VIPER

NRCS' Newest Forecasting Tool

….a snow surveyor does forecasting

Adam Mazurkiewicz

Hetch Hetchy Water &Power

2008 California Co-op Snow Survey Meeting

SFPUC

WATER
WASTEWATER
POWER

Outline

- VIPER
- Statistical Analysis
- What does it look like?
- Getting started
- Optimization
- Performance
- Equations
- Summary

?

?

SFPUC
WATER
WASTEWATER
POWER

Credits

USDA Natural Resources Conservation Service
National Water & Climate Center

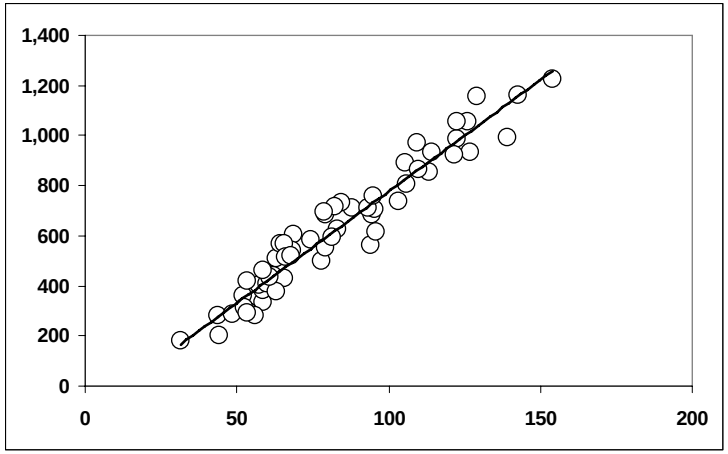
- *Project Coordinator*
  - Tom Perkins – Lead Forecast Hydrologist

- *Technical Contributors*
  - David Garen – Modeling Hydrologist
  - Tom Pagano – Forecast Hydrologist
  - Jim Marron – Resource Conservationist
  - Paul Jaeger – Programmer

SFPUC
WATER
WASTEWATER
POWER

VIPER

Visually Interactive Prediction and Estimation Routines



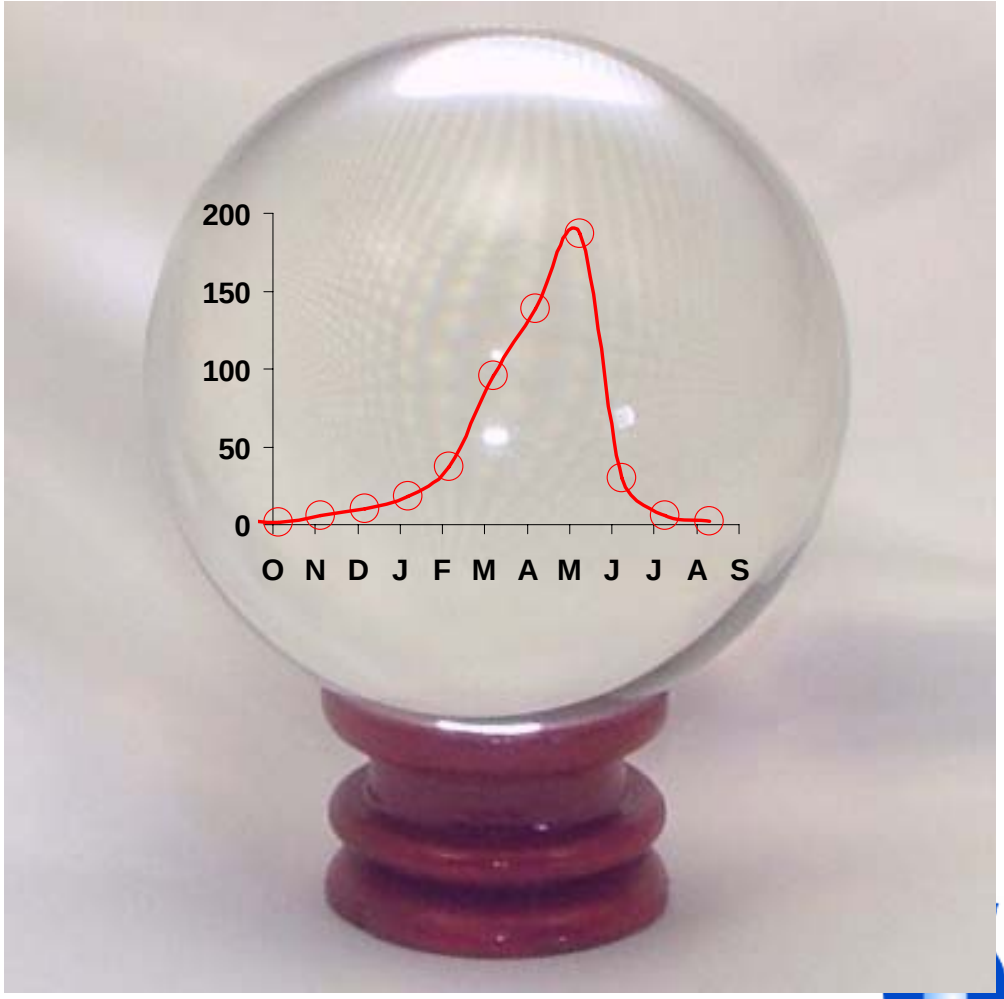
..a tool for developing forecast equations and a platform for executing operational forecasts

..not just another great hydrologic acronym

SFPUC
WATER
WASTEWATER
POWER

Why use VIPER?

- Old tools are clumsy
  - or dusty
- Expert developed
- Statistically sound
- Analytically robust



WATER
WASTEWATER
POWER

Key Features

- ***Excel Platform***
- ***Graphical output***
- ***Predictor search routines***
- ***Predictor time optimization***
- Report generation
- Exporting to GIS
- ***Saving and managing forecast equations/models***
- Configuration management
- *Automated data acquisition from web*
- Data management and editing
- ***Forecast coordination and management***
- Downstream relational forecasts

SFPUC
WATER
WASTEWATER
POWER

Statistics – PCA

$$PC_1 = \sum_{j=1}^{n} e_{1j} X_j$$

Principal Components Regression –

- Commonly used in situations where there are many intercorrelated independent variables.

- Principal components regression is just like standard regression except the independent variables are principal components rather than the original X variables.

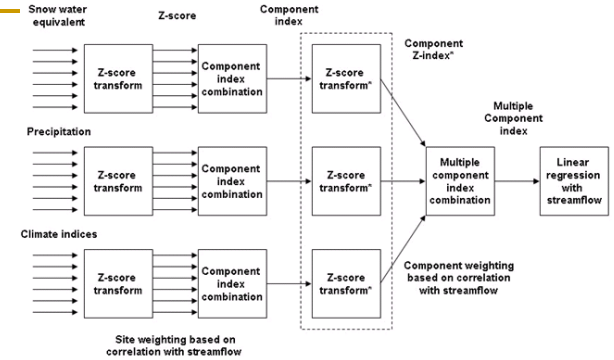
- Principal components are linear combinations of the X's.

$$PC_2 = \sum_{j=1}^{n} e_{2j} X_j$$

- The weights on variables within a given PC will be relatively large on some and noticeably smaller on others. A large number of variables can be reduced to a few PC's that still explain most of the variance in the data set.

SFPUC
WATER
WASTEWATER
POWER

Statistics – Z-score

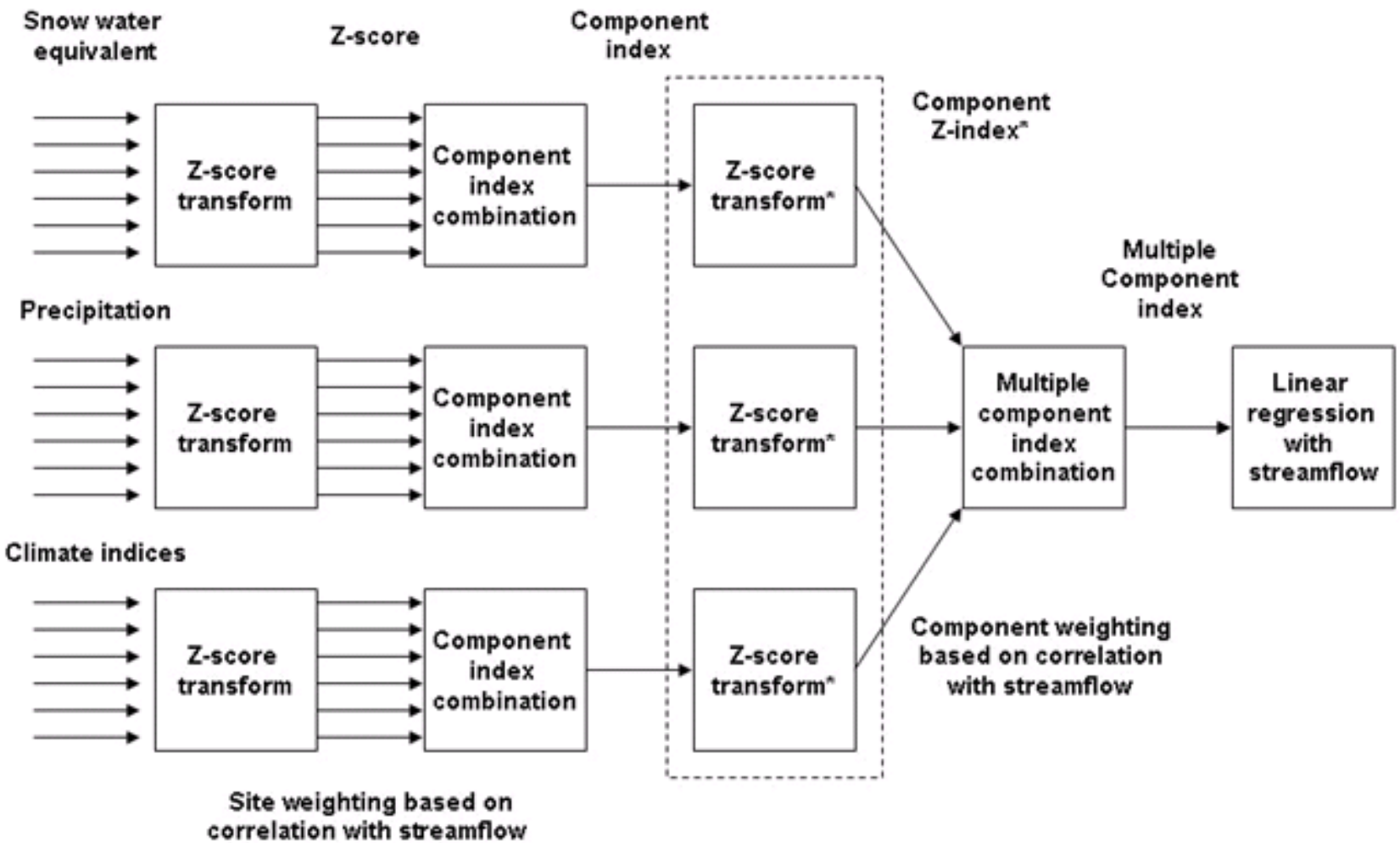


Z-score regression –

- A regression methodology that, within reason, can handle uneven record lengths and missing data.
- It groups stations into indices, emphasizing good stations, minimizing the effect of poor stations.
- Multiple signals can be managed (e.g. snow, fall precip, baseflow).
- Can be abused especially if the input data set is highly uneven.

SFPUC
WATER
WASTEWATER
POWER

Statistics – Z-Score Regression



WATER
WASTEWATER
POWER

Statistics – PCA vs. Z-Regression

- With Z-score, still need enough years to get good estimate of mean and standard deviation
- PC does “grouping” implicitly, Z-score explicitly
- PC and Z-score often give similar forecast results, but coefficients differ
- Z-score assumes you have good representation of each main signal

‘nuff Stats – What’s it look like?

