The REG procedure fits linear regression models by least-squares. Subsets of independent variables that ~best~ predict the dependent or response variable can be determined by various model-selection methods.

PROC REG performs the following regression techniques with flexibility: (some omitted here, jh)

- provides nine model-selection methods
- generates scatter plots of data and various statistics
- "paints" or highlights scatter plots
- produces partial regression leverage plots
- computes collinearity diagnostics
- prints predicted values, residuals, studentized residuals, confidence limits, and influence statistics and can output these items to a SAS data set
- performs weighted least-squares regression.

Nine model-selection methods are available in PROC REG. The simplest method is also the default, where REG fits the complete model you specify. The other eight methods involve various ways of including or excluding variables from the model. These methods are specified with the SELECTION= option in the MODEL statement. The methods are identified below and explained in detail in **Model-Selection Methods** later in this chapter.

NONE	no model selection. This is the default. The complete model specified in the MODEL statement is fit