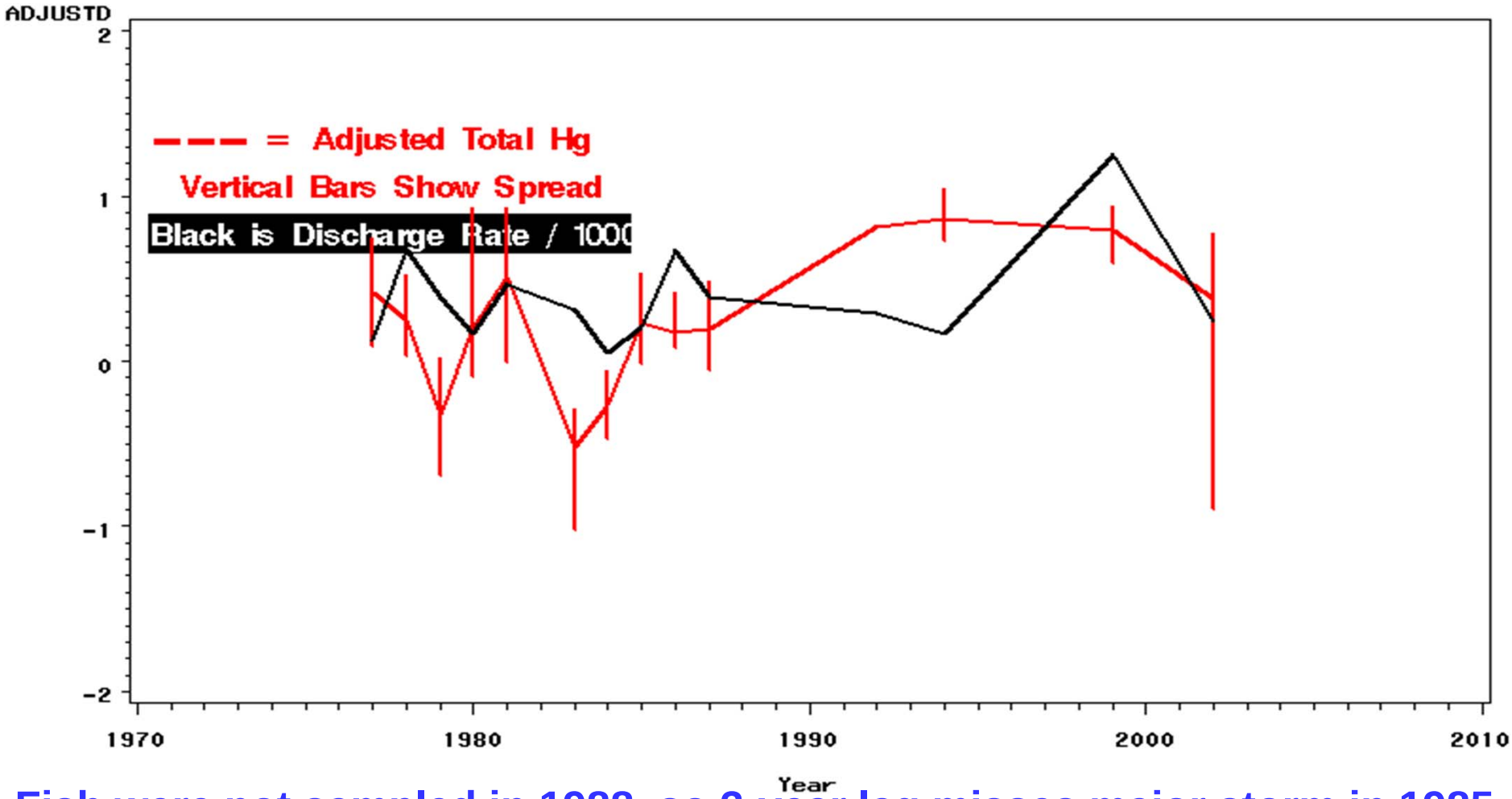
SMB at Station 7, Grottoes, VA near Grand Caverns bridge
Discharge Measured at Harriston, 2 Years Previous
Discharge rates Divided by 10000 for Ploting



Good tracking except for 1985 major storm event.

Log(Adjusted Hg) vs Maximum Daily Discharge Rate

SMB at Station 7, Grottoes, VA near Grand Caverns bridge
Discharge Measured at Harriston, 3 Years Previous
Discharge rates Divided by 10000 for Ploting



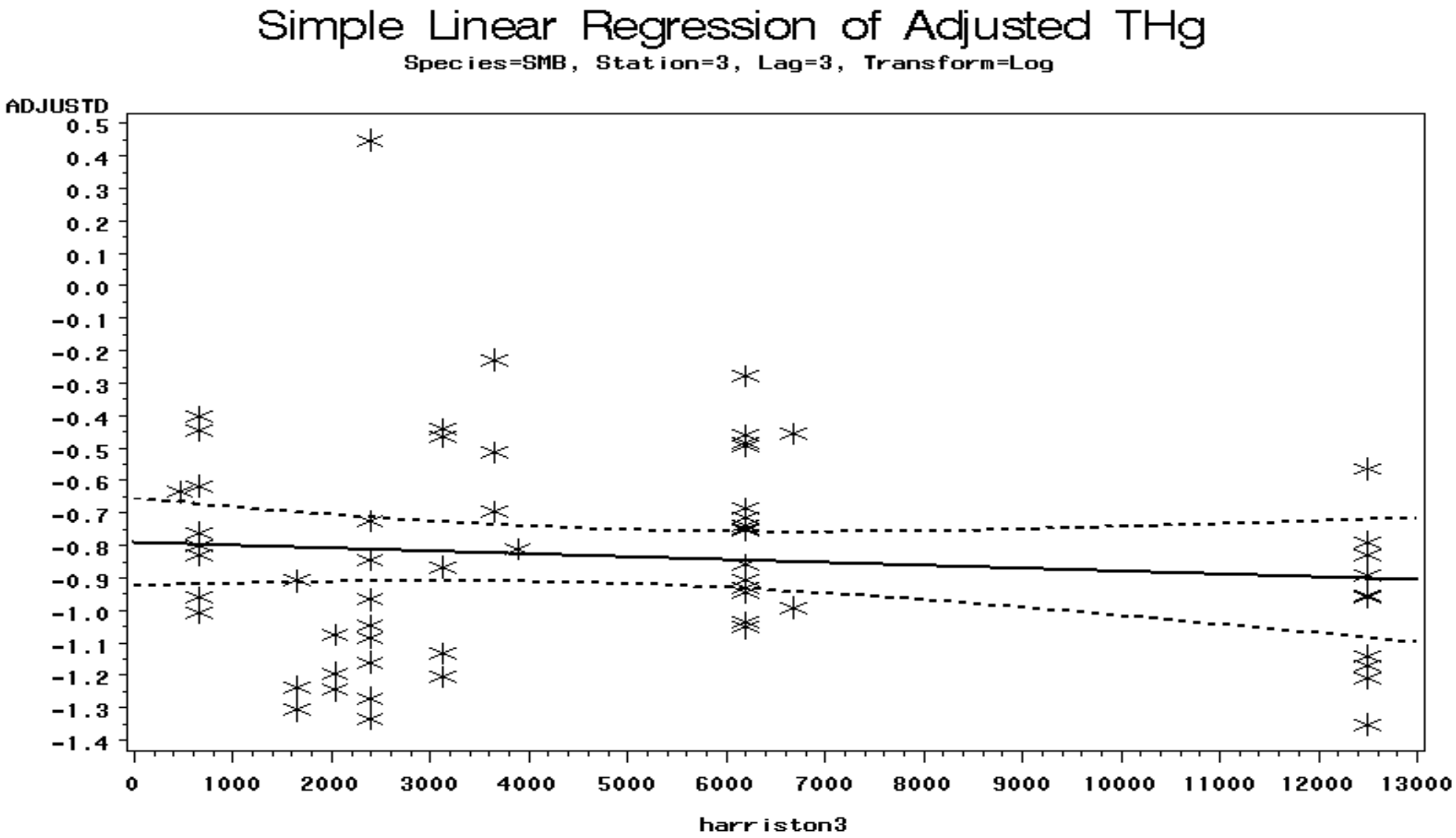
Fish were not sampled in 1988, so 3-year lag misses major storm in 1985. Relationship between 3-year lag discharge rate and Hg less compelling than that between 2-year lag, but this corrects for 1985 major storm effect not seen in previous plot.

Regression of Log(Adjusted Hg) vs Year
SMB at Station 3, Waynesboro City Park north of DuPont footbridge

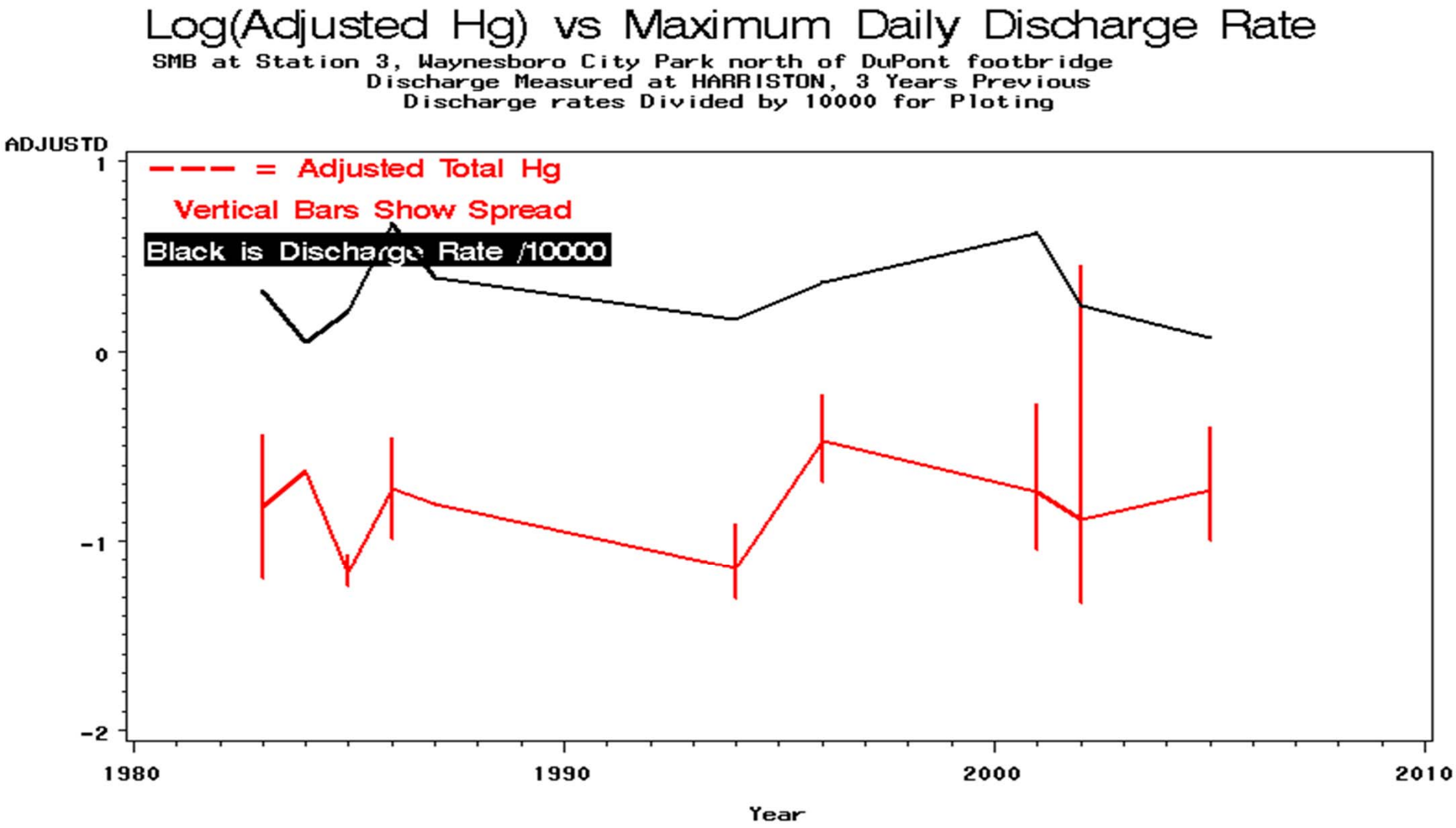
model	ratio/rsqr	rsqr	Source	DF	FValue	ProbF
Year	.	.	Model	10	1.65	0.1188
Year	.	.	Error	50	_	_
Year	.	.	Corrected Total	60	_	_
Year	0.248505	.	R-Square	.	.	.
Hydro	.	.	Model	4	0.22	0.9265
Hydro	.	.	Error	47	_	_
Hydro	.	.	Corrected Total	51	_	_
Hydro	0.018301	7	R-Square	.	.	.
Hydro	.	.	harriston0	1	0.00	0.9713
Hydro	.	.	harriston1	1	0.02	0.8890
Hydro	.	.	harriston2	1	0.60	0.4408
Hydro	.	.	harriston3	1	0.20	0.6603

Parameter	Estimate	StdErr	tValue	Probt
Intercept	-.7662702239	0.12127298	-6.32	<.0001
harriston0	0.0000004623	0.00001276	0.04	0.9713
harriston1	-.0000023116	0.00001647	-0.14	0.8890
harriston2	-.0000164938	0.00002122	-0.78	0.4408
harriston3	-.0000065884	0.00001489	-0.44	0.6603

Total Hg values at station 3 were uniformly low. Regression and plots do not indicate relationship where none exist.