Microsoft Excel - VIPER 1.7.4_HH_v2.xls

	Type	Target	Start	End
	CSV File	HH_target.csv	Apr	Aug

	Type	Predictors	Start	End
1	CSV File	NorthMidp.csv	Apr	Apr
2	CSV File	NorthHighp.csv	Apr	Apr
3	CSV File	SouthMidp.csv	Apr	Apr
4	CSV File	SouthHighp.csv	Apr	Apr
5	CSV File	LowSnowp.csv	Apr	Apr
6	CSV File	lowsnow_focus_ln.csv	Apr	Apr
7	CSV File	PCPOctMarDecayedl.csv	Apr	Apr
8	CSV File	pcp_index.csv	Apr	Jul
9	CSV File	pcp_index_pinecrest.csv	Dec-1	Apr
10	CSV File	pcp_index_HHWP.csv	Dec-1	Mar
11	CSV File	PCPAprJulBoost.csv	Apr	Apr
12	CSV File	PCPAnnualdecay.csv	Apr	Apr
13	CSV File	PCPOctMarWted.csv	Apr	Apr
14	CSV File	PCPOctDecadw.csv	Apr	Apr
15	CSV File	PCPWYearWTED.csv	Apr	Apr
16	CSV File	PCPJanMarl.csv	Apr	Apr
17	CSV File	PCPOctDecBEST.csv	Apr	Apr
18	CSV File	PCPBestAprJul.csv	Apr	Apr

Global month changes: Instantaneous Apr F Accumulated Oct-1 F Jul L

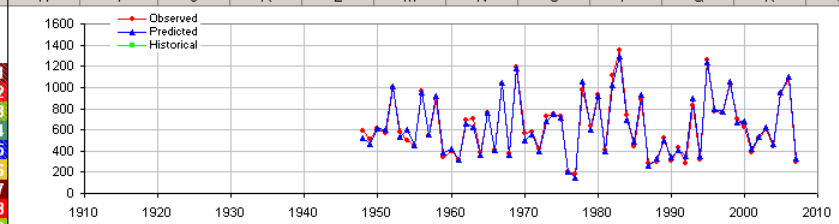
Station	1	2	3	4	5	6
Correl	0.927	0.942		0.958		
Years	60/60	60/60		60/60		
CurrZ	-0.123	-0.252		0.028		
PctNorm	95%	90%		101%		
Pred	603.46	568.63		642.99		

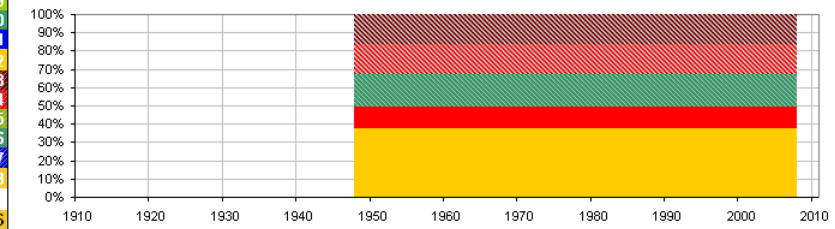
Station	7	8	9	10	11	12
Correl		0.541				0.970
Years		60/60				60/60
CurrZ		-1.365				-0.827
PctNorm		20%				71%
Pred		428.10				410.28

Station	13	14	15	16	17	18
Correl						
Years						
CurrZ						
PctNorm						
Pred						

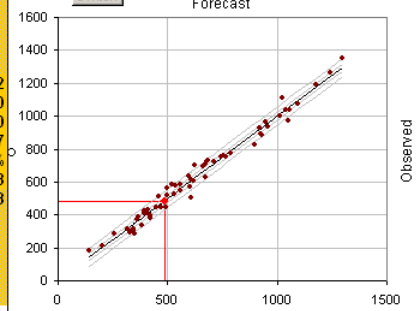
Group		SNOW	PRCP	SNOWI	SNOWDEN	RESINFLOW
Correl			0.944	0.956		
Years			60	60		
CurrZ			-0.954	-0.114		
Pred			360.39	604.38		

Group	SOI	All
Correl		0.990
Years		60
CurrZ		-0.529
Pred		485.587811

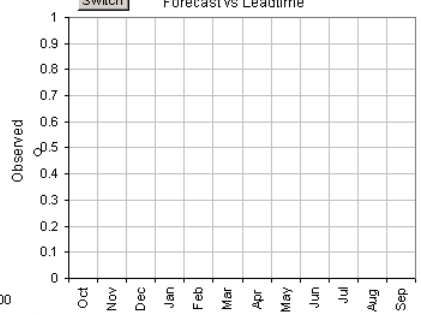




Switch Forecast



Switch Forecast vs Leadtime



Analysis Type	Principal Components	
Forecast	Volume	Pct
10	540.74	86%
30	508.02	81%
50	485.59	77%
70	463.16	73%
90	430.44	68%
	standard	jackknife
r2	0.980	0.978

Transformation	None
First Year Used	1948
Last Year Used	2008
Target Data Src	AWDB
Publication Date	April
Published?	☑

More Predictors... (none active)
Advanced Settings... (none active)
Helper Predictand... (not active)
Clear
Recalculate

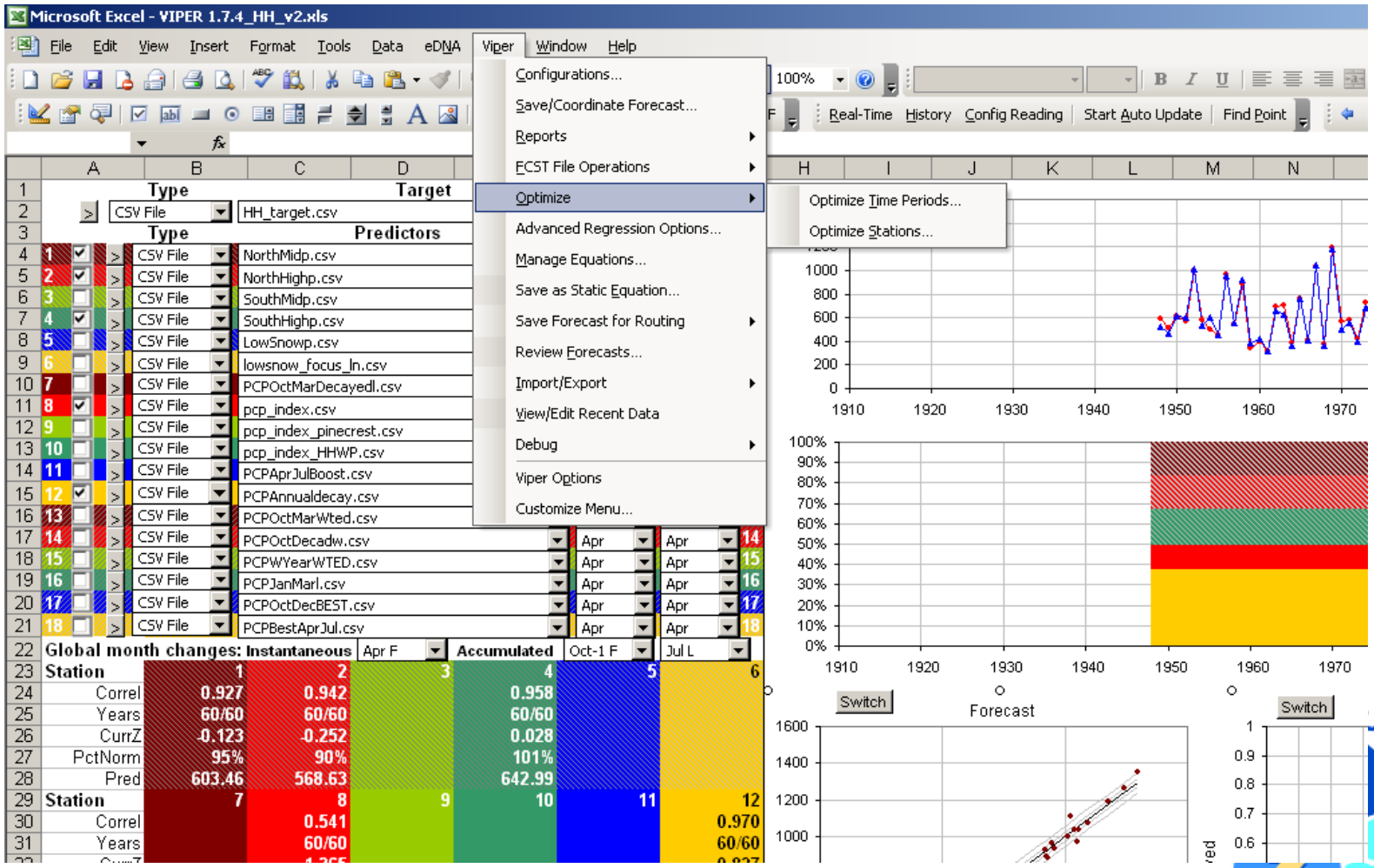
PCA
comp 2
% var 0.798

Main / Stations / Calculation / HistoricalFcsts / Forecast / SecondaryForecast / Configurations / Equations / FMT / Averages / Routing / ExternalData / DataViewer

Ready Calculate NUM

PUC WATER WASTEWATER POWER

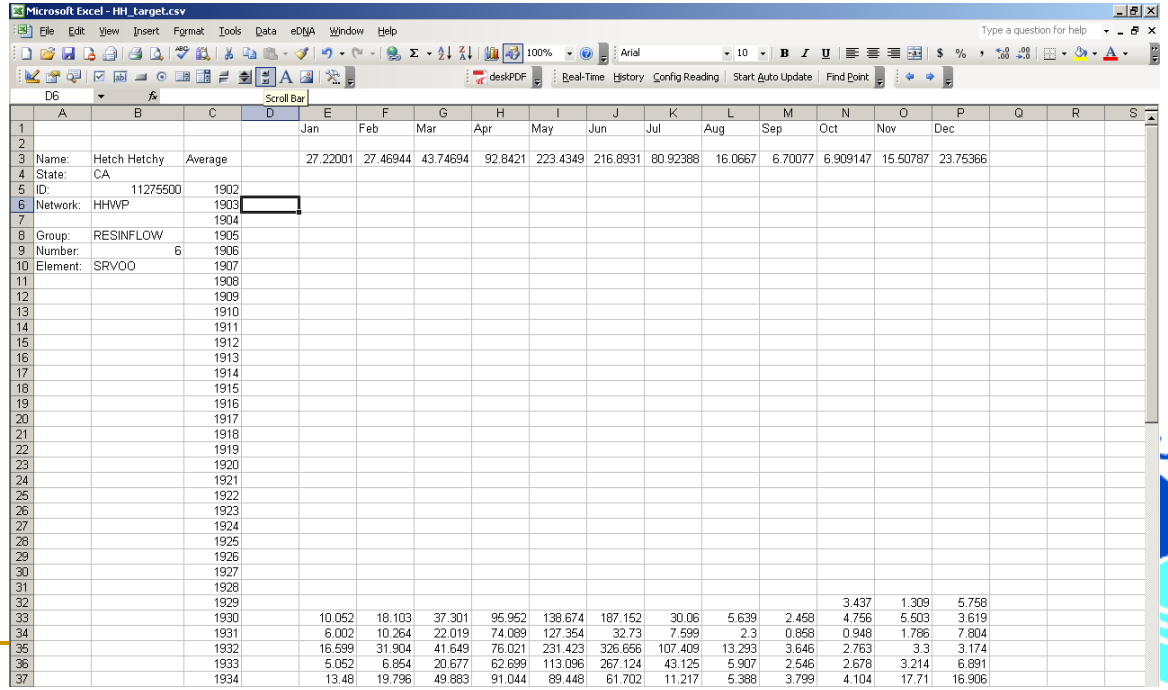
What's it look like?



WATER
WASTEWATER
POWER

Getting Started

- VIPER data input
  - Database access (NRCS corporate database)
  - USGS database
  - CSV files



Microsoft Excel - HH_target.csv

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1					Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2																
3	Name:	Hetch Hetchy	Average		27.22001	27.46944	43.74694	92.8421	223.4349	216.8931	80.92388	16.0667	6.70077	6.909147	15.50787	23.75366
4	State:	CA														
5	ID:	11275500	1902													
6	Network:	HHWP	1903													
7			1904													
8	Group:	RESINFLOW	1905													
9	Number:	6	1906													
10	Element:	SRVOO	1907													
11			1908													
12			1909													
13			1910													
14			1911													
15			1912													
16			1913													
17			1914													
18			1915													
19			1916													
20			1917													
21			1918													
22			1919													
23			1920													
24			1921													
25			1922													
26			1923													
27			1924													
28			1925													
29			1926													
30			1927													
31			1928													
32			1929											3.437	1.309	5.758
33			1930		10.052	18.103	37.301	95.952	138.674	187.152	30.06	5.639	2.458	4.756	5.503	3.619
34			1931		6.002	10.264	22.019	74.089	127.354	32.73	7.599	2.3	0.858	0.948	1.786	7.804
35			1932		16.599	31.904	41.649	76.021	231.423	326.656	107.409	13.293	3.646	2.763	3.3	3.174
36			1933		5.052	6.854	20.677	62.699	113.096	267.124	43.125	5.907	2.546	2.678	3.214	6.891
37			1934		13.48	19.796	49.883	91.044	89.448	61.702	11.217	5.388	3.799	4.104	17.71	16.906

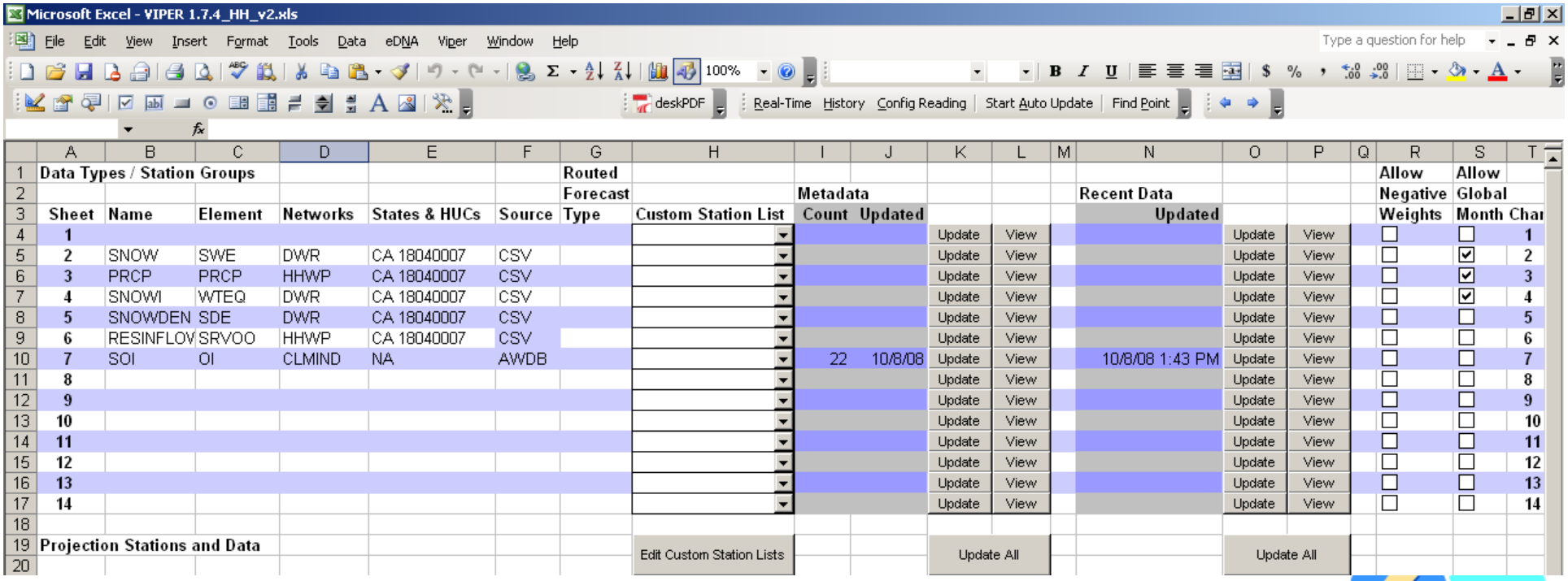
WATER
WASTEWATER
POWER

Getting started – USGS data

- Feature will retrieve monthly volumes for stream gaging stations.
- Search by state, preferably station ID to reduce retrieval time

Getting Started

- Define forecast point
- Set-up station groupings for Z-score



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Data Types / Station Groups						Routed											Allow	Allow	
2							Forecast		Metadata					Recent Data				Negative	Global	
3	Sheet	Name	Element	Networks	States & HUCs	Source	Type	Custom Station List	Count	Updated				Updated				Weights	Month	Char
4	1										Update	View			Update	View				1
5	2	SNOW	SWE	DWR	CA 18040007	CSV					Update	View			Update	View				2
6	3	PRCP	PRCP	HHWP	CA 18040007	CSV					Update	View			Update	View				3
7	4	SNOWI	WTEQ	DWR	CA 18040007	CSV					Update	View			Update	View				4
8	5	SNOWDEN	SDE	DWR	CA 18040007	CSV					Update	View			Update	View				5
9	6	RESINFLOV	SRVOO	HHWP	CA 18040007	CSV					Update	View			Update	View				6
10	7	SOI	OI	CLMIND	NA	AWDB			22	10/8/08	Update	View		10/8/08 1:43 PM	Update	View				7
11	8										Update	View			Update	View				8
12	9										Update	View			Update	View				9
13	10										Update	View			Update	View				10
14	11										Update	View			Update	View				11
15	12										Update	View			Update	View				12
16	13										Update	View			Update	View				13
17	14										Update	View			Update	View				14
18																				
19	Projection Stations and Data							Edit Custom Station Lists			Update All				Update All					
20																				

WATER
WASTEWATER
POWER

Z-score Groupings

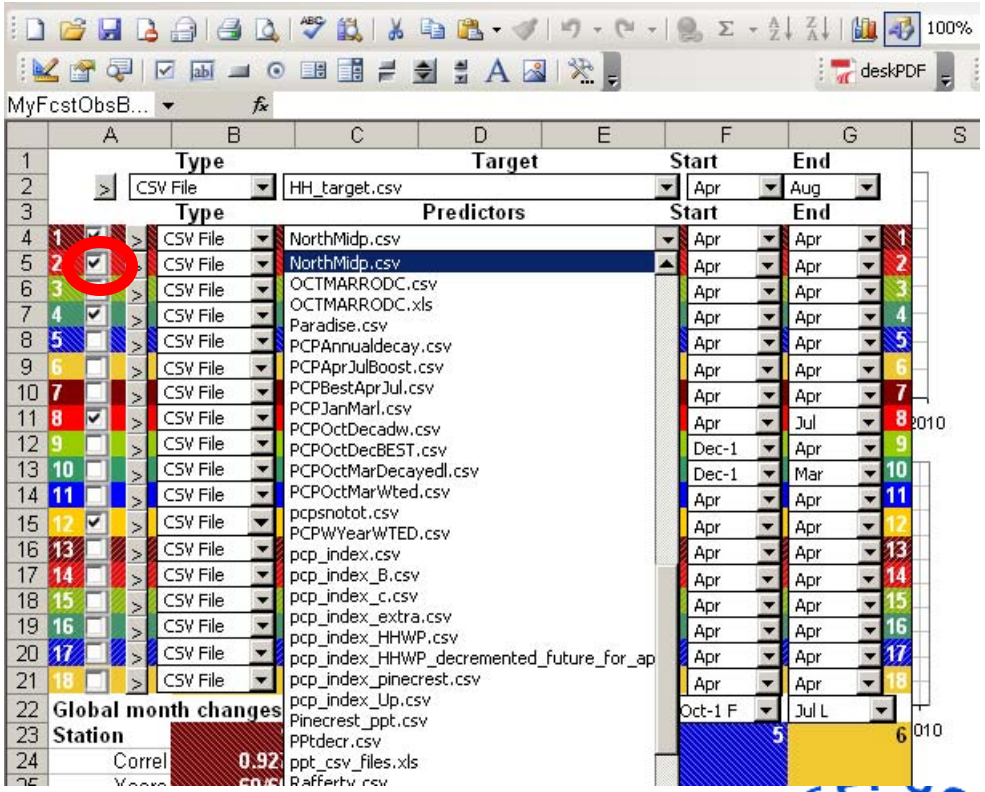
Microsoft Excel - VIPER 1.7.4_HH_v2.xls

			Station Name	Station ID	Network	State	Latitude	Longitude	Elevation	Acton ID	Shef ID	HUC	Country
Name:	PRCP		HHppt	HHPPT	HHWP	CA							
Element:	PRCP		Cherryppt	ChPPT	HHWP	CA							
Networks:	HHWP, DWR		Excppt	ExcPPT	HHWP	CA							
States:	CA		Grovelandppt	GVDPPT	HHWP	CA							
Source:	CSV		Pinecrestppt	PCPPT	HHWP	CA							
Count:	10		Sonorappt	SNPPT	HHWP	CA							
Last Updated:	10/1/08		Yosemiteppt	YOPPT	HHWP	CA							
Routed Type:			Moccppt	MoccPPT	HHWP	CA							
			Earlyppt	ElPPT	HHWP	CA							
			TuolumePPT	TUMPPT	HHWP	CA							