Partial regression plot showing little relationship of THg vs Lag3 discharge rate after correcting for lags 0, 1, and 2. This corresponds to ANOVA table on previous slide.

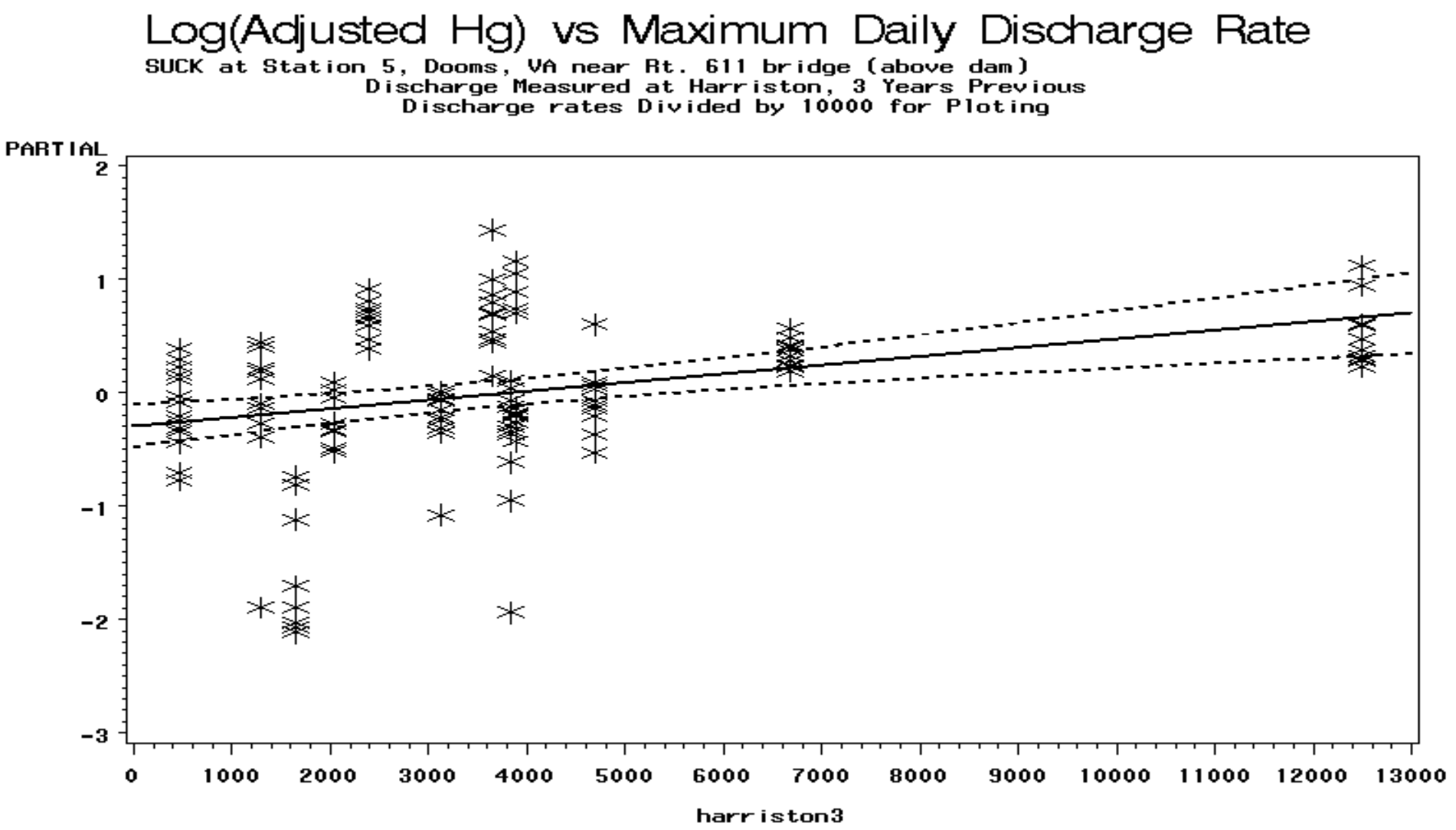


Tracks poorly up to 1985, well 1986-1997, poorly 1997-2001 and 2002-2005. Weak correlations in line with preceding ANOVA table.

Regression of Log(Adjusted Hg) vs Year
SUCK at Station 5, Dooms, VA near Rt. 611 bridge (above dam)

model	rsquare	ratio/rsqr	Source	DF	FValue	ProbF
Year	.	.	Model	12	17.11	<.0001
Year	.	.	Error	119	_	_
Year	.	.	Corrected Total	131	_	_
Year	0.633048	.	R-Square	.	.	.
Hydro	.	.	Model	4	4.88	0.0011
Hydro	.	.	Error	118	_	_
Hydro	.	.	Corrected Total	122	_	_
Hydro	0.142053	22	R-Square	.	.	.
Hydro	.	.	harriston0	1	1.91	0.1698
Hydro	.	.	harriston1	1	0.27	0.6023
Hydro	.	.	harriston2	1	0.13	0.7141
Hydro	.	.	harriston3	1	18.50	<.0001

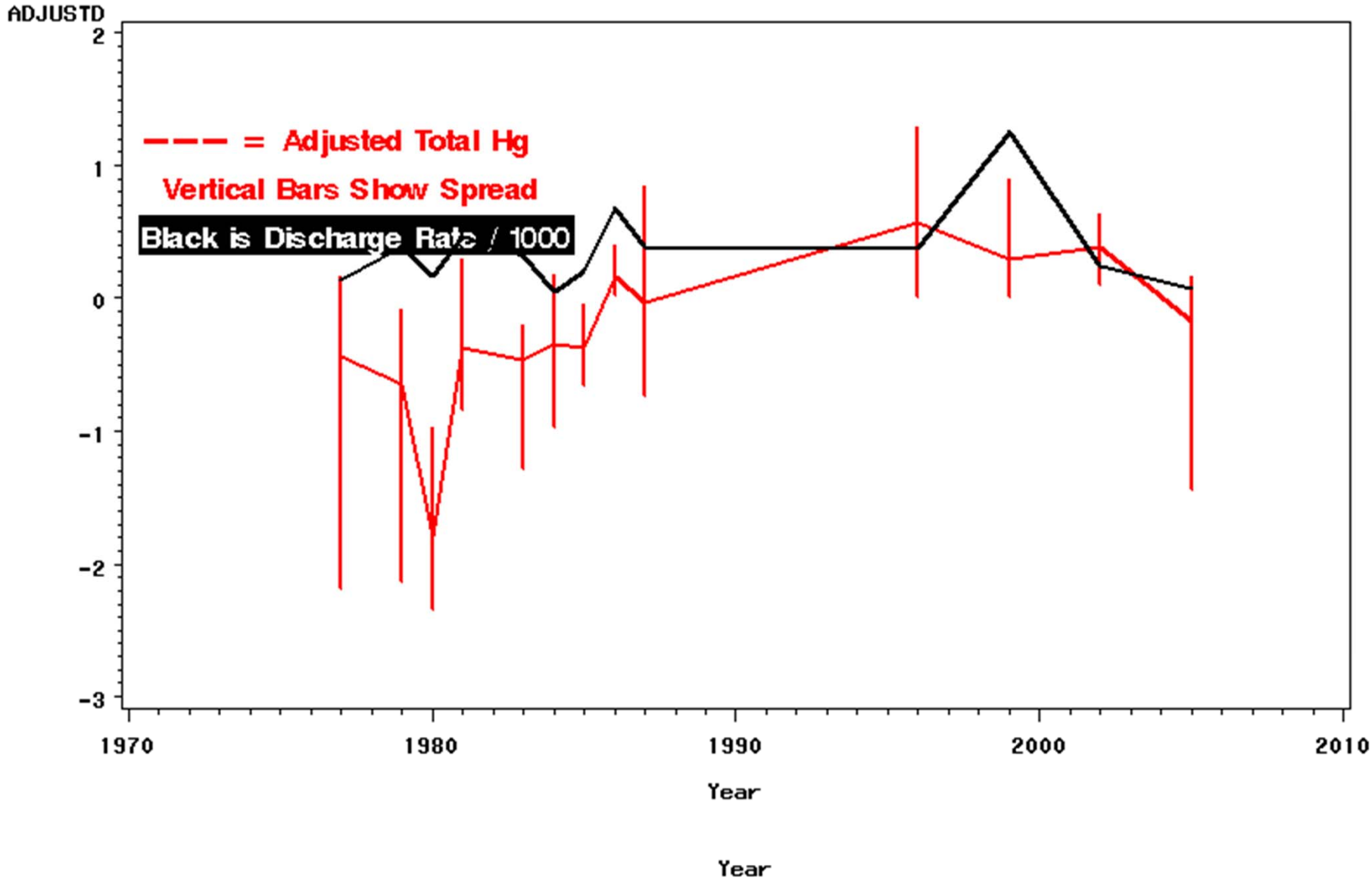
Parameter	Estimate	StdErr	tValue	Probt
Intercept	-.5869554382	0.16737606	-3.51	0.0006
harriston0	0.0000183933	0.00001332	1.38	0.1698
harriston1	-.0000082671	0.00001582	-0.52	0.6023
harriston2	-.0000050969	0.00001388	-0.37	0.7141
harriston3	0.0000859273	0.00001998	4.30	<.0001



Partial regression plot showing relationship of THg vs Lag3 discharge rate after correcting for lags 0, 1, and 2. This corresponds to ANOVA table on previous slide.

Log(Adjusted Hg) vs Maximum Daily Discharge Rate

SUCK at Station 5, Dooms, VA near Rt. 611 bridge (above dam)
Discharge Measured at Harriston, 3 Years Previous
Discharge rates Divided by 10000 for Ploting



Year

Slopes of Significant Regression
-------- species=SUNFISH--------

station	YEAR 0	1	2	3
3				
5				9.86
6		2.41		5.31
7			2.14	3.84
8				

Slopes of Significant Regression
-------- species=SMB ---------

station	YEAR 0	1	2	3
3				
5		-6.54		6.48
6		-5.92		4.02
7	-2.81		3.28	5.32
8	3.42			

Slopes of Significant Regression
-------- species=REDBREAST------

station	YEAR 0	1	2	3
3				13.24
5	3.39			10.23
6		3.21	0.96	5.26
7			-2.92	5.27
8		3.85		

Slopes of Significant Regression
-------- species=SUCKER ------

station	YEAR 0	1	2	3
3		-9.12	5.15	
5				8.59
6	2.37		-2.84	8.53
7	1.99	2.24		2.37
8		-6.97		10.75

There is some consistency in the slopes wrt discharge 3 years previous at stations 5, 6, 7 (Dooms, Crimora, Grottoes).

Note: Slopes multiplied by 100000 for easy reference.

Summary of Significant Regressions

------species=SUNFISH ------

station	YEAR 0	1	2	3
3	0	0	0	0
5	0	0	0	1
6	0	1	0	1
7	0	0	1	1
8	0	0	0	0

-------- species=SMB ----------

station	YEAR 0	1	2	3
3	0	0	0	0
5	0	-1	0	1
6	0	-1	0	1
7	-1	0	1	1
8	1	0	0	0

------species=REDBREAST ---

station	YEAR 0	1	2	3
3	0	0	0	1
5	1	0	0	1
6	0	1	1	1
7	0	0	-1	1
8	0	1	0	0

-------- species=SUCKER -------

station	YEAR 0	1	2	3
3	0	-1	1	0
5	0	0	0	1
6	1	0	-1	1
7	1	1	0	1
8	0	-1	0	1

1=significant positive correlation
-1=significant negative correlation
0=non-significant correlation

Summary of Significant Regressions

------species=LMB ----------

station	YEAR 0	1	2	3
3	1	0	0	0
5	0	0	0	0
6	0	0	0	0
7	1	-1	1	-1
8	-1	1	0	0

1=significant positive correlation
-1=significant negative correlation
0=non-significant correlation

There were relatively few large mouth bass caught at these stations (next slide), which may account for the different patterns for this species.