FPUC
WATER
WASTEWATER
POWER

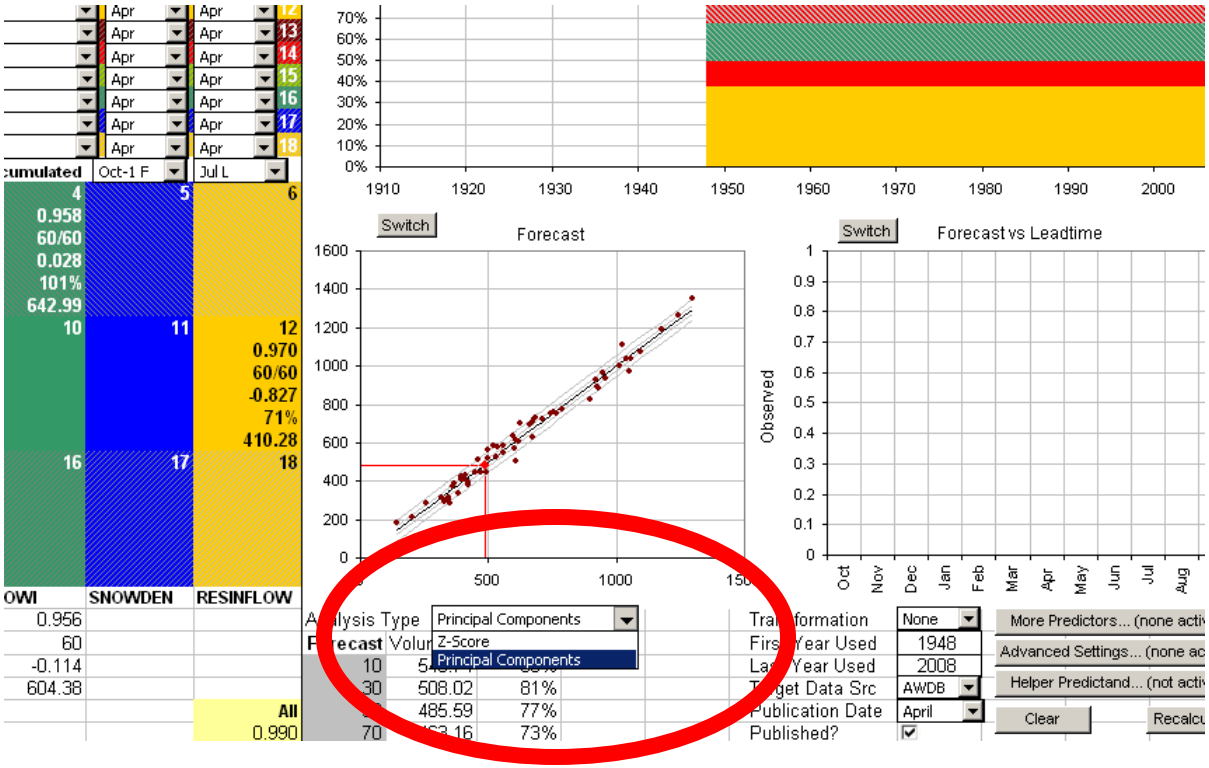
Load data

- From your directory select csv files
- Checkbox activates variable in equation



WATER
WASTEWATER
POWER

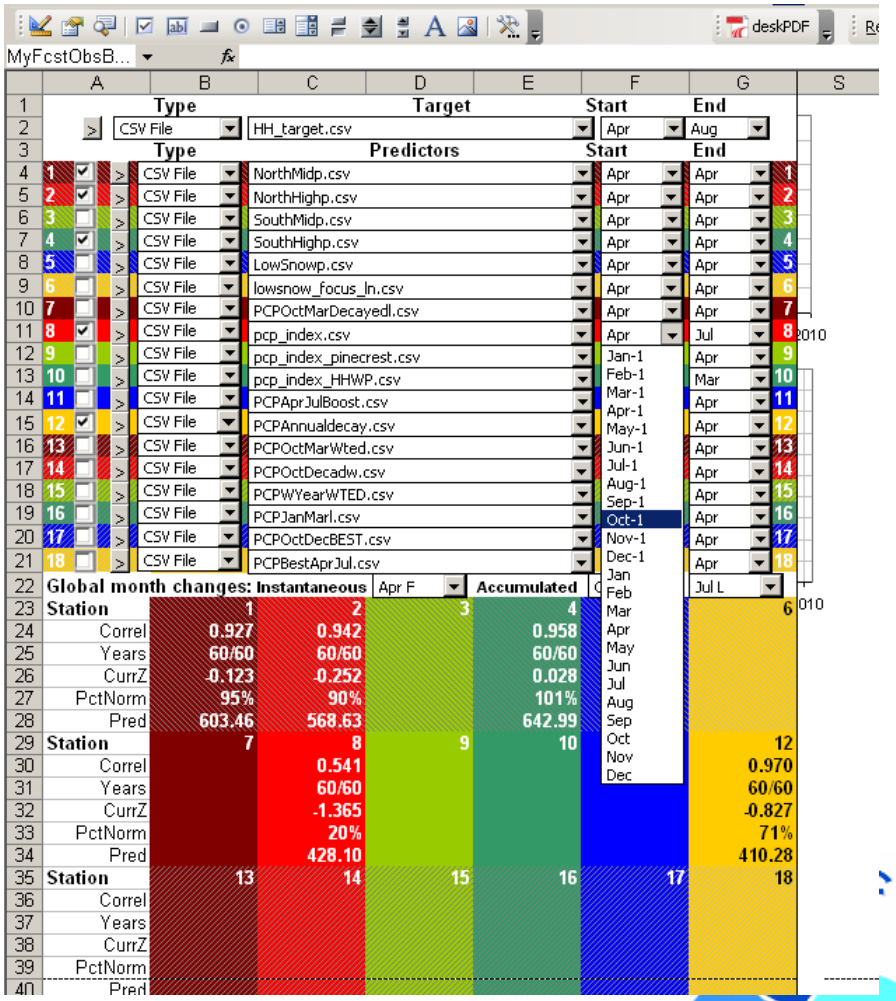
Choose Z-score or PCA



- Choosing PCA gives results for both
- Final Equation uses the analysis chosen here

Select Time Periods

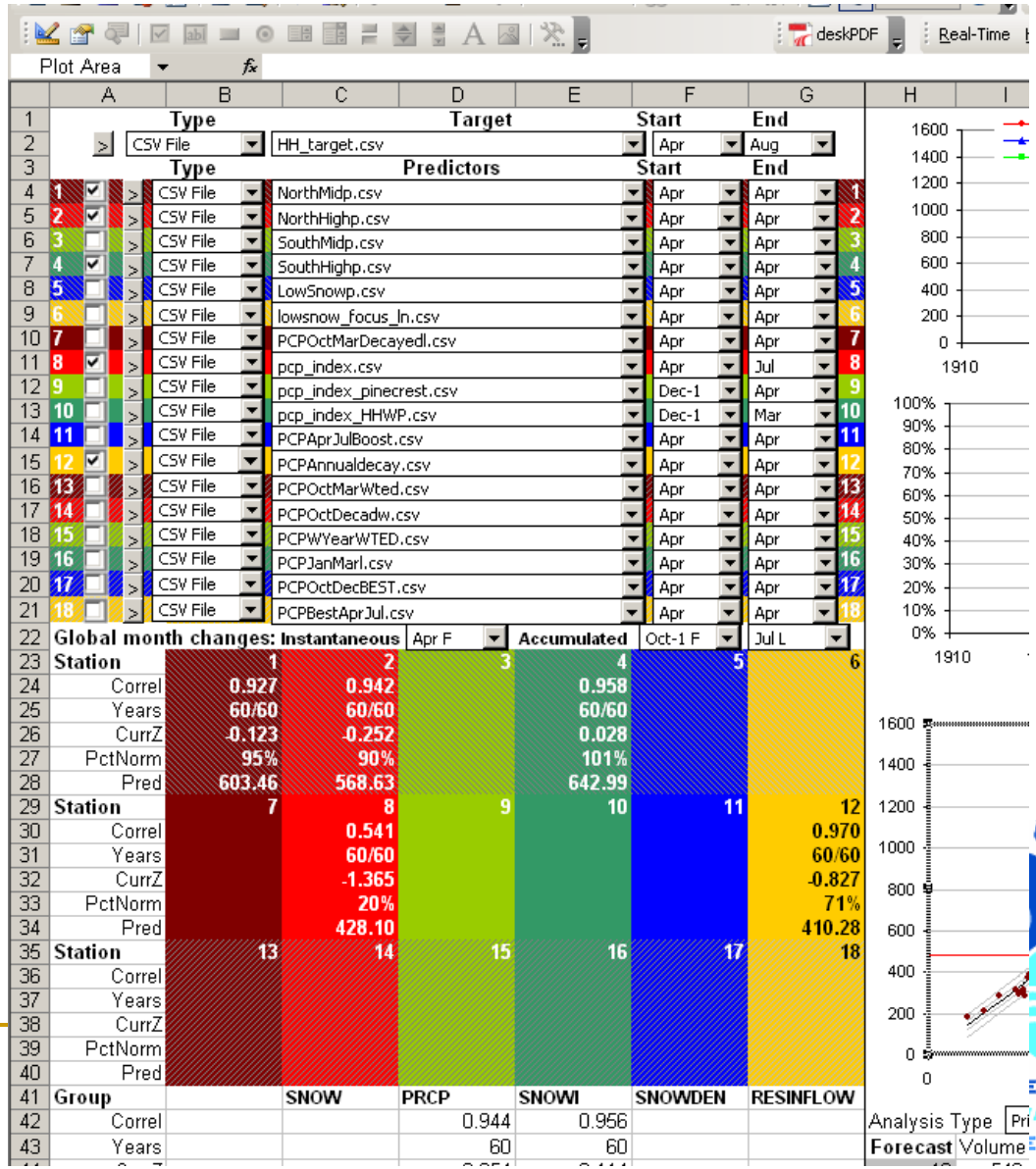
- Variables are defined as cumulative of instantaneous in their element type
- i.e. selecting a PRCP site will accumulate precip over the time period selected



WATER
WASTEWATER
POWER

Target and Predictor Variables

- 18 Predictors on Front Page
- 12 additional available



	Type	Target	Start	End
	CSV File	HH_target.csv	Apr	Aug

		Type	Predictors	Start	End	
1	☑	CSV File	NorthMidp.csv	Apr	Apr	1
2	☑	CSV File	NorthHighp.csv	Apr	Apr	2
3	☐	CSV File	SouthMidp.csv	Apr	Apr	3
4	☑	CSV File	SouthHighp.csv	Apr	Apr	4
5	☐	CSV File	LowSnowp.csv	Apr	Apr	5
6	☐	CSV File	lowsnow_focus_ln.csv	Apr	Apr	6
7	☐	CSV File	PCPOctMarDecayedl.csv	Apr	Apr	7
8	☑	CSV File	pcp_index.csv	Apr	Jul	8
9	☐	CSV File	pcp_index_pinecrest.csv	Dec-1	Apr	9
10	☐	CSV File	pcp_index_HHWP.csv	Dec-1	Mar	10
11	☐	CSV File	PCPAprJulBoost.csv	Apr	Apr	11
12	☑	CSV File	PCPAnnualdecay.csv	Apr	Apr	12
13	☐	CSV File	PCPOctMarWted.csv	Apr	Apr	13
14	☐	CSV File	PCPOctDecadw.csv	Apr	Apr	14
15	☐	CSV File	PCPWYearWTED.csv	Apr	Apr	15
16	☐	CSV File	PCPJanMarl.csv	Apr	Apr	16
17	☐	CSV File	PCPOctDecBEST.csv	Apr	Apr	17
18	☐	CSV File	PCPBestAprJul.csv	Apr	Apr	18

Global month changes: Instantaneous Apr F Accumulated Oct-1 F Jul L

Station	1	2	3	4	5	6
Correl	0.927	0.942		0.958		
Years	60/60	60/60		60/60		
CurrZ	-0.123	-0.252		0.028		
PctNorm	95%	90%		101%		
Pred	603.46	568.63		642.99		

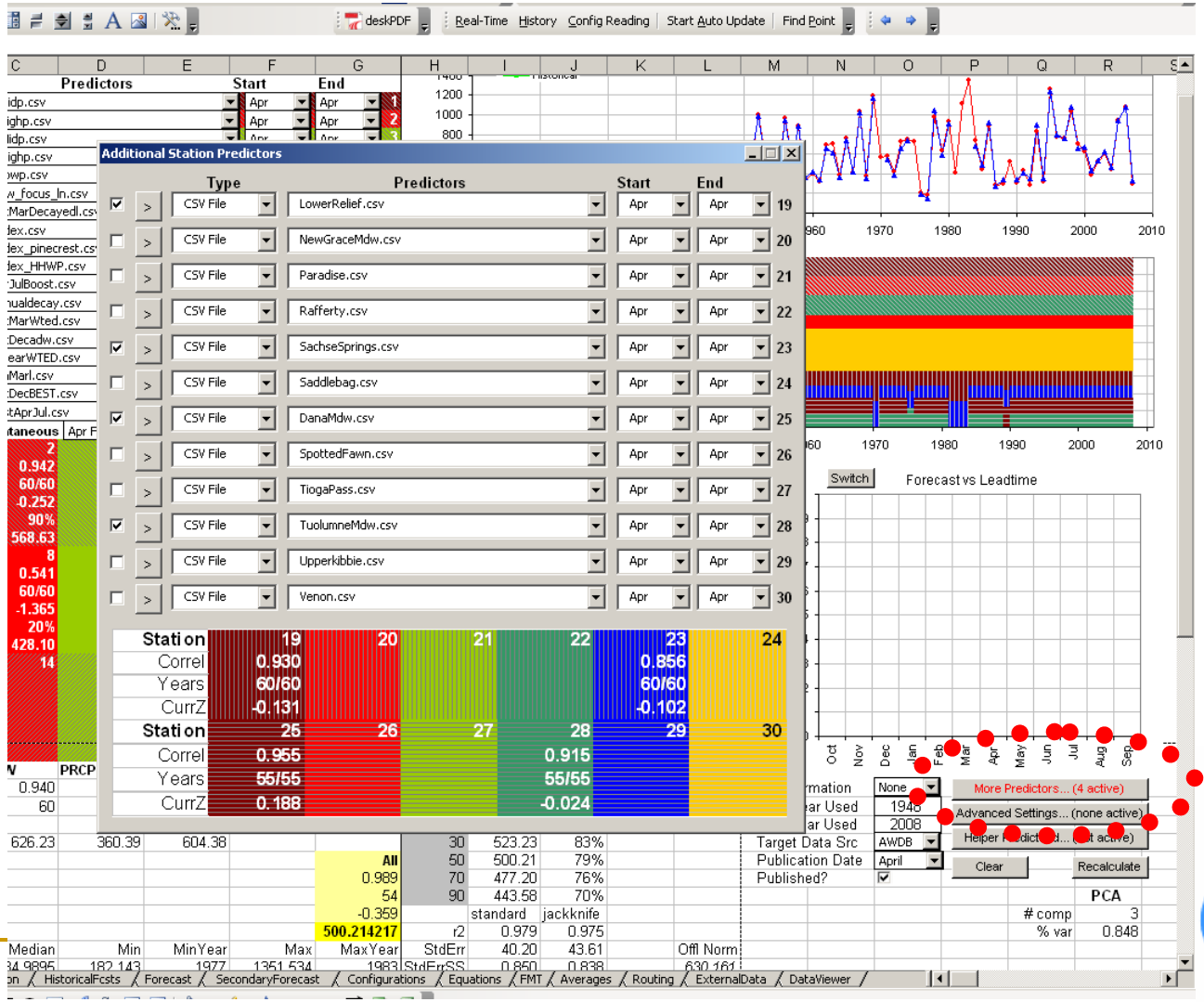
Station	7	8	9	10	11	12
Correl		0.541				0.970
Years		60/60				60/60
CurrZ		-1.365				-0.827
PctNorm		20%				71%
Pred		428.10				410.28

Station	13	14	15	16	17	18
Correl						
Years						
CurrZ						
PctNorm						
Pred						

Group		SNOW	PRCP	SNOWI	SNOWDEN	RESINFLOW
Correl			0.944	0.956		
Years			60	60		

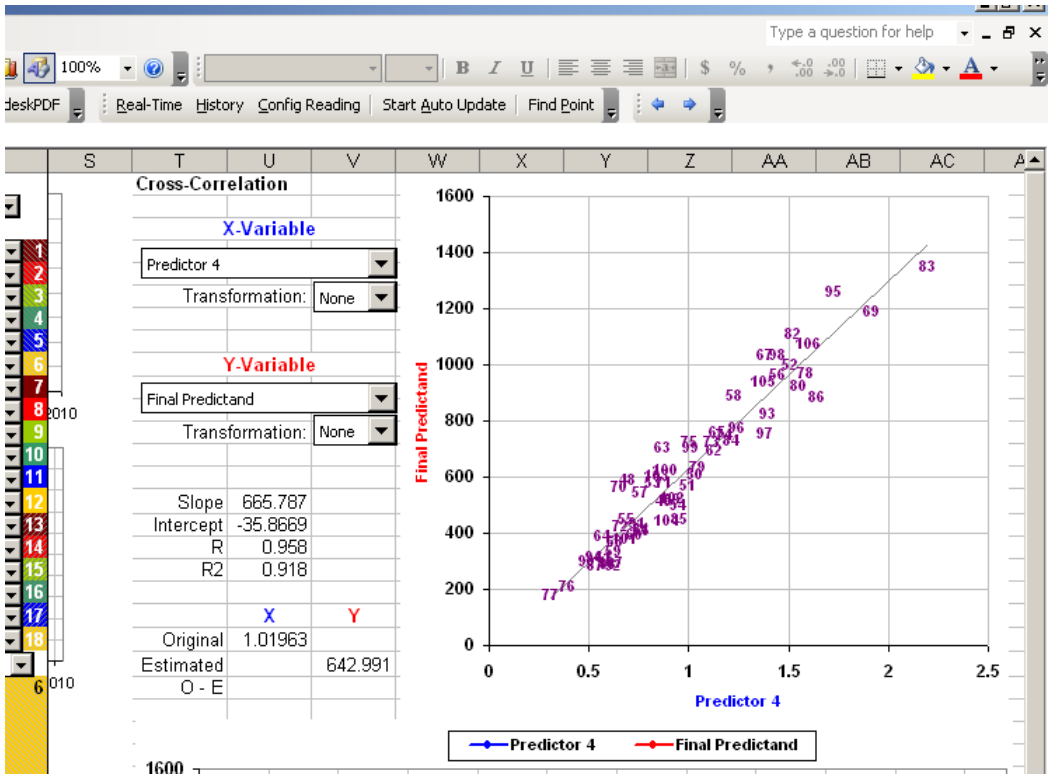
Analysis Type Pri
Forecast Volume

More Predictors



SFPUC
WATER
WASTEWATER
POWER

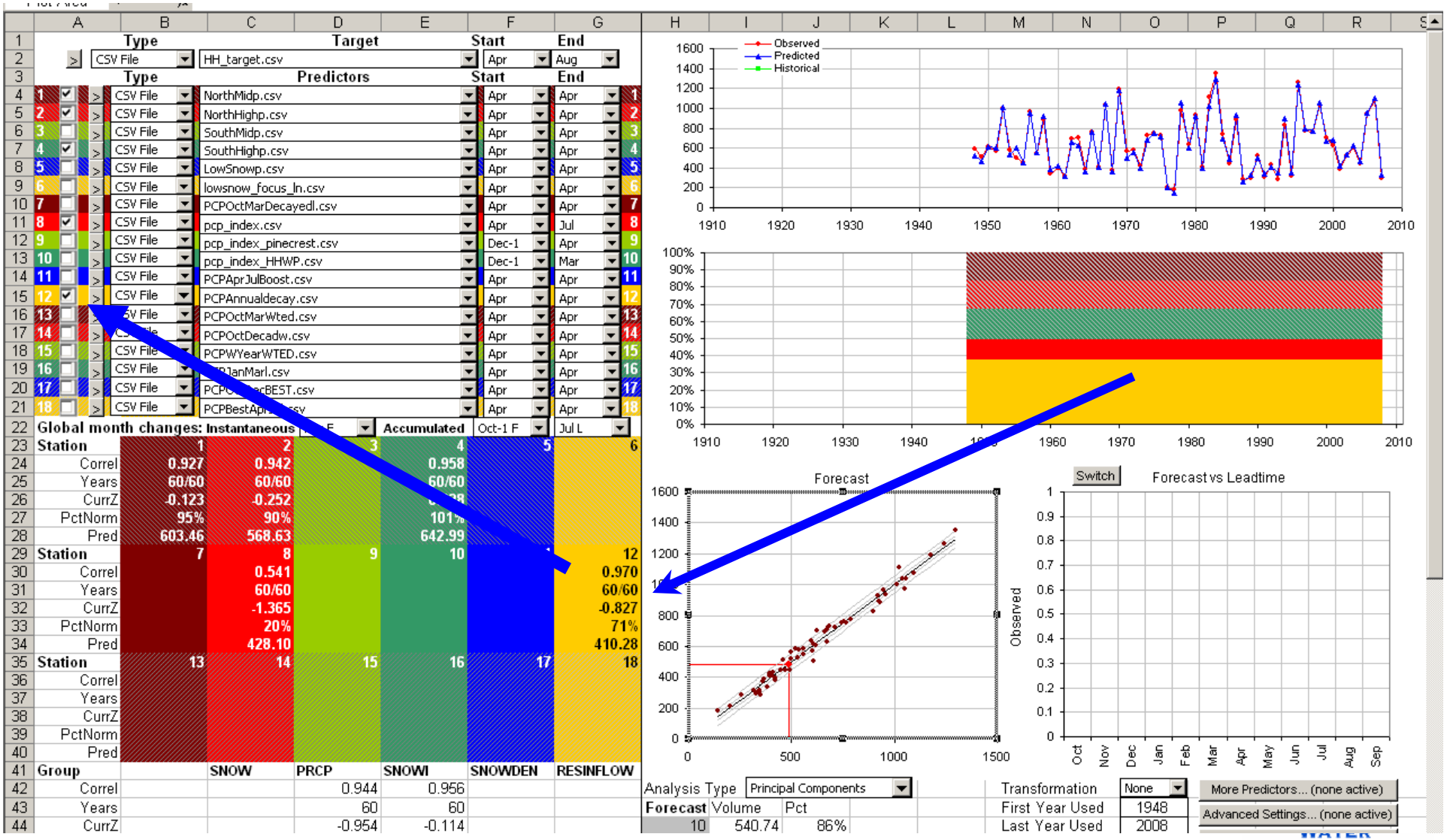
Predictor correlation



Are there outliers?

Is it linear or need to be transformed?

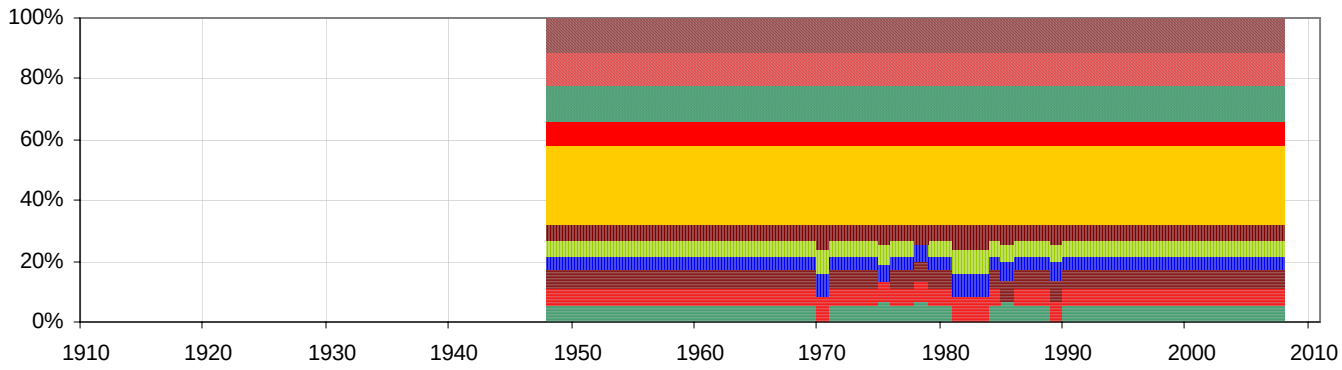
Predictor correlation



WATER
WASTEWATER
POWER

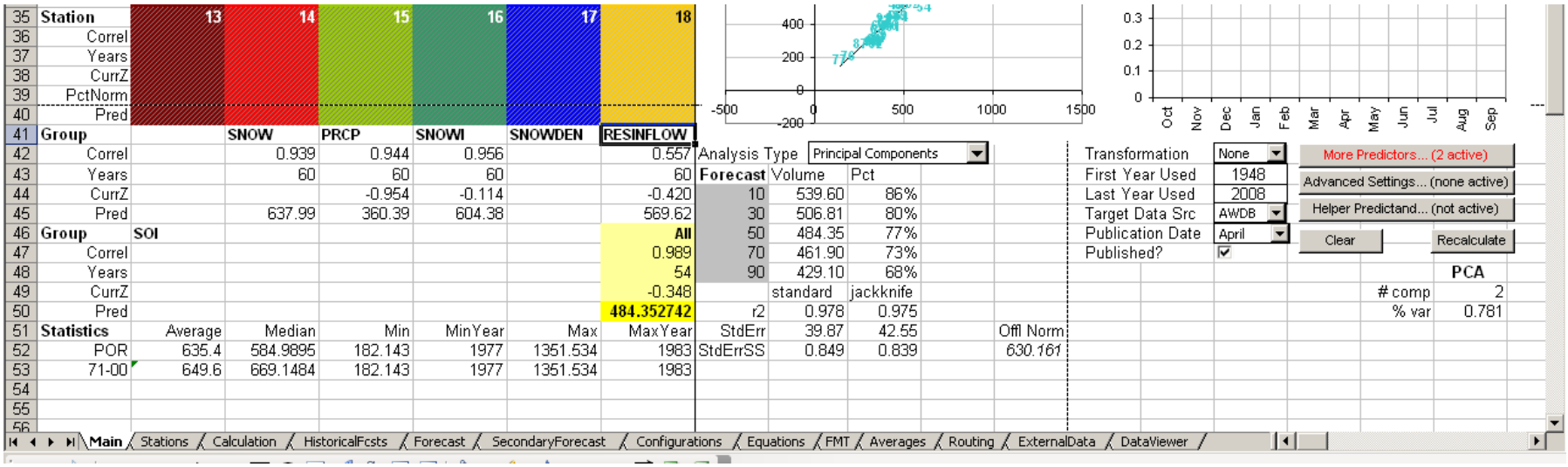
Predictor period of record

- The bar plot not only shows you the relevant correlation, but also data gaps.