Fish Tissue Sample Sizes

	Full Sample Size				
	species				
station	LMB	REDB	SMB	SUCK	SUN
3	3	76	61	160	150
5	44	104	34	132	192
6	28	386	48	168	521
7	22	89	90	137	172
8	9	103	88	127	167

Summary

- **A significant percent of variation in adjusted fish tissue Hg is "explained" by the maximum daily discharge rate in the 3 years prior to fish sampling**
  - **In most cases, there is an apparent *3-year lag* between high discharge rates and high Hg**
- **Note: The month of fish sampling is often not known**
  - **Adds some vagueness to time lag**
  - **Some current year sampling may come after time of highest discharge rate**

Possible Follow-up

- **Use maximum of 2- or 3-day moving average instead of maximum daily rate**
  - **Sustained storm vs limited duration**
- **Date time from fish sample, not calendar year**
- **Regress only over period 1977-1987**
  - **Yearly fish sample available, allows exploration of apparent 3-year time delay**
  - **Period may be too short to be meaningful**

Additional Results

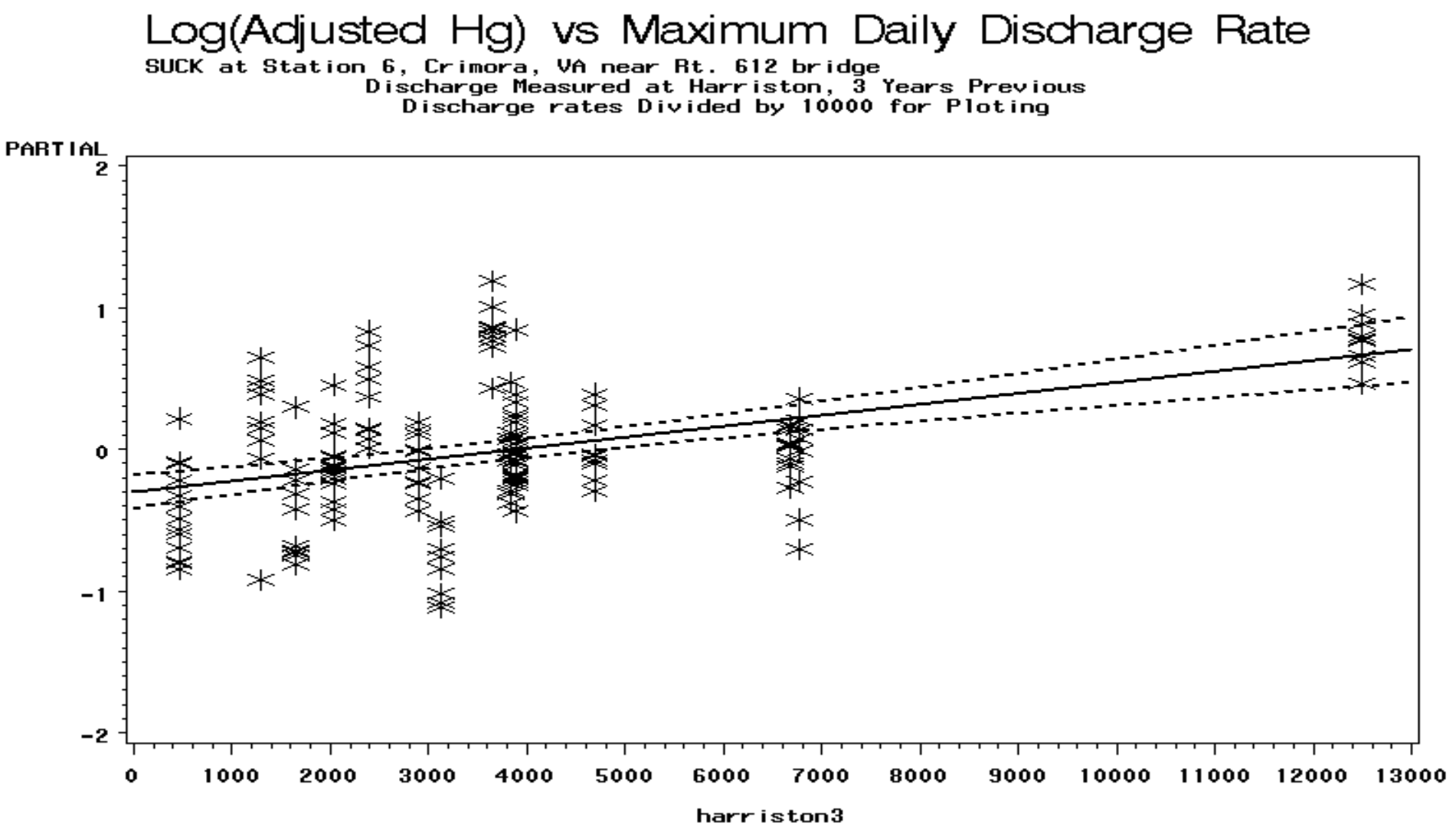
- **Additional, similar slides are provided for the other species (Sunfish, LMB, Redbreast, more sucker)**

- **An appendix is given indicating how the assumptions underlying the modeling were done and what the results were**

Regression of Log(Adjusted Hg) vs Year
SUCK at Station 6, Crimora, VA near Rt. 612 bridge

model	rsqr	ratio/rsqr	Source	DF	FValue	ProbF
Year	.	.	Model	14	29.15	<.0001
Year	.	.	Error	153	_	_
Year	.	.	Corrected Total	167	_	_
Year	0.727317	.	R-Square	.	.	.
Hydro	.	.	Model	4	16.00	<.0001
Hydro	.	.	Error	154	_	_
Hydro	.	.	Corrected Total	158	_	_
Hydro	0.293609	40	R-Square	.	.	.
Hydro	.	.	harriston0	1	9.06	0.0030
Hydro	.	.	harriston1	1	0.05	0.8273
Hydro	.	.	harriston2	1	15.20	0.0001
Hydro	.	.	harriston3	1	44.42	<.0001

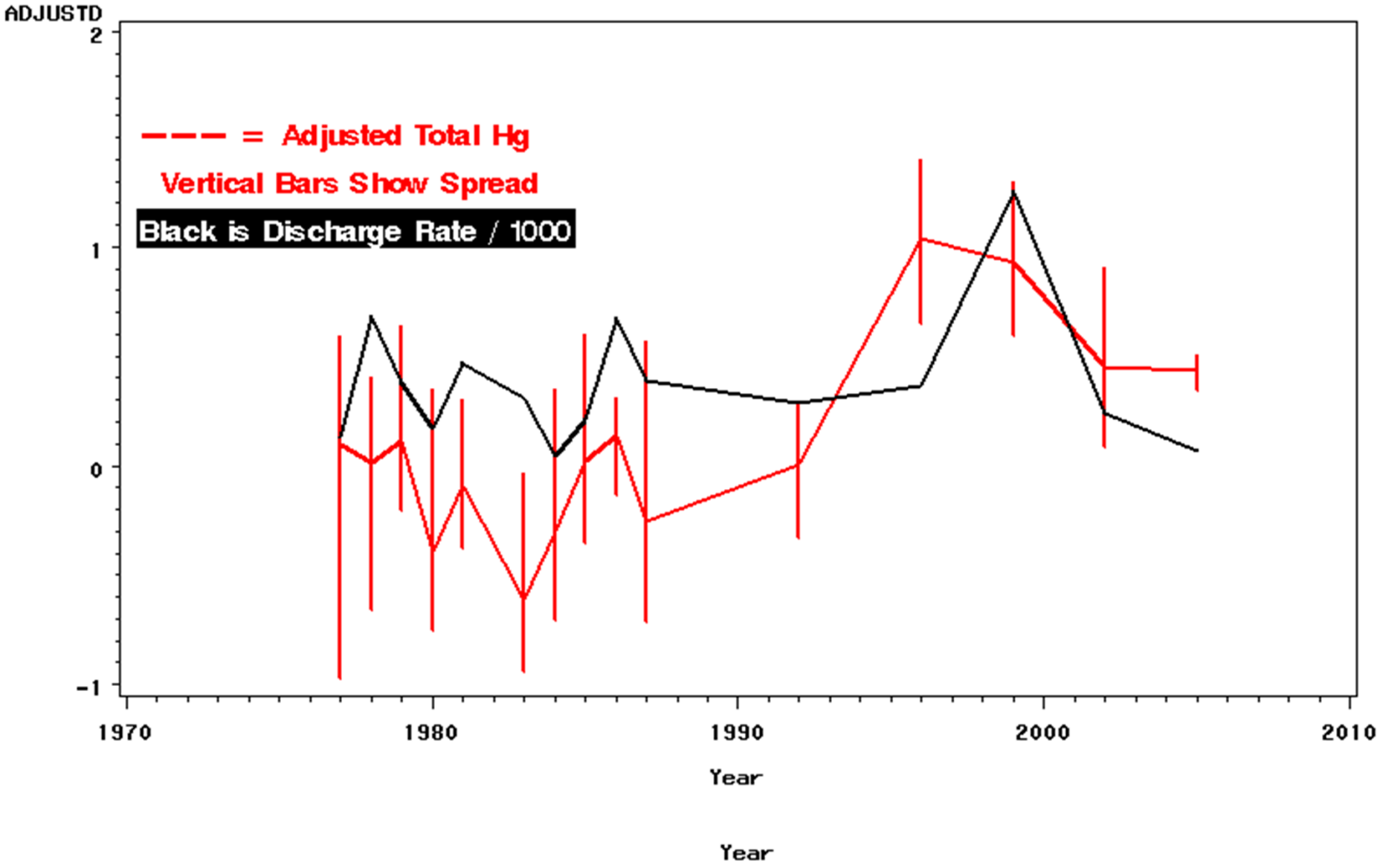
Parameter	Estimate	StdErr	tValue	Probt
Intercept	-.2633878561	0.10327152	-2.55	0.0117
harriston0	0.0000236978	0.00000787	3.01	0.0030
harriston1	-.0000021183	0.00000969	-0.22	0.8273
harriston2	-.0000284142	0.00000729	-3.90	0.0001
harriston3	0.0000853321	0.00001280	6.66	<.0001



Partial regression plot showing relationship of THg vs Lag3 discharge rate after correcting for lags 0, 1, and 2. This corresponds to ANOVA table on previous slide.

Log(Adjusted Hg) vs Maximum Daily Discharge Rate

SUCK at Station 6, Crimora, VA near Rt. 612 bridge
Discharge Measured at Harriston, 3 Years Previous
Discharge rates Divided by 10000 for Ploting

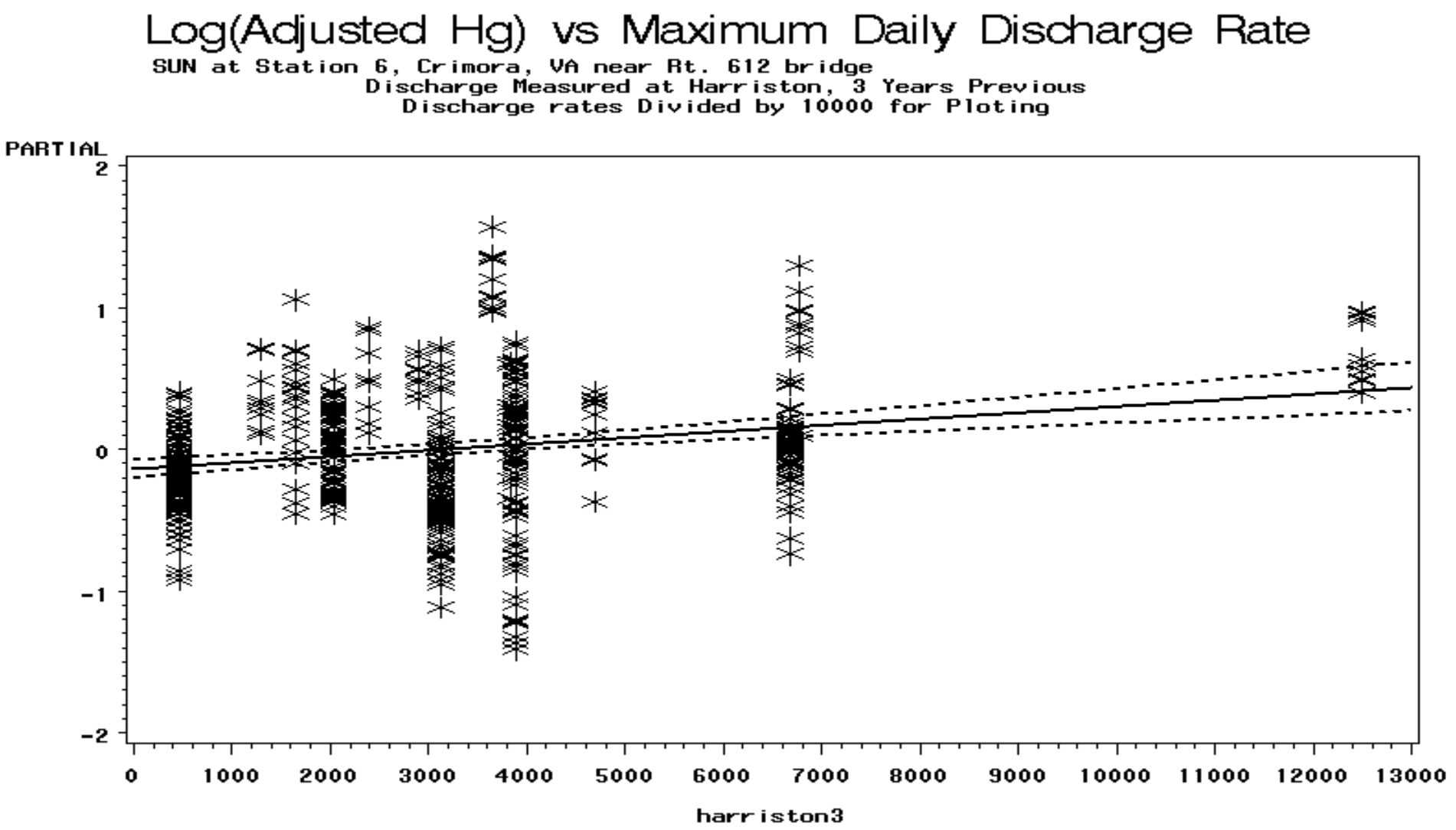


Year

Regression of Log(Adjusted Hg) vs Year
SUN at Station 6, Crimora, VA near Rt. 612 bridge

model	rsquare	ratio/rsqr	Source	DF	FValue	ProbF
Year	.	.	Model	15	34.34	<.0001
Year	.	.	Error	505	_	_
Year	.	.	Corrected Total	520	_	_
Year	0.504942	.	R-Square	.	.	.
Hydro	.	.	Model	4	22.70	<.0001
Hydro	.	.	Error	507	_	_
Hydro	.	.	Corrected Total	511	_	_
Hydro	0.151891	30	R-Square	.	.	.
Hydro	.	.	harriston0	1	0.11	0.7358
Hydro	.	.	harriston1	1	20.68	<.0001
Hydro	.	.	harriston2	1	3.63	0.0574
Hydro	.	.	harriston3	1	32.03	<.0001

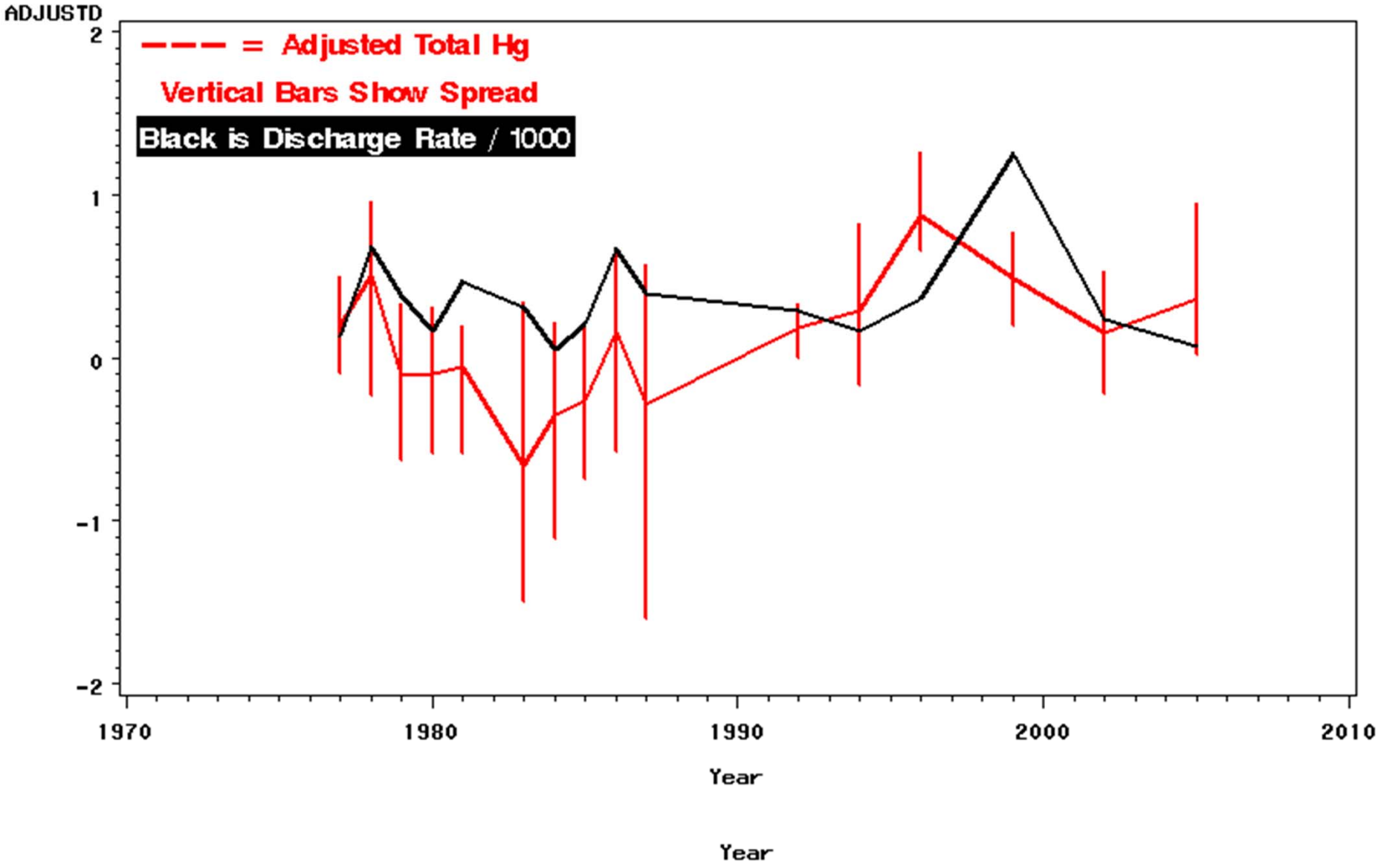
Parameter	Estimate	StdErr	tValue	Probt
Intercept	-.5316221515	0.05914700	-8.99	<.0001
harriston0	-.0000016259	0.00000482	-0.34	0.7358
harriston1	0.0000241113	0.00000530	4.55	<.0001
harriston2	0.0000076074	0.00000399	1.90	0.0574
harriston3	0.0000530703	0.00000938	5.66	<.0001



Partial regression plot showing relationship of THg vs Lag3 discharge rate after correcting for lags 0, 1, and 2. This corresponds to ANOVA table on previous slide.

Log(Adjusted Hg) vs Maximum Daily Discharge Rate

SUN at Station 6, Crimora, VA near Rt. 612 bridge
Discharge Measured at Harriston, 3 Years Previous
Discharge rates Divided by 10000 for Ploting

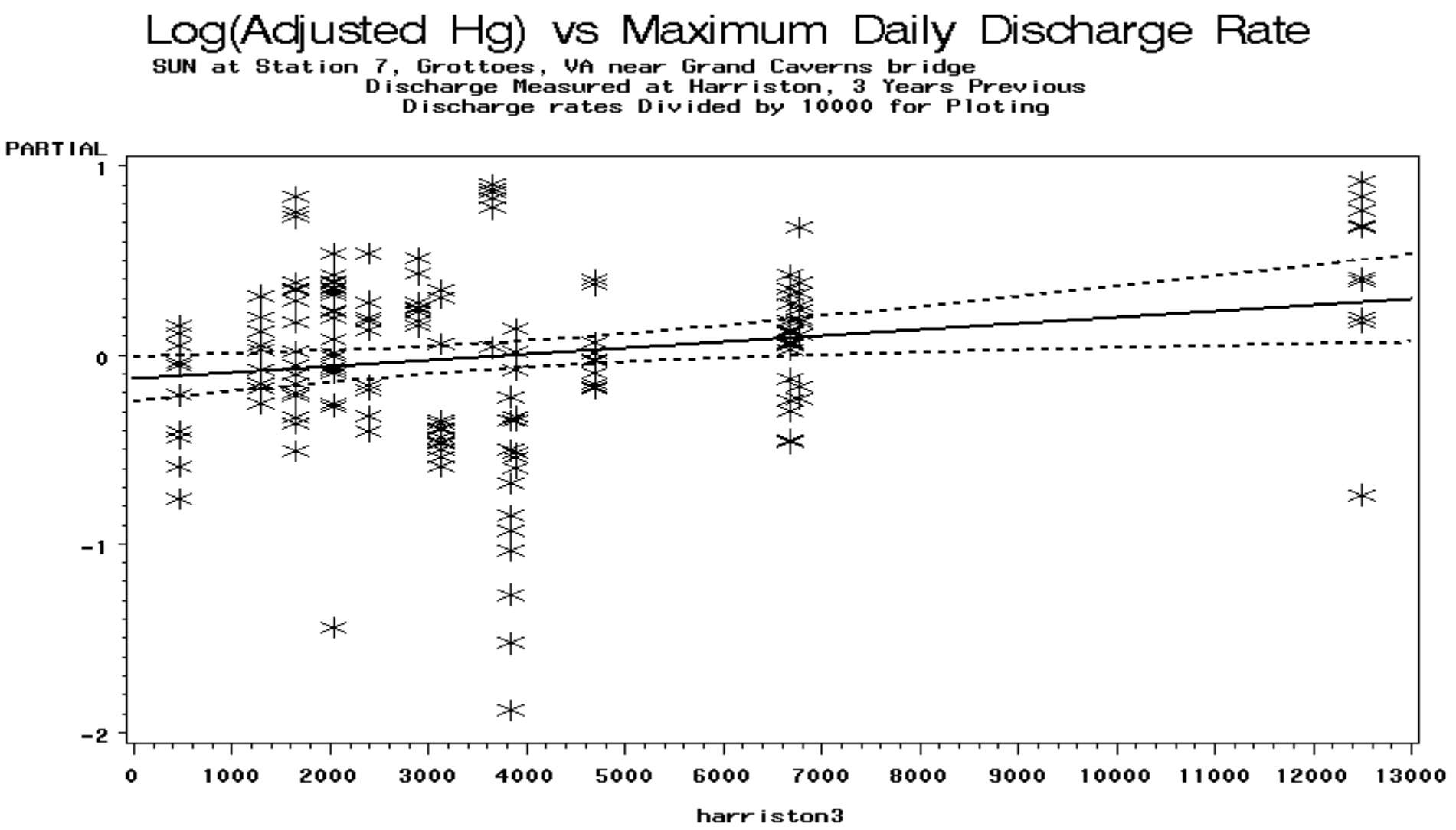


Year

Regression of Log(Adjusted Hg) vs Year
SUN at Station 7, Grottoes, VA near Grand Caverns bridge

model	rsquare	ratio/rsqr	Source	DF	FValue	ProbF
Year	.	.	Model	15	15.72	<.0001
Year	.	.	Error	156	_	_
Year	.	.	Corrected Total	171	_	_
Year	0.601895	.	R-Square	.	.	.
Hydro	.	.	Model	4	4.85	0.0010
Hydro	.	.	Error	158	_	_
Hydro	.	.	Corrected Total	162	_	_
Hydro	0.109319	18	R-Square	.	.	.
Hydro	.	.	harriston0	1	0.68	0.4123
Hydro	.	.	harriston1	1	2.13	0.1464
Hydro	.	.	harriston2	1	4.75	0.0308
Hydro	.	.	harriston3	1	8.10	0.0050

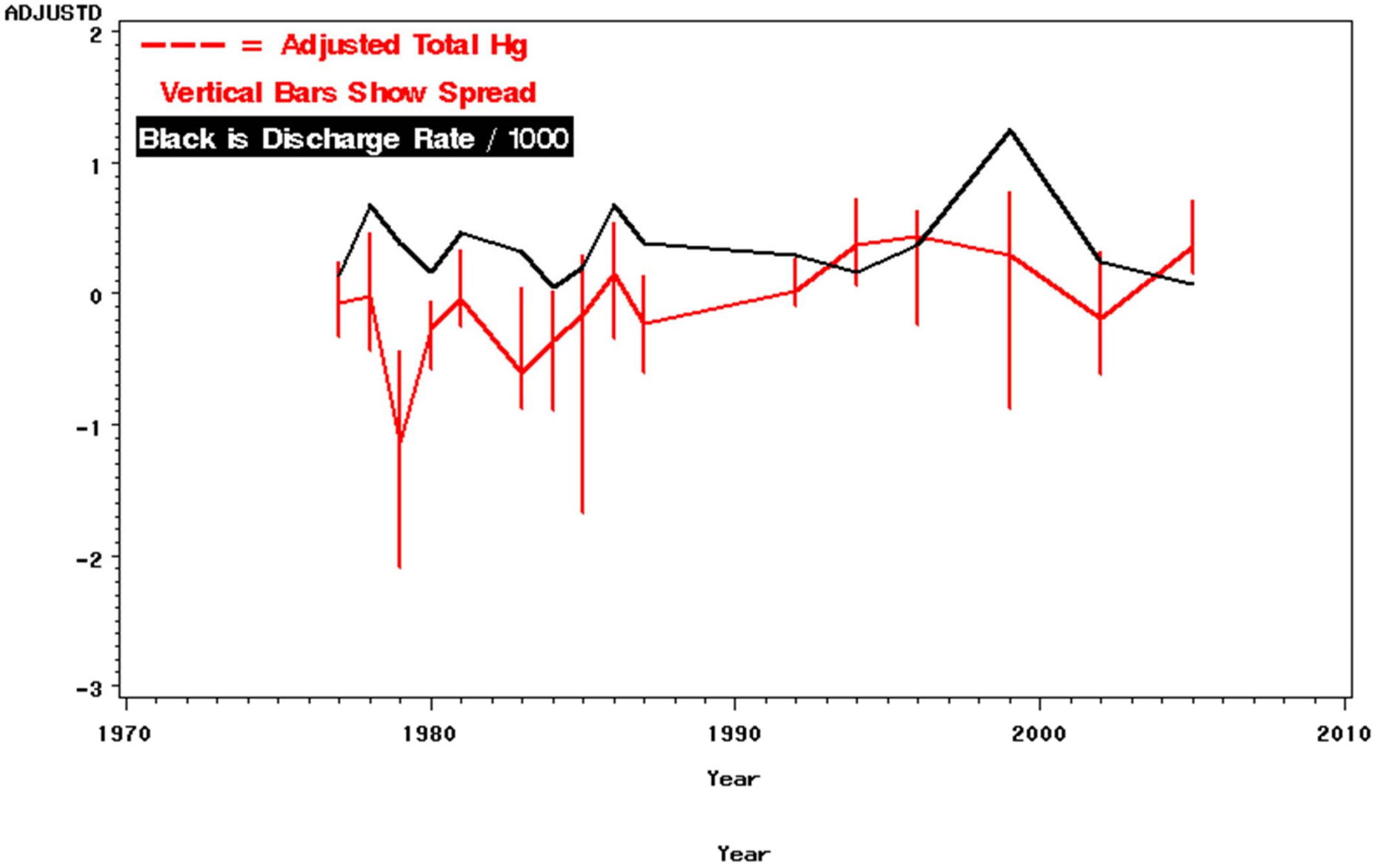
Parameter	Estimate	StdErr	tValue	Probt
Intercept	-.4198366322	0.10326778	-4.07	<.0001
harriston0	-.0000064860	0.00000789	-0.82	0.4123
harriston1	0.0000132981	0.00000911	1.46	0.1464
harriston2	0.0000213795	0.00000981	2.18	0.0308
harriston3	0.0000384239	0.00001350	2.85	0.0050



Partial regression plot showing relationship of THg vs Lag3 discharge rate after correcting for lags 0, 1, and 2. This corresponds to ANOVA table on previous slide.