- Here you can decide: Z-score or estimate missing data to use PCA

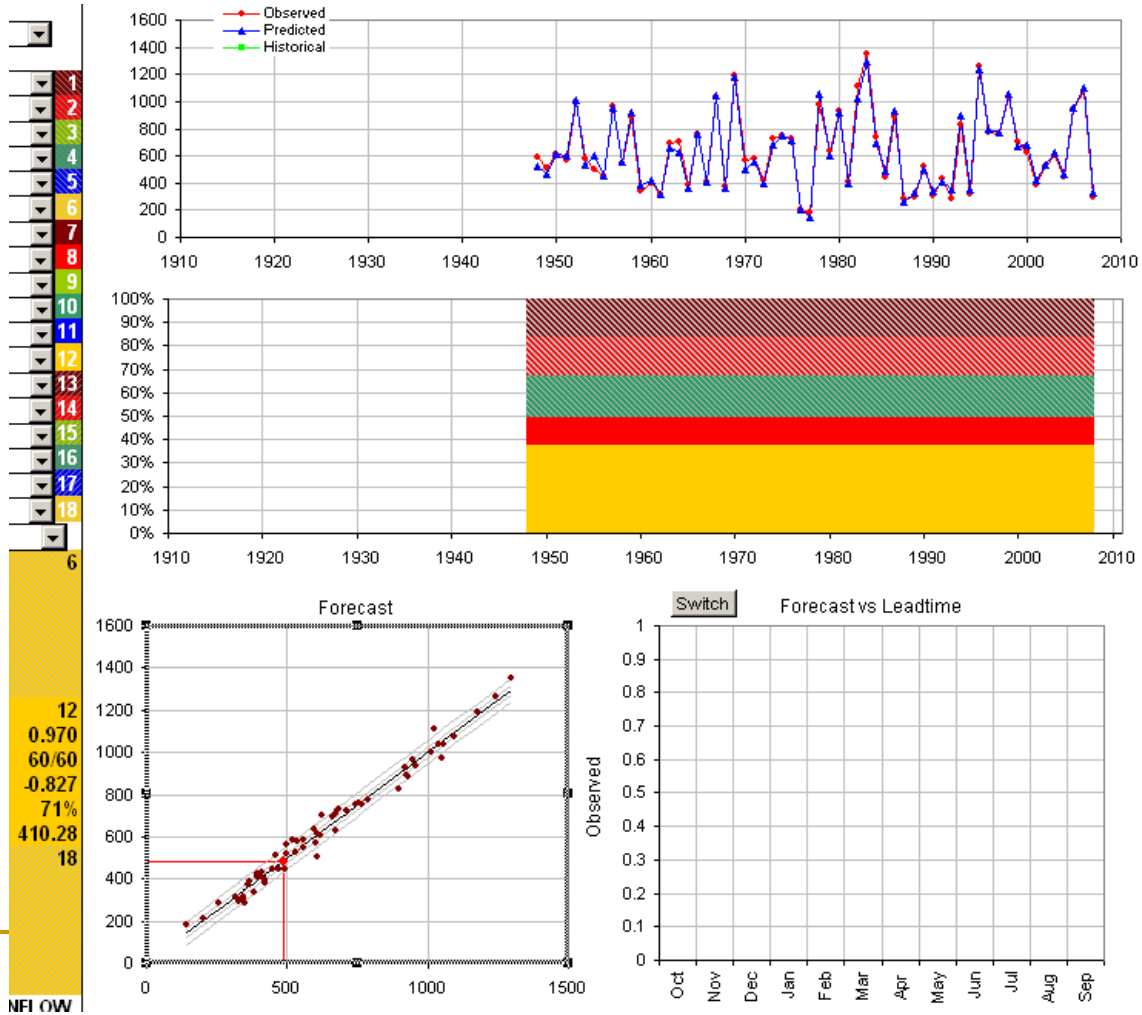
Z-score Groups



- Z-score group correlation and predictions

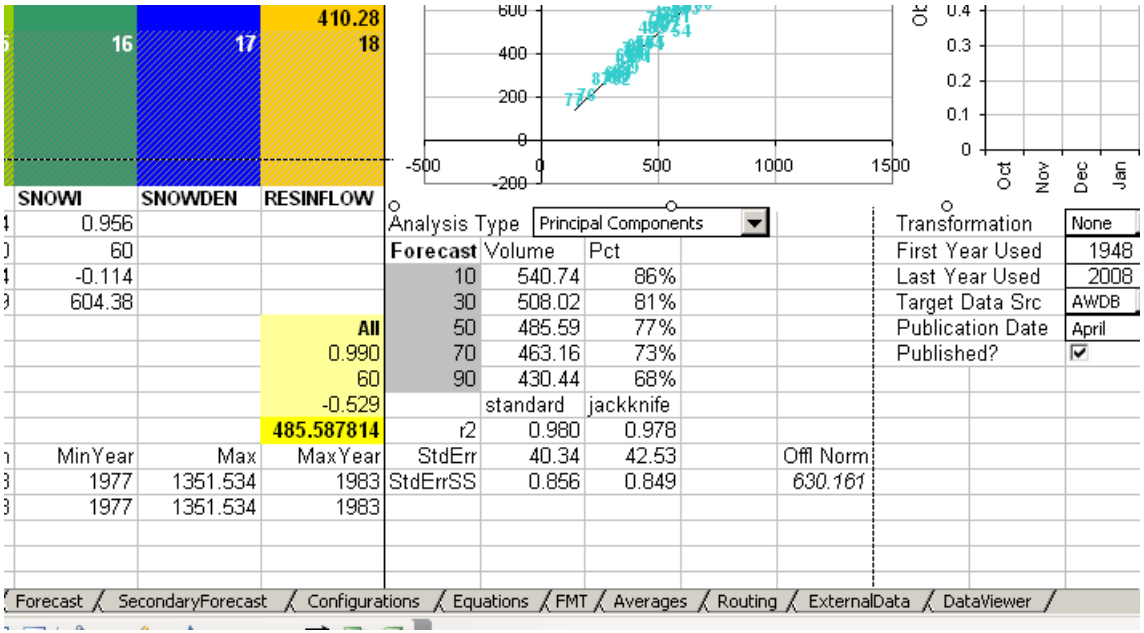
SFPUC
WATER
WASTEWATER
POWER

Model Performance



SFPUC
WATER
WASTEWATER
POWER

Forecast Volume

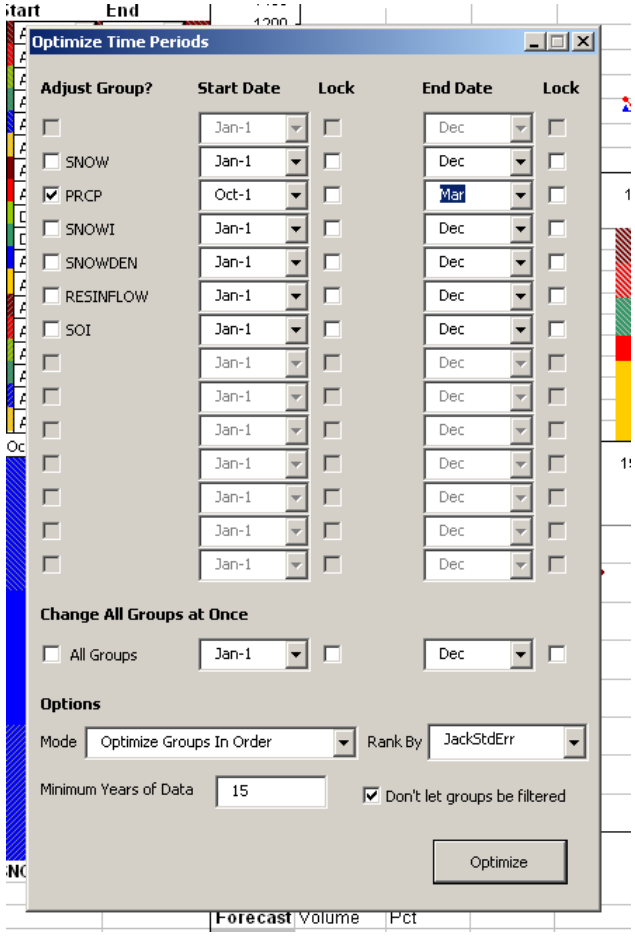


- Forecast volume, exceedence forecasts, and standard error for the forecast period
- ….How do I improve the forecast equation?

  <u>OPTIMIZATION!!</u>

SFPUC
WATER
WASTEWATER
POWER

Time Period - Optimization

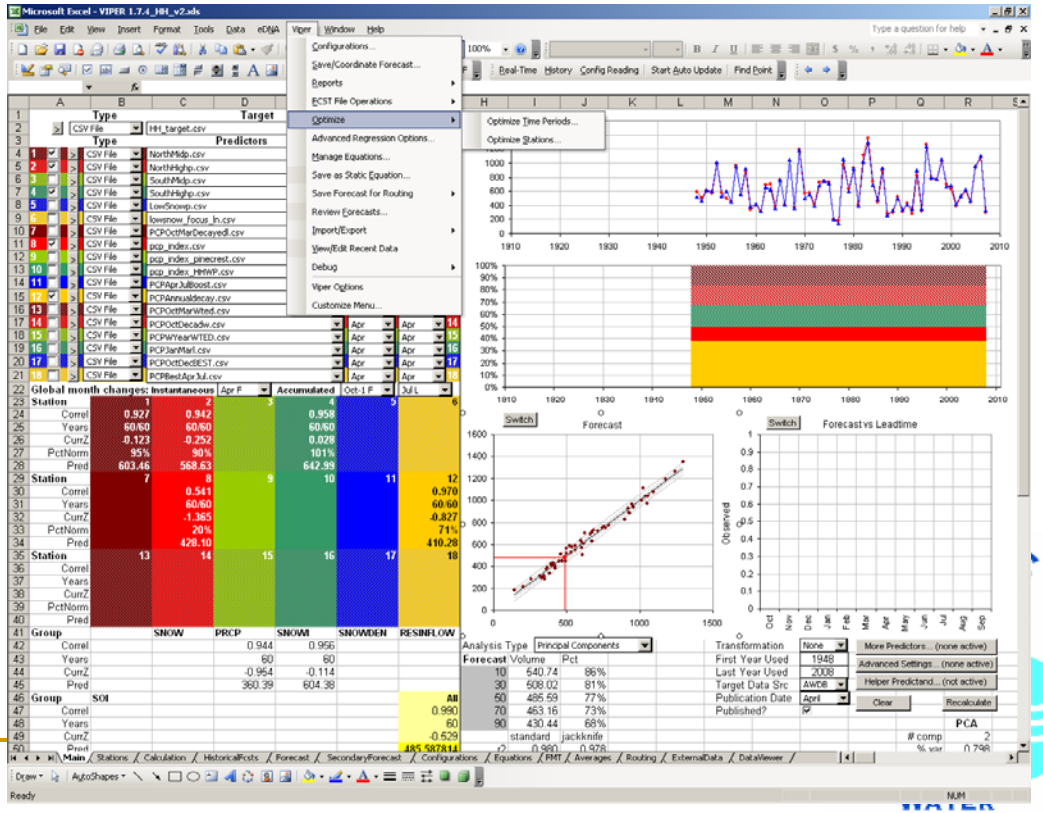


- Maybe you know Hetchy precip should be a predictor of Apr-Jul FNF
- But, what timeframe?
  - Oct-Mar?
  - Jan-Mar?
  - Last November??
- Time optimization varies stations within groups one at a time, looking for the optimal combination.
- It does not have a fancy search algorithm, it tries all combinations.

SFPUC
WATER
WASTEWATER
POWER

Station Optimization

- Great, now 30 options for variables – now what?
- *Answer* – Use station optimization function.



WATER
WASTEWATER
POWER

Microsoft Excel - VIPER 1.7.4_HH_v2.xls

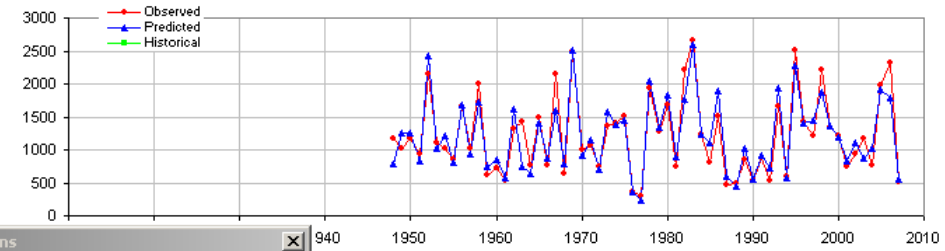
File Edit View Insert Format Tools Data eDNA Viper Window Help

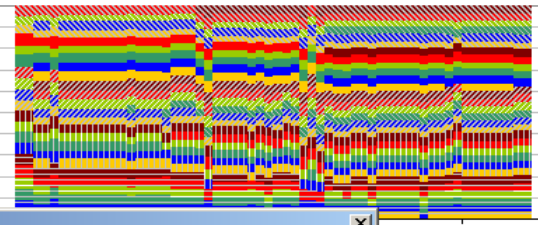
Type a question for help

deskPDF | Real-Time History Config Reading | Start Auto Update | Find Point

	Type	Target	Start	End
>	CSV File	LaG_Target.csv	Apr	Jul

		Type	Predictors	Start	End	
1	>	CSV File	VirginiaLakes.csv	Apr	Apr	1
2	>	CSV File	LowerRelief.csv	Apr	Apr	2
3	>	CSV File	BellMdw.csv	Apr	Apr	3
4	>	CSV File	Gianelli.csv	Apr	Apr	4
5	>	CSV File	GemPass.csv	Apr	Apr	5
6	>	CSV File	GinFlat.csv	Apr	Apr	6
7	>	CSV File	Gianelli.csv	Apr	Apr	7
8	>	CSV File	SonoraPass.csv	Apr		
9	>	CSV File	Beehive.csv	Apr		
10	>	CSV File	BondPass.csv	Apr		
11	>	CSV File	ElleryLake.csv	Apr		
12	>	CSV File	GemLake.csv	Apr		
13	>	CSV File	GemPass.csv	Apr		
14	>	CSV File	HorseMdw.csv	Apr		
15	>	CSV File	Huckleberry.csv	Apr		
16	>	CSV File	Kerrick.csv	Apr	Apr	16
17	>	CSV File	LakeTenaya.csv	Apr	Apr	17
18	>	CSV File	LowerKibbie.csv	Apr	Apr	18





Optimize Stations

Evaluating equation 4747...

Optimize Stations

Force These Stations to be Included:

1	SNOW: 377, CA, Virgina Lakes Ridge Snow Course	11 SNOW: 286, CA, Ellery Lakes Snow Course	21 SNOW: 167, CA, Paradise Meadow Snow Course
2	SNOW: 138, CA, Lower Relief Snow Course	12 SNOW: 282, CA, Gem Lake Snow Course	22 SNOW: 158, CA, Rafferty Meadow Snow Course
3	SNOW: 172, CA, Bell Meadow Snow Course	13 SNOW: 281, CA, Gem Pass Snow Course	23 SNOW: 165, CA, Sachse Springs Snow Course
4	SNOW: 427, CA, Gianelli Snow Course	14 SNOW: 162, CA, Horse Meadow Snow Course	24 SNOW: 287, CA, Saddlebag Lakes Snow Course
5	SNOW: 281, CA, Gem Pass Snow Course	15 SNOW: 166, CA, Huckleberry Lake Snow Course	25 SNOW: 157, CA, Dana Meadow Snow Course
6	SNOW: 179, CA, GinFlatSnowCourse	16 SNOW: 348, CA, Kerrick Corral Snow Course	26 SNOW: 164, CA, Spotted Fawn Snow Course
7	SNOW: 427, CA, Gianelli Snow Course	17 SNOW: 178, CA, Lake Tenaya Snow Course	27 SNOW: 181, CA, Tioga Pass Snow Course
8	SNOW: 152, CA, Sonora Pass Snow Course	18 SNOW: 173, CA, Lower Kibbie Ridge Snow Course	28 SNOW: 161, CA, Tuolumne Meadow Snow Course
9	SNOW: 171, CA, Beehive Meadow Snow Course	19 SNOW: 138, CA, Lower Relief Snow Course	29 SNOW: 168, CA, Upper Kibbie Ridge Snow Course
10	SNOW: 159, CA, Bond Pass Snow Course	20 SNOW: 368, CA, New Grace Meadow Snow Course	30 SNOW: 169, CA, Vernon Lake Snow Course

Rank By: JackStdErr | Components to Retain: Automatic | T-statistic: 1.6 | Optimize

Global month changes:		
Station		1
	Correl	0.939
	Years	38/38
	CurrZ	0.258
	PctNorm	112%
	Pred	1377.28
Station		7
	Correl	0.921
	Years	24/24
	CurrZ	-0.384
	PctNorm	85%
	Pred	946.46
Station		13
	Correl	0.932
	Years	56/56
	CurrZ	0.231
	PctNorm	110%
	Pred	1334.54
Group		
	Correl	
	Years	
	CurrZ	
	Pred	1178.93
Group		SOI
	Correl	
	Years	
	CurrZ	
	Pred	

			All
			0.931
			60
			-0.054
			1178.93321

30	1295.15	107%	
50	1178.93	97%	
70	1062.72	88%	
90	893.23	74%	
	standard	jackknife	
r2	0.866		

Target Data Src: AWDB
Publication Date: April
Published? ☑

Clear | Recalculate

Helper Predictand... (not active)

Settings... (none active)

edictors... (12 active)

PCA

comp

% var

Main / Stations / Calculation / HistoricalFcsts / Forecast / SecondaryForecast / Configurations / Equations / FMT / Averages / Routing / ExternalData / DataViewer

Draw | AutoShapes

Ready | Calculate | NUM

Model Evaluation - Equations

```
VARIABLES:

    Y1   11275500,APR-AUG,SRVOO,HHWP,CA,AWDB,Hetch Hetchy
    H1

    X1   NMSWEp,APR,WTEQ,DWR,CA,CSV,NorthMidp
    X2   NHSWEp,APR,WTEQ,DWR,CA,CSV,NorthHighp
    X3   SMSWEp,APR,WTEQ,DWR,CA,CSV,SouthMidp
    X4   SHSWEp,APR,WTEQ,DWR,CA,CSV,SouthHighp
    X7   2,APR,PRCP,HHWP,CA,CSV,PCPOCTMARDecayed
    X8   2,APR-JUL,PRCP,HHWP,CA,CSV,PCPindex
   X11   2,APR,PRCP,HHWP,CA,CSV,PCPAPRILjULBOOST
   X12   2,APR,PRCP,HHWP,CA,CSV,PCPAnnualDecay
   X13   2,APR,PRCP,HHWP,CA,CSV,PCPOctMarWTD
   X14   2,APR,PRCP,HHWP,CA,CSV,PCPOctDecadw
   X15   2,APR,PRCP,HHWP,CA,CSV,PCPWYearEted
   X16   2,APR,PRCP,HHWP,CA,CSV,PCPJanMar
   X17   2,APR,PRCP,HHWP,CA,CSV,PCPOctDecBEST
   X18   2,APR,PRCP,HHWP,CA,CSV,PCPBESTaprJul

EQUATION SUMMARY:
```

RANK	VARIABLES 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	JACKKNIFE STANDARD ERROR	JACK. CORR. COEF.	NUM. OBS. USED	NUM. PC'S
1	X	X		X				X			X																				42.534	0.989	60	2
2	X	X		X				X			X			X			X														42.806	0.989	60	2
3	X	X		X				X				X																			42.877	0.989	60	2
4	X	X	X	X				X				X		X																	43.007	0.989	60	2
5	X	X	X	X				X			X			X																	43.097	0.989	60	2
6	X	X	X	X				X				X					X														43.134	0.989	60	2
7	X	X		X				X				X						X													43.212	0.988	60	1
8	X	X		X				X			X						X														43.215	0.989	60	3
9	X	X		X				X			X			X																	43.236	0.989	60	3
10		X	X	X				X				X						X													43.348	0.988	60	1
11	X	X		X				X				X		X			X														43.348	0.989	60	3
12	X	X		X				X				X					X														43.374	0.989	60	3
13		X	X	X				X			X																				43.385	0.988	60	2
14	X	X	X	X				X			X																				43.399	0.988	60	2
15		X		X				X			X																				43.403	0.988	60	1
16	X	X		X				X				X		X																	43.410	0.989	60	3
17	X	X	X	X				X		X							X														43.422	0.988	60	2
18	X	X	X	X				X			X	X	X																		43.440	0.988	60	2
19		X	X	X				X				X																			43.445	0.988	60	2
20	X	X	X	X				X		X				X																	43.457	0.988	60	2
21		X	X					X			X																				43.486	0.988	60	1
22	X	X	X	X				X			X			X			X														43.524	0.989	60	3
23	X	X	X	X				X				X	X				X														43.568	0.989	60	3
24	X	X	X	X				X			X	X					X														43.584	0.988	60	2
25	X	X	X	X				X			X						X														43.604	0.988	60	2
26	X	X	X	X				X				X																			43.693	0.988	60	2
27	X	X	X	X				X			X							X													43.844	0.988	60	1
28	X	X		X				X		X																					43.871	0.988	60	2
29	X	X		X				X			X	X																			43.952	0.988	60	3
30	X	X	X	X				X		X	X			X																	43.964	0.988	60	2

```
DETAILED OUTPUT
```

Model Evaluation – Variables used

```
VARIABLES:

    Y1   11275500,APR-AUG,SRVOO,HHWP,CA,AWDB,Hetch Hetchy
    H1

    X1   NMSWEp,APR,WTEQ,DWR,CA,CSV,NorthMidp
    X2   NHSWEp,APR,WTEQ,DWR,CA,CSV,NorthHighp
    X3   SMSWEp,APR,WTEQ,DWR,CA,CSV,SouthMidp
    X4   SHSWEp,APR,WTEQ,DWR,CA,CSV,SouthHighp
    X7   2,APR,PRCP,HHWP,CA,CSV,PCPOCTMARDecayed
    X8   2,APR-JUL,PRCP,HHWP,CA,CSV,PCPindex
   X11   2,APR,PRCP,HHWP,CA,CSV,PCPAPRILjULBOOST
   X12   2,APR,PRCP,HHWP,CA,CSV,PCPAnnualDecay
   X13   2,APR,PRCP,HHWP,CA,CSV,PCPOctMarWTD
   X14   2,APR,PRCP,HHWP,CA,CSV,PCPOctDecadw
   X15   2,APR,PRCP,HHWP,CA,CSV,PCPWYearEted
   X16   2,APR,PRCP,HHWP,CA,CSV,PCPJanMar
   X17   2,APR,PRCP,HHWP,CA,CSV,PCPOctDecBEST
   X18   2,APR,PRCP,HHWP,CA,CSV,PCPBESTaprJul

EQUATION SUMMARY:
```

Recognize useful variables

…and others not so much

Choose what stays and what goes using your "hydrologic expertise"

RANK	VARIABLES 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	JACKKNIFE STANDARD ERROR	JACK. CORR. COEF.	NUM. OBS. USED	NUM. PC'S
1	X	X		X				X			X																				42.534	0.989	60	2
2	X	X		X				X			X		X				X														42.806	0.989	60	2
3	X	X		X				X				X																			42.877	0.989	60	2
4	X	X	X	X				X				X	X																		43.007	0.989	60	2
5	X	X	X	X				X			X		X																		43.097	0.989	60	2
6	X	X	X	X				X				X					X														43.134	0.989	60	2
7	X	X		X				X				X						X													43.212	0.988	60	1
8	X	X		X				X			X						X														43.215	0.989	60	3
9	X	X		X				X			X		X																		43.236	0.989	60	3
10		X	X	X				X				X						X													43.348	0.988	60	1
11	X	X		X				X				X	X				X														43.348	0.989	60	3
12	X	X		X				X				X					X														43.374	0.989	60	3
13		X	X	X				X			X																				43.385	0.988	60	2
14	X	X	X	X				X			X																				43.399	0.988	60	2
15		X		X				X			X																				43.403	0.988	60	1
16	X	X		X				X				X	X																		43.410	0.989	60	3
17	X	X	X	X				X		X							X														43.422	0.988	60	2
18	X	X	X	X				X			X	X	X																		43.440	0.988	60	2
19		X	X	X				X				X																			43.445	0.988	60	2
20	X	X	X	X				X		X			X																		43.457	0.988	60	2
21		X	X					X			X																				43.486	0.988	60	1
22	X	X	X	X				X			X		X				X														43.524	0.989	60	3
23	X	X	X	X				X				X	X				X														43.568	0.989	60	3
24	X	X	X	X				X			X	X					X														43.584	0.988	60	2
25	X	X	X	X				X			X						X														43.604	0.988	60	2
26	X	X	X	X				X				X																			43.693	0.988	60	2
27	X	X	X	X				X			X							X													43.844	0.988	60	1
28	X	X		X				X		X																					43.871	0.988	60	2
29	X	X		X				X				X	X																		43.952	0.988	60	3
30	X	X	X	X				X		X	X																				43.964	0.988	60	2

```
DETAILED OUTPUT
```

Detailed Results

```
DETAILED OUTPUT

RANKED REGRESSION EQUATIONS:

*****
 RANK     1

     coef   var
     ----   ----
 157.27767  X1    NMSWEp,APR,WTEQ,DWR,CA,CSV,NorthMidp
 158.80452  X2    NHSWEp,APR,WTEQ,DWR,CA,CSV,NorthHighp
 161.21853  X4    SHSWEp,APR,WTEQ,DWR,CA,CSV,SouthHighp
   5.87691  X8   2,APR-JUL,PRCP,HHWP,CA,CSV,PCPindex
   2.45318  X12  2,APR,PRCP,HHWP,CA,CSV,PCPAnnualDecay
-127.73178  Constant

      #obs=          60
      #pc =           2
        Jr =       0.989
   JStdErr =      42.534
JStdErrSS =       0.849
         r =       0.990
    StdErr =      40.338
 StdErrSS =       0.856
```

# of PC's used

Model standard error and Jackknife standard error

YEAR	OBSERVED	JCK REG COMPUTED	JCK REG ERROR	STD REG COMPUTED	STD REG ERROR
1948	587.56	513.75	-73.81	521.74	-65.82
1949	514.90	457.93	-56.97	460.11	-54.78
1950	611.98	608.37	-3.61	608.48	-3.50
1951	568.27	599.83	31.56	599.73	31.47
1952	1000.41	1014.02	13.60	1011.98	11.57
1953	580.43	534.84	-45.59	536.09	-44.35
1954	500.90	607.37	106.48	605.55	104.65
1955	448.87	448.56	-0.31	448.57	-0.30
1956	963.04	947.98	-15.06	948.52	-14.52
1957	546.39	558.75	12.36	558.39	12.01
1958	889.17	925.40	36.23	923.60	34.43
1959	335.50	382.20	46.70	380.67	45.16
1960	393.95	422.94	29.00	422.18	28.24
1961	313.23	316.75	3.51	316.65	3.42
1962	694.73	661.37	-33.37	662.99	-31.74
1963	702.89	614.76	-88.13	625.59	-77.30
1964	388.94	364.78	-24.15	365.81	-23.13
1965	761.30	754.83	-6.47	754.92	-6.38
1966	411.34	406.84	-4.50	407.01	-4.33
1967	1035.72	1040.05	4.33	1039.29	3.57
1968	370.01	363.25	-6.76	363.55	-6.45
1969	1188.20	1173.64	-14.56	1174.90	-13.30
1970	563.59	494.45	-69.14	496.08	-67.51
1971	582.42	558.74	-23.67	559.35	-23.07
1972	424.78	390.63	-34.14	391.73	-33.05
1973	724.16	677.15	-47.01	679.82	-44.34
1974	751.10	745.03	-6.07	745.12	-5.98
1975	725.56	712.02	-13.54	712.24	-13.31
1976	208.66	201.51	-7.15	201.88	-6.78
1977	182.14	140.00	-42.14	143.39	-38.75
1978	972.00	1057.58	85.57	1052.75	80.75
1979	633.19	597.91	-35.27	599.11	-34.08
1980	926.40	918.14	-8.26	918.37	-8.03
1981	412.74	393.80	-18.94	394.41	-18.32
1982	1111.05	1020.00	-91.05	1024.32	-86.72
1983	1351.53	1290.34	-61.19	1295.63	-55.90
1984	732.44	685.40	-47.04	686.68	-45.76
1985	448.16	495.43	47.27	493.28	45.12

Calculated results for Each year in your dataset

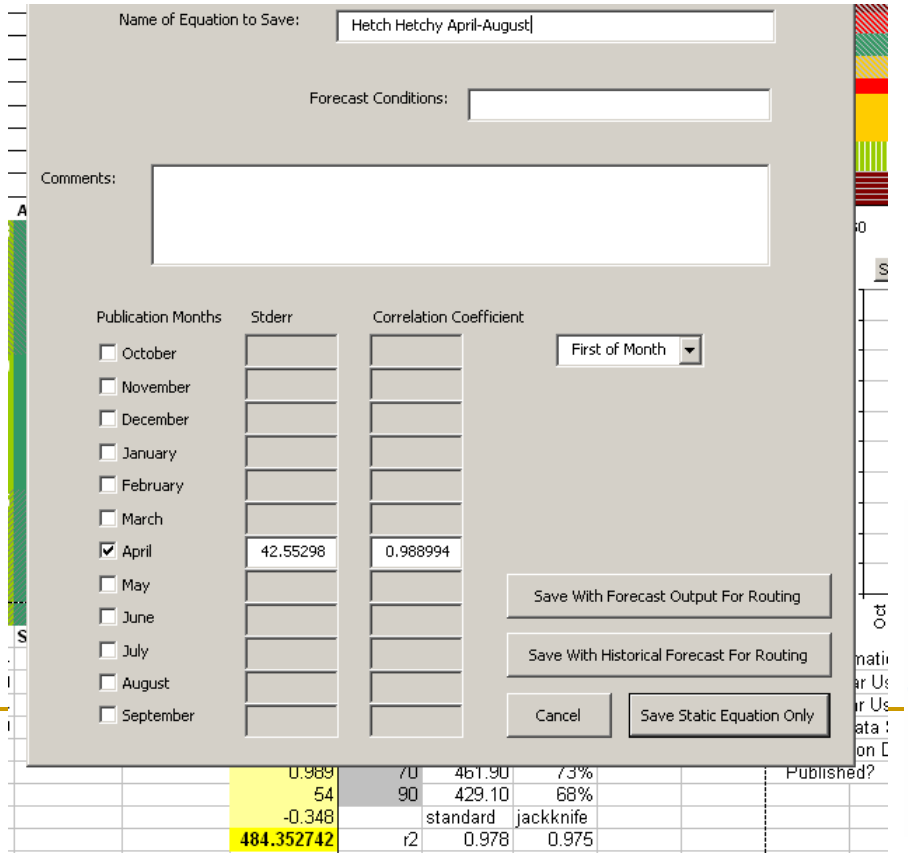
SFPUC
WATER
WASTEWATER
POWER

Continue VIPERing

- Continue Optimization process until you are "satisfied" with an equation that is:
  - Statistically valid
  - Simple and easily explained
  - Physically meaningful
  - Operationally useful

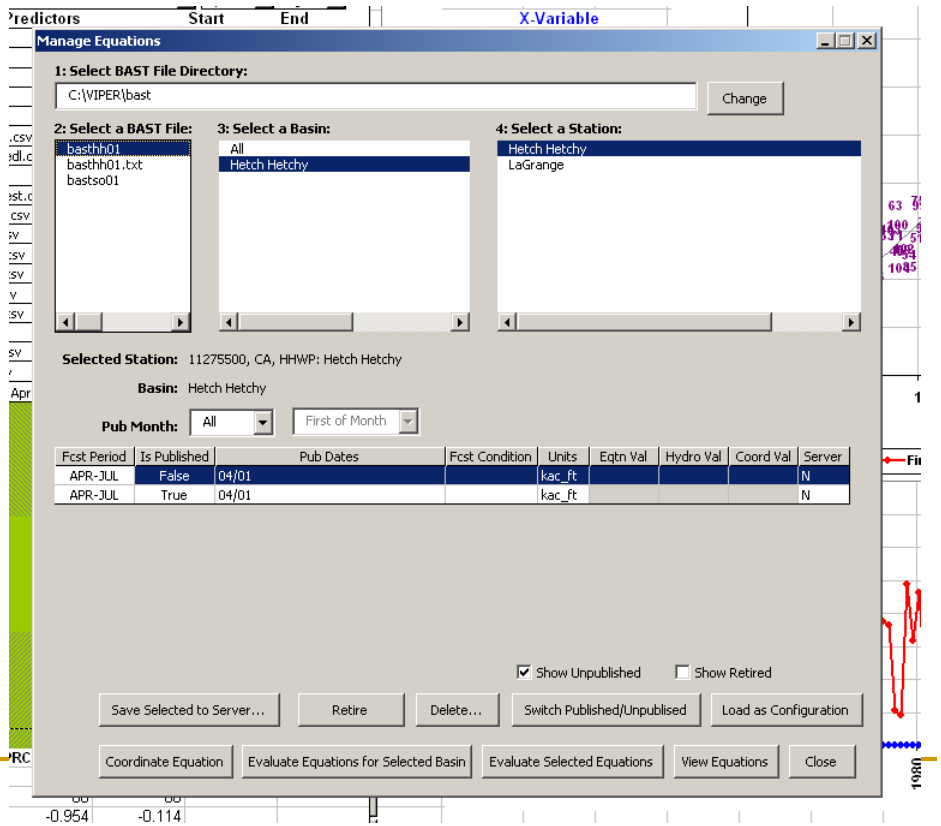
Now you can

- Save and publish the equation
- The equation is published on the equations worksheet



Operational Function of VIPER

- Once equations are developed, you can execute!! – *although clumsy for non-NRCS users*



SFPUC
WATER
WASTEWATER
POWER

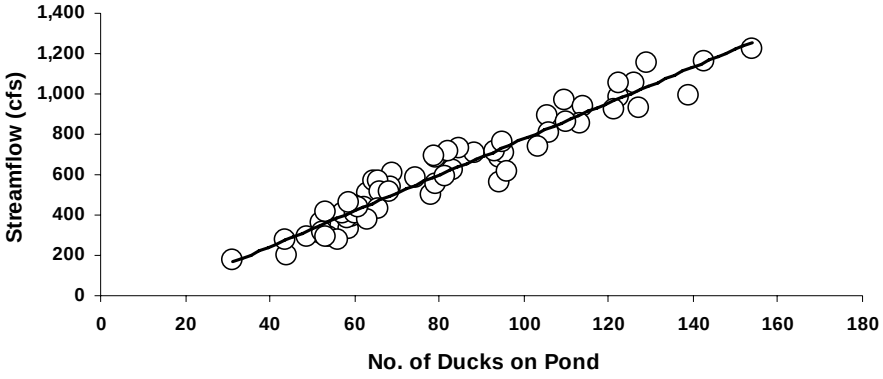
What else does this have?

- Ability to edit NRCS database data
- Archive old forecast equations
- Save configurations
- Execute, review, and publish NRCS forecast

ALL OF WHICH ARE BUILT INTO VIPER, BUT ONLY PERTAIN TO NRCS SYSTEM

SFPUC
WATER
WASTEWATER
POWER

Dangers



- Easy to add irrelevant data
- Optimization process does not replace your hydrologic expertise.
- Z-score easily abused
- Boss may want you to re-work forecast procedure… ….often.

SFPUC
WATER
WASTEWATER
POWER

Other applications

- Estimating missing data
- Identifying station groupings for indices

Summary

USDA Natural Resources Conservation Service
National Water & Climate Center

- VIPER is a new powerful tool which the NRCS developed to make BETTER forecast equations, easier and faster.

- A tool similar to VIPER which is capable of interacting with CDEC would provide everyone access to a powerful forecasting tool.

SFPUC
WATER
WASTEWATER
POWER

Statistics – PCA

- Principal components analysis is a standard multivariate statistical procedure
- Can be used for descriptive purposes to reduce the dimensionality of correlated variables
- Can be taken a step further to provide new, non-correlated independent variables for regression
- PC's taken in order, subject to t-test and sign test
- Final model is expressed in terms of original X variables

SFPUC
WATER
WASTEWATER
POWER