Log(Adjusted Hg) vs Maximum Daily Discharge Rate

SUN at Station 7, Grottoes, VA near Grand Caverns bridge
Discharge Measured at Harriston, 3 Years Previous
Discharge rates Divided by 10000 for Ploting

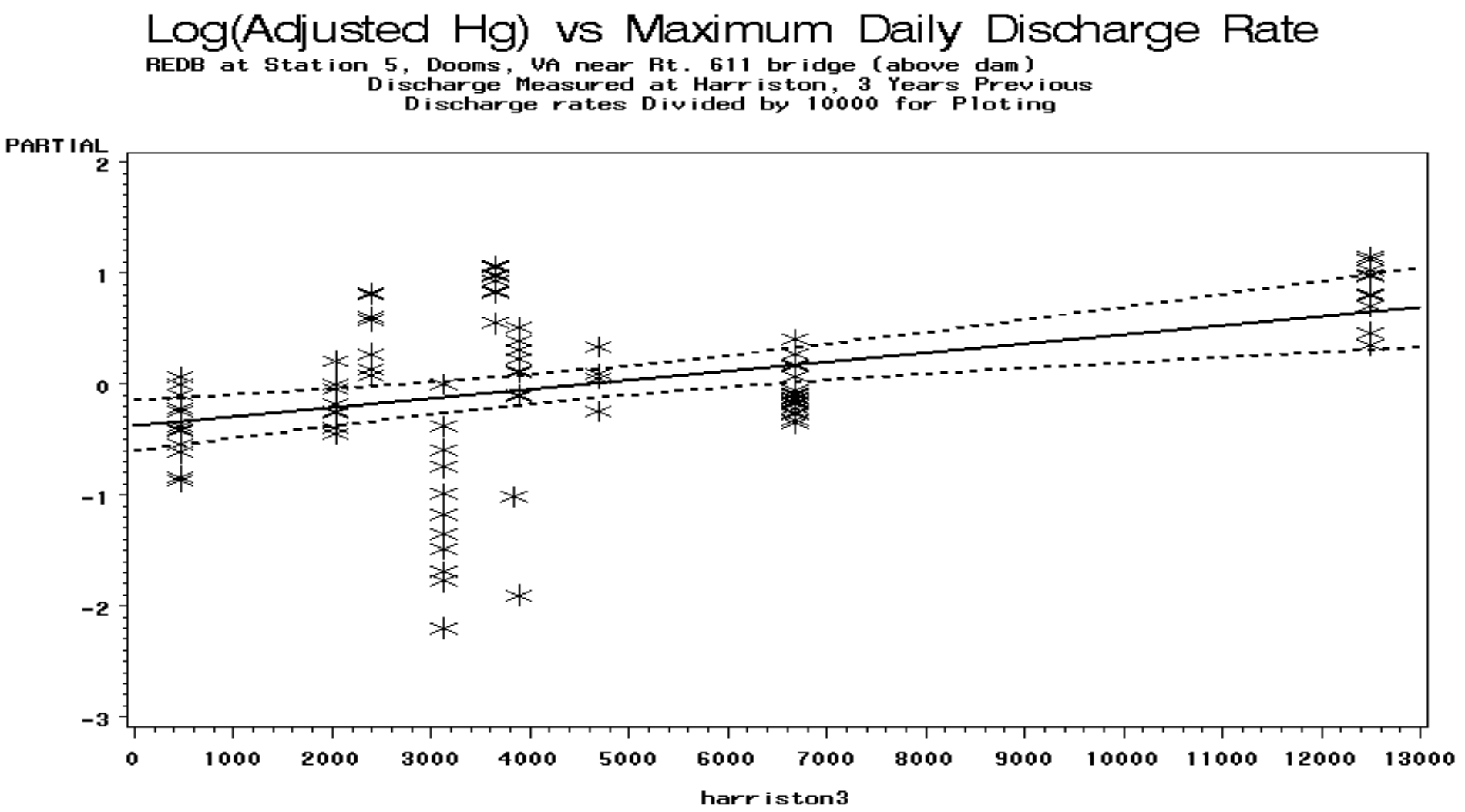


Year

Regression of Log(Adjusted Hg) vs Year
REDB at Station 5, Dooms, VA near Rt. 611 bridge (above dam)

model	rsquare	ratio/rsqr	Source	DF	FValue	ProbF
Year	.	.	Model	10	27.55	<.0001
Year	.	.	Error	93	_	_
Year	.	.	Corrected Total	103	_	_
Year	0.747618	.	R-Square	.	.	.
Hydro	.	.	Model	4	5.99	0.0003
Hydro	.	.	Error	90	_	_
Hydro	.	.	Corrected Total	94	_	_
Hydro	0.210316	28	R-Square	.	.	.
Hydro	.	.	harriston0	1	4.65	0.0338
Hydro	.	.	harriston1	1	0.02	0.8844
Hydro	.	.	harriston2	1	0.67	0.4135
Hydro	.	.	harriston3	1	21.08	<.0001

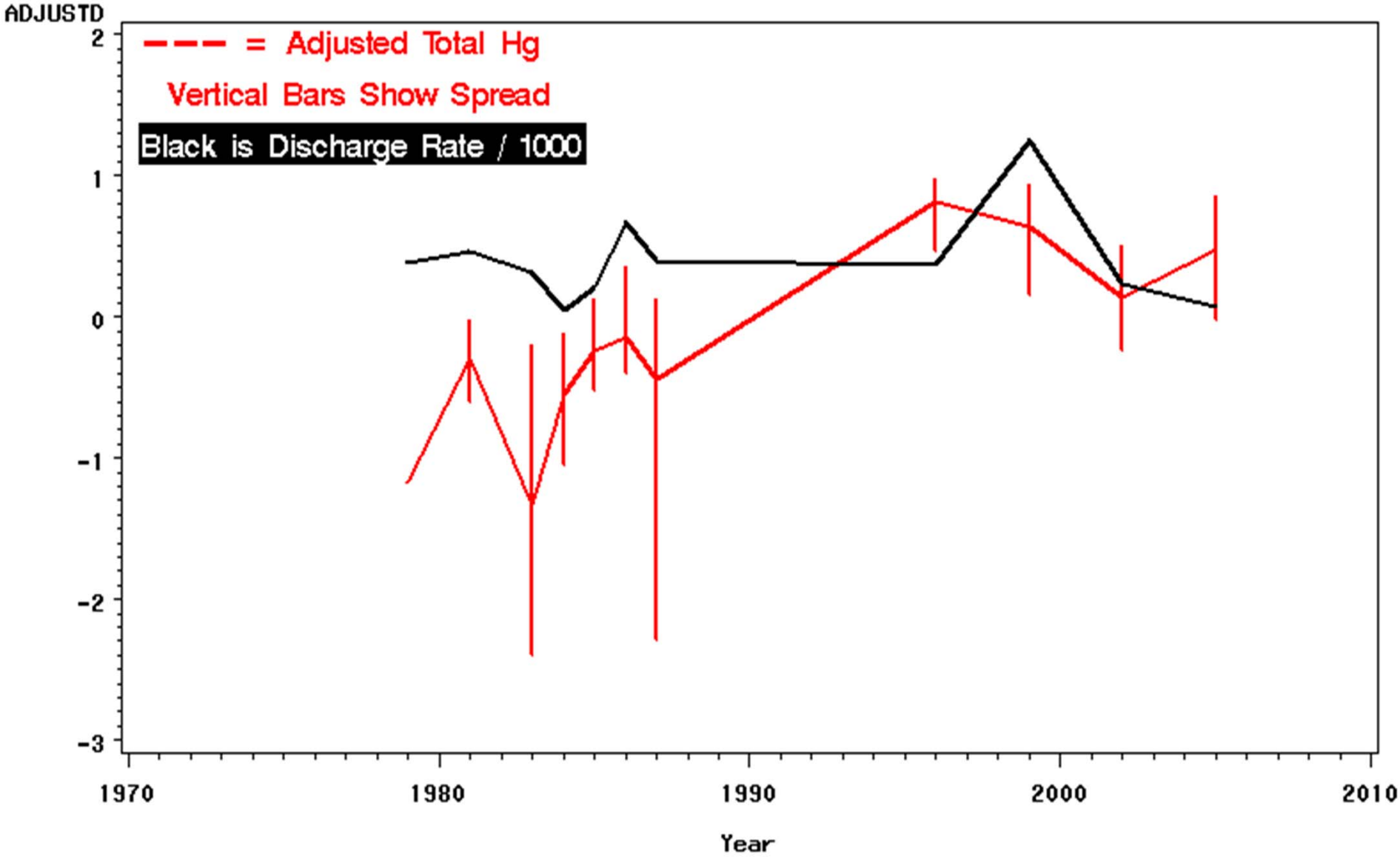
Parameter	Estimate	StdErr	tValue	Probt
Intercept	-.7906267198	0.19378276	-4.08	<.0001
harriston0	0.0000338906	0.00001572	2.16	0.0338
harriston1	0.0000020405	0.00001399	0.15	0.8844
harriston2	-.0000127505	0.00001552	-0.82	0.4135
harriston3	0.0001023029	0.00002228	4.59	<.0001



Partial regression plot showing relationship of THg vs Lag3 discharge rate after correcting for lags 0, 1, and 2. This corresponds to ANOVA table on previous slide.

Log(Adjusted Hg) vs Maximum Daily Discharge Rate

REDB at Station 5, Dooms, VA near Rt. 611 bridge (above dam)
Discharge Measured at Harriston, 3 Years Previous
Discharge rates Divided by 10000 for Ploting



Appendix

- **The following slides indicate the process followed to check the models for statistical appropriateness.**

Check of Assumptions Underlying ANCOVA, Regression

- **Normality checked by**
  - **Shapiro-Wilk test**
  - **QQ-plot**
  - **histogram with fitted normal pdf**
- **Variance homogeneity checked by**
  - **Levene test (in the so-called “W50” form)**
  - **Box plots**
- **Outliers (in total Hg) checked by**
  - **Tukey outlier rule**
  - **Re-analysis with outliers omitted**

Example Assumption Check

SUN at Station 7 Transform=LOG

EXPLANATORY VARIABLE IS YEAR

TESTS OF NORMALITY OF ADJUSTD: FULL DATA SET

Test	Label	Stat	pType	Sign	pValue
Shapiro-Wilk	W	0.939295	Pr < W	<	0.0001
Kolmogorov-Smirnov	D	0.070765	Pr > D		0.0346
Cramer-von Mises	W-Sq	0.138101	Pr > W-Sq		0.0358
Anderson-Darling	A-Sq	1.167841	Pr > A-Sq	<	0.0050

POSSIBLE OUTLIERS FROM ANOVA ON ADJUSTD

Obs	year	ADJUSTD	Pred	Resid	LB	UB
1	1979	-2.10215	-1.14665	-0.95551	-0.72261	0.76721
2	1985	-1.68159	-0.15912	-1.52247	-0.72261	0.76721
3	1999	-0.88512	0.28657	-1.17169	-0.72261	0.76721

LEVENE TEST FOR ADJUSTD

Effect	DF	LEVENE	P_VALUE
year	15	1.45033	0.13063

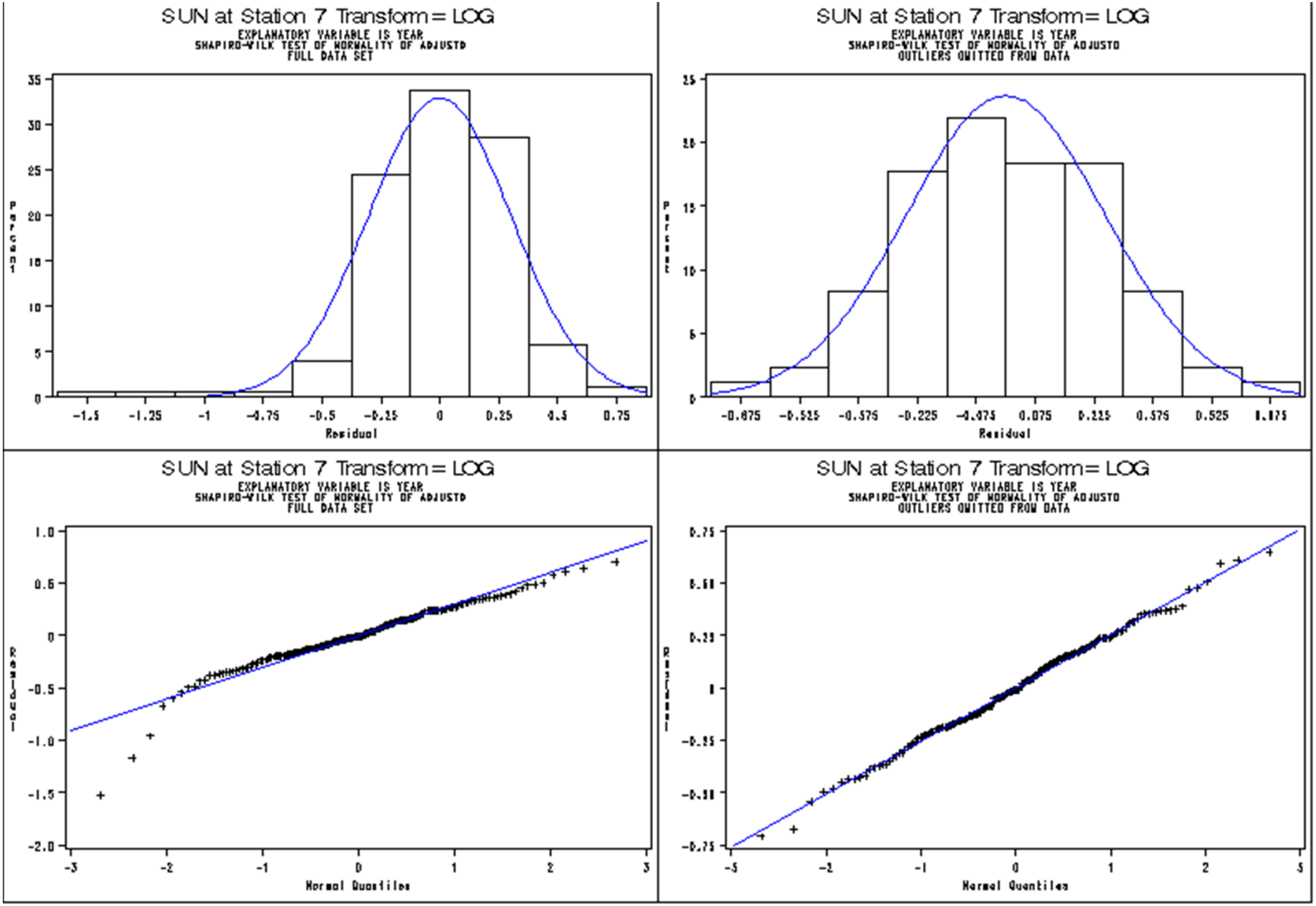
SUN at Station 7 Transform=LOG
EXPLANATORY VARIABLE IS YEAR
TESTS OF NORMALITY OF ADJUSTD: Outlier Omitted DATA SET

Test	Label	Stat	pType	Sign	pValue
Shapiro-Wilk	W	0.995791	Pr < W		0.9184
Kolmogorov-Smirnov	D	0.039848	Pr > D	>	0.1500
Cramer-von Mises	W-Sq	0.035728	Pr > W-Sq	>	0.2500
Anderson-Darling	A-Sq	0.217891	Pr > A-Sq	>	0.2500

LEVENE TEST FOR ADJUSTD

Effect	DF	LEVENE	P_VALUE
year	15	1.20951	0.27009

Omission of three low values eliminates the significant formal tests for normality.
The QQ-plot makes clear that the data are normally distributed.



Normal Quantiles

Tukey Outlier Summary

	Number of Outliers Found				
	species				
station	LMB	REDB	SMB	SUCK	SUN
3	0	1	1	8	3
5	1	4	4	3	13
6	1	8	6	3	16
7	1	2	2	3	3
8	0	5	1	4	7

Summary of Shapiro-Wilk Tests

Shapiro-Wilk Tests on Full Dataset

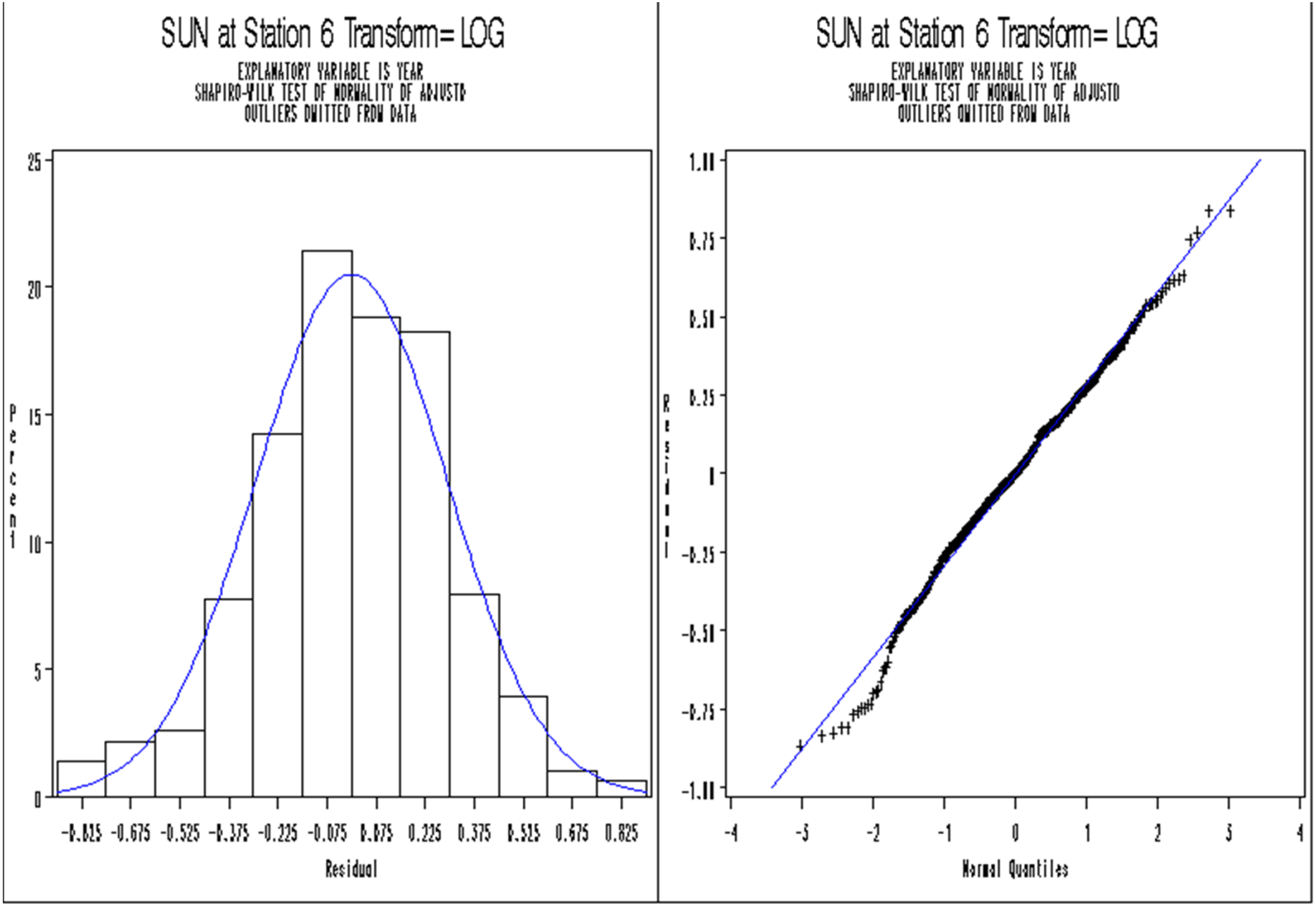
station	species LMB	REDB	SMB	SUCK	SUN
3	1.000	0.116	0.000	0.008	0.000
5	0.311	0.000	0.002	0.000	0.000
6	0.381	0.001	0.043	0.009	0.000
7	0.878	0.000	0.000	0.000	0.000
8	0.496	0.071	0.007	0.000	0.004

Shapiro-Wilk w/ Outliers Omitted

station	species LMB	REDB	SMB	SUCK	SUN
3	1.000	0.222	0.135	0.784	0.069
5	0.414	0.697	0.196	0.451	0.757
6	0.111	0.089	0.653	0.572	0.007
7	0.996	0.348	0.621	0.070	0.918
8	0.496	0.977	0.278	0.965	0.340

So all but one dataset tests as normal after removal of a small number of outliers (ranging from 0 to 8, plus one each of 13 or 16).

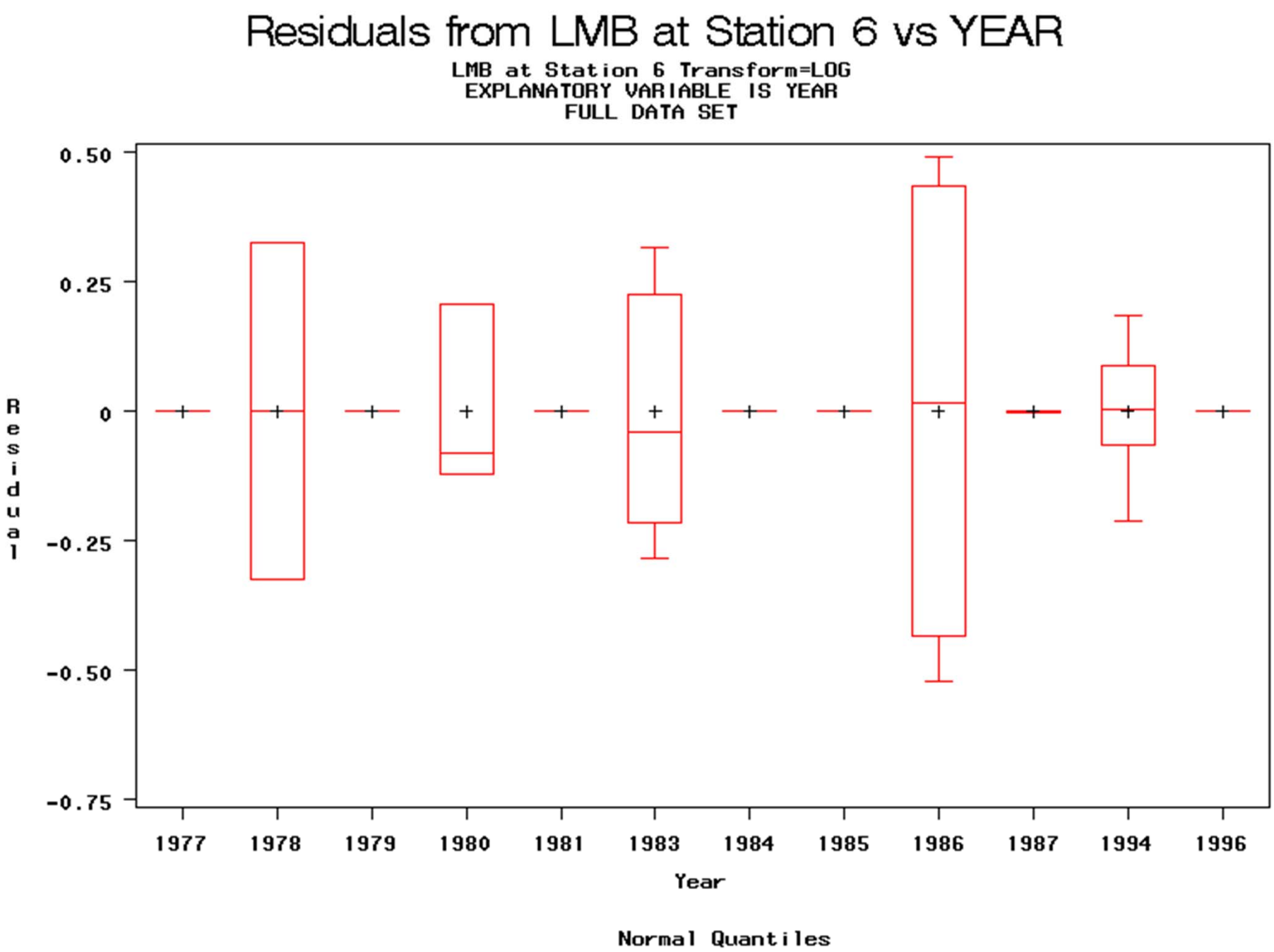
The plots following for the sole exception show a slight skewness, but little reason to question the applicability of normal-based methods.


SUN at Station 6 Transform= LOG
EXPLANATORY VARIABLE IS YEAR
SHAPIRO-WILK TEST OF NORMALITY OF ADJUSTD
OUTLIERS OMITTED FROM DATA
Percent
Residual
SUN at Station 6 Transform= LOG
EXPLANATORY VARIABLE IS YEAR
SHAPIRO-WILK TEST OF NORMALITY OF ADJUSTD
OUTLIERS OMITTED FROM DATA
Residual
Normal Quantiles

Residual

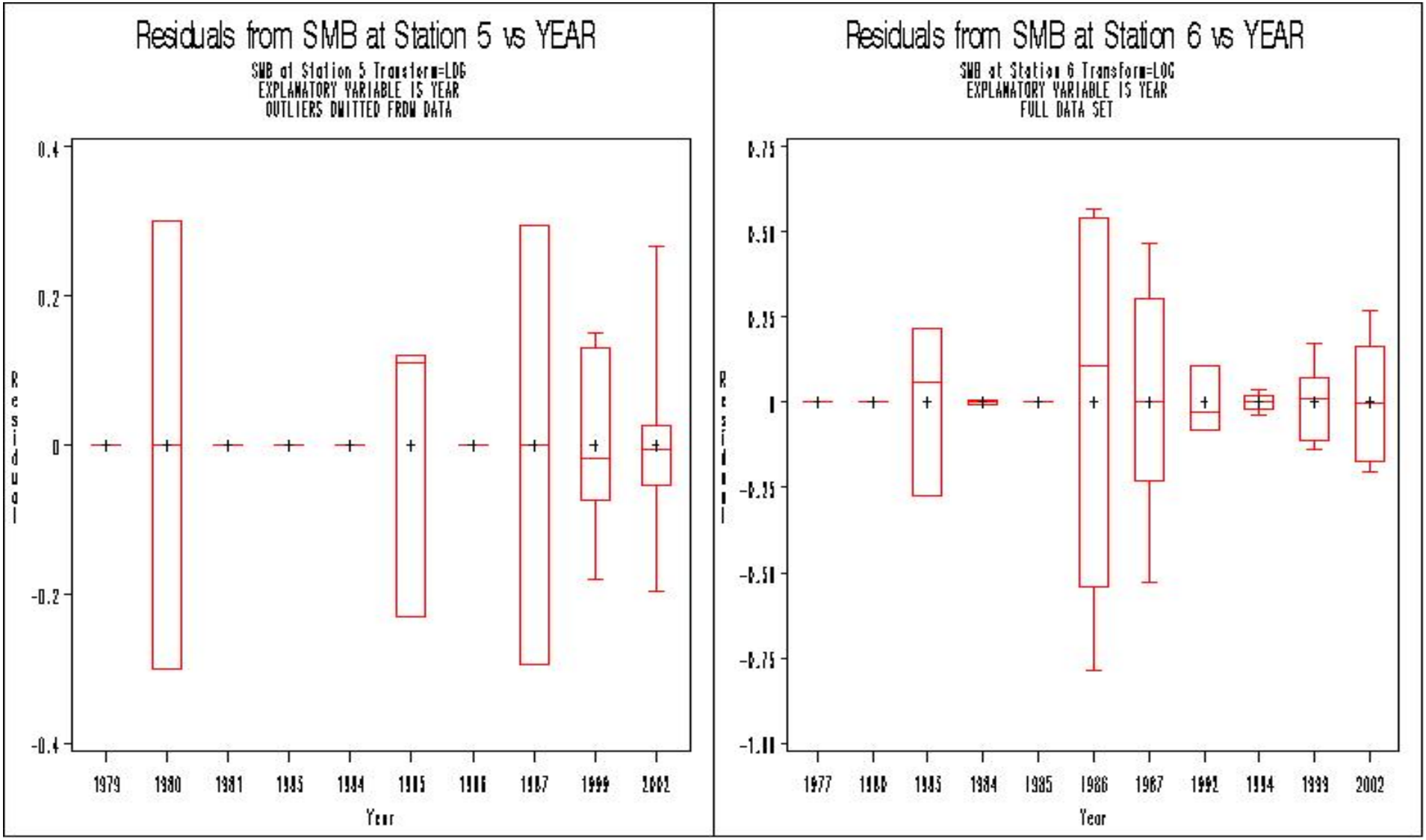
Example of Significant Levene Test for Variance Homogeneity

- **The next example makes clear that formal tests for variance homogeneity should not be taken as "true" without further examination.**
- **The Levene test is highly significant, indicating variance heterogeneity. The box plot makes clear the problem is several years with a single observation.**
  - **There is no meaningful variance heterogeneity in these data**

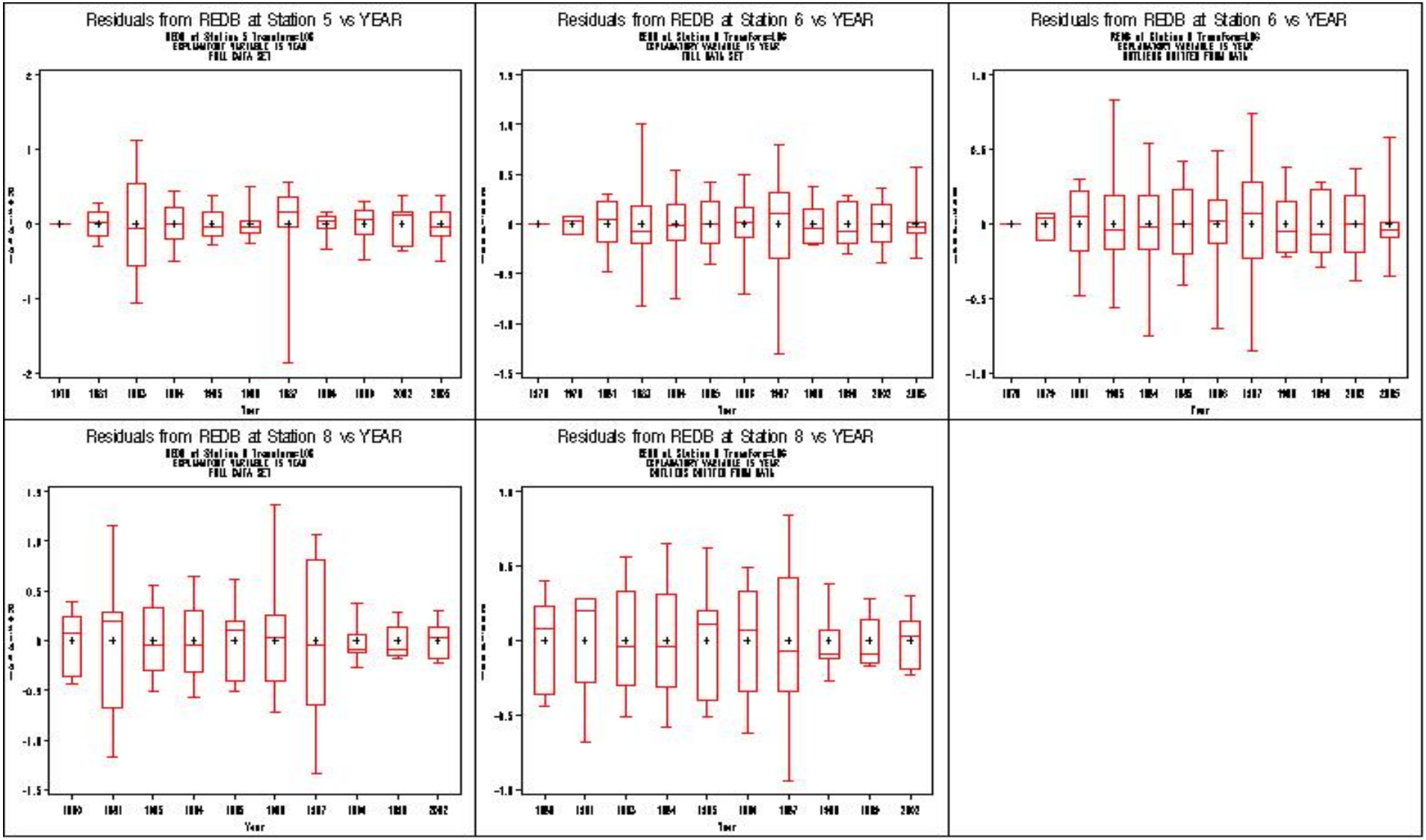

Residuals from LMB at Station 6 vs YEAR
LMB at Station 6 Transform=LOG
EXPLANATORY VARIABLE IS YEAR
FULL DATA SET
0.50
0.25
0
-0.25
-0.50
-0.75
Residual
1977
1978
1979
1980
1981
1983
1984
1985
1986
1987
1994
1996
Year
Normal Quantiles

Variance Boxplots

- **Boxplot drawn only where a significant Levene test reported**
  - **If outlier-omitted data had non-significant Levene test, boxplot not given for that**
- **If only an outlier-omitted boxplot is given then Levene test on full dataset was not significant and outlier-omitted analysis not needed**

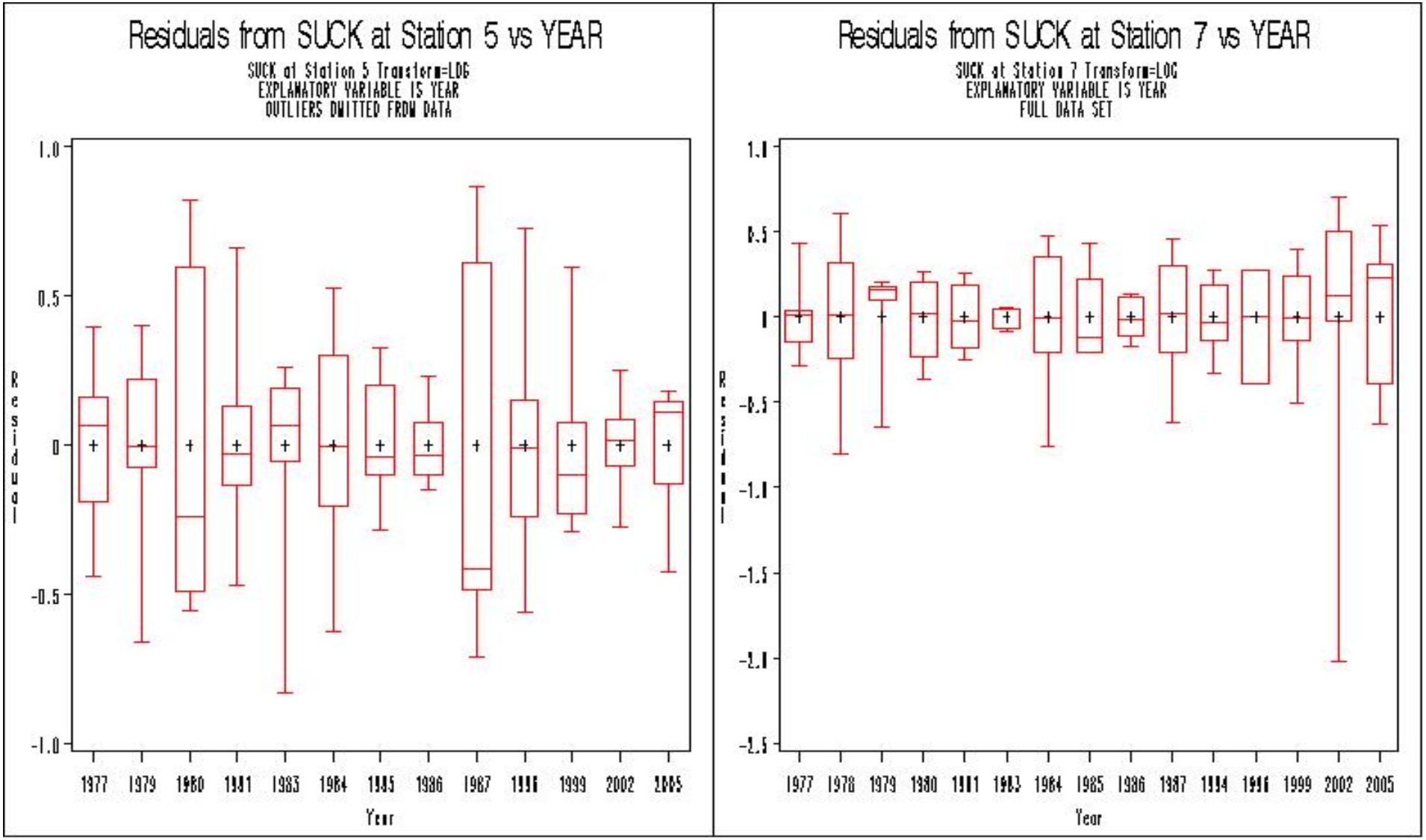


Significant Levene tests arise from years with only 1 or 2 observations



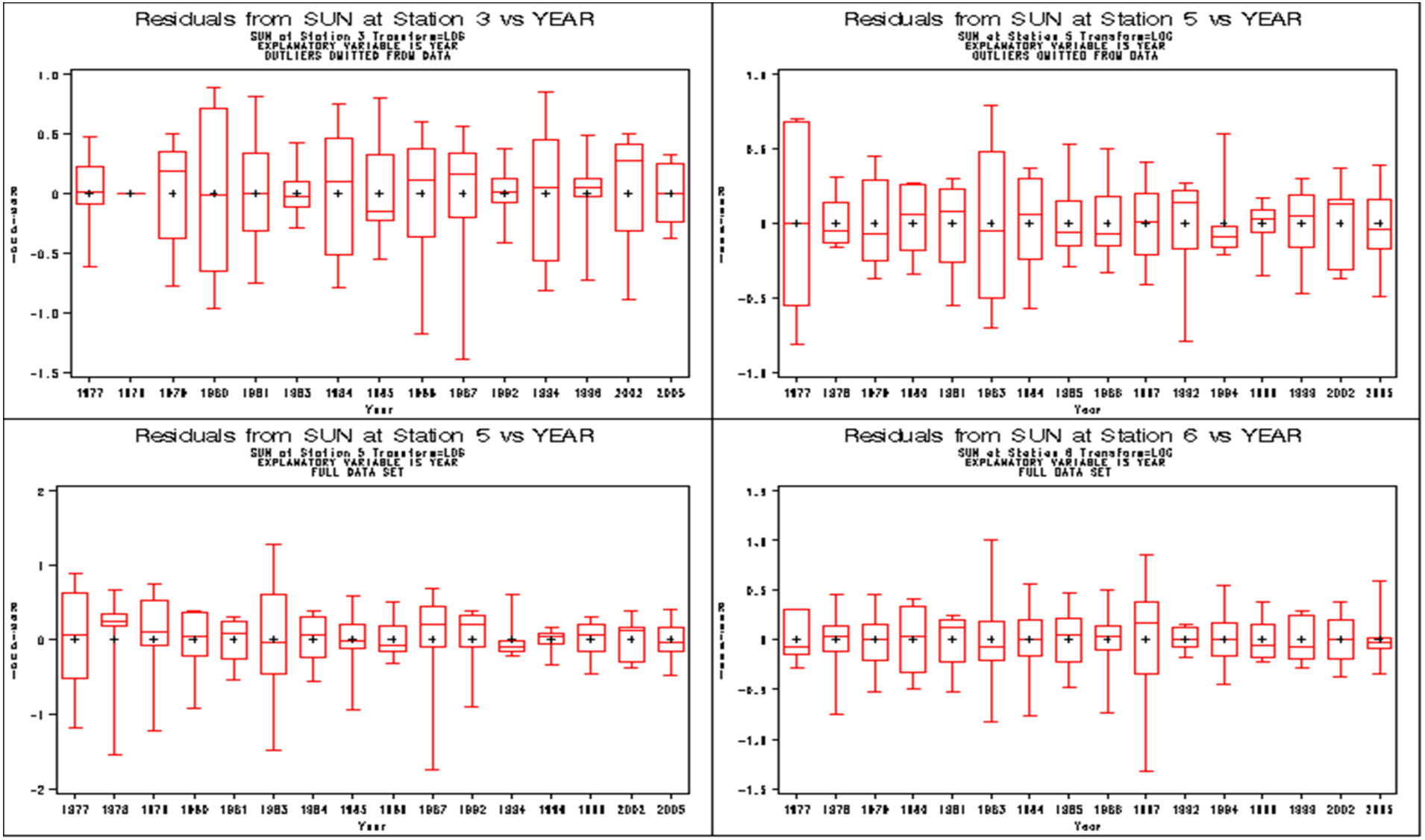
Significant Levene tests at stations 5 and 6 (top row) arise from year with only 1 observation.

Significant Levene test at station 8 is result of reduced variability in latest 3 years. ANOVA and regression results likely little affected.



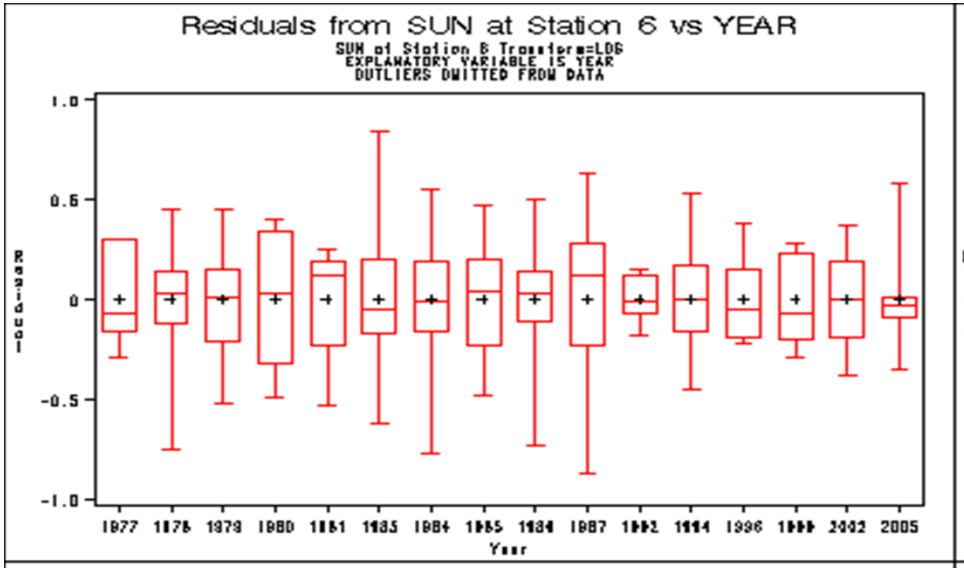
Significant Levene test for station 5 appears only in outlier-omitted analysis, so not relevant

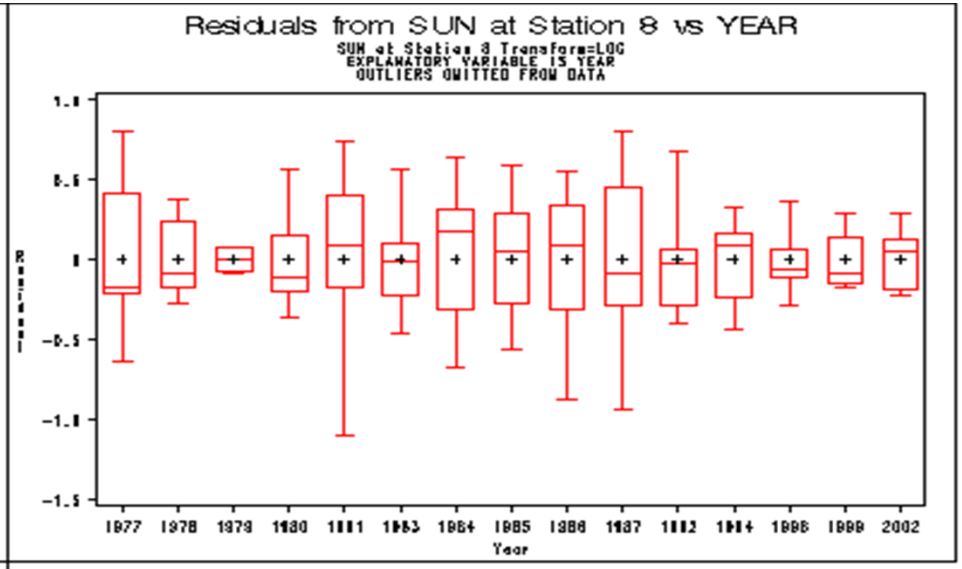
Significant Levene test for station 7 disappears in outlier-omitted analysis

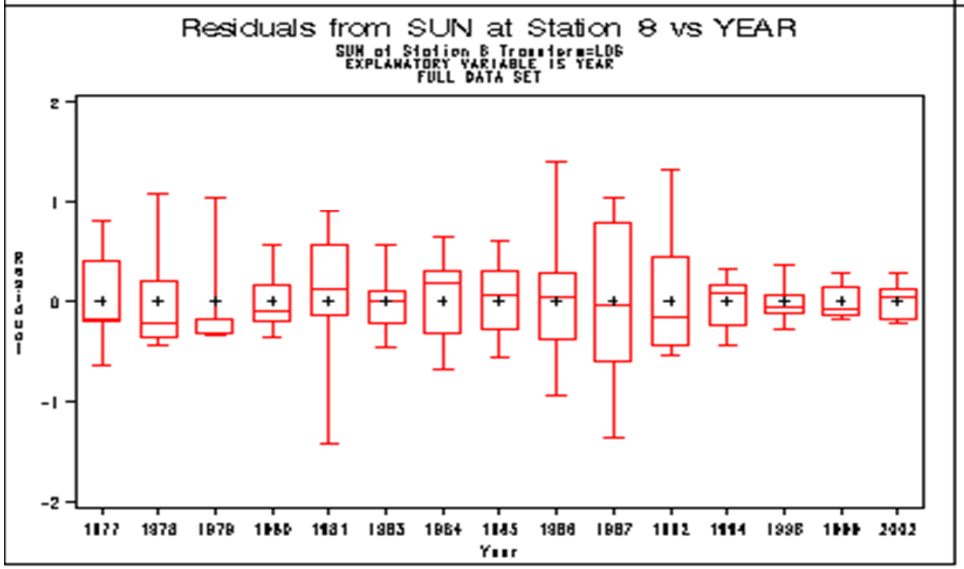


Significant Levene test for station 3 only occurs in outlier-omitted analysis, so irrelevant.

Significant Levene test for station 5 is cause for concern. Square-root transform eliminates problem, has very little effect on regression.

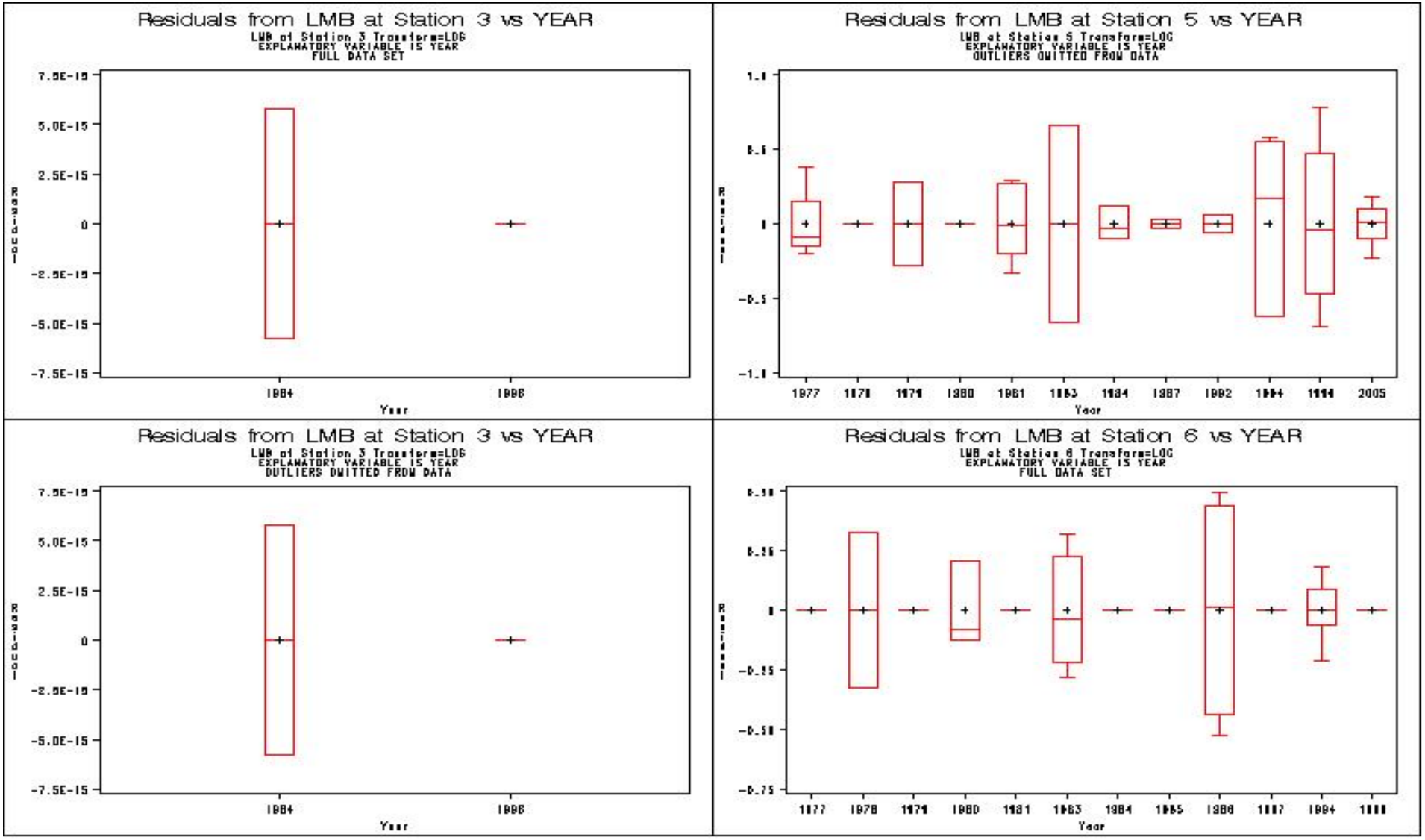






Significant Levene test for station 6 due to low variance in one year, 2005. Should not have much effect on results.

Significant Levene test for station 8 due primarily to low variance in one year, 1979. Should not have much effect on results.



All significant Levene tests arise from years with only 1 or 2 observations

Summary for Assumptions Check

- **Log-transform normalizes results (with a few outliers omitted), and stabilizes variances except for one species and station. Square-root transform eliminates that concern and has only trivial affect on regression results**
- **Virtue of common approach to all species and stations outweighs technical issue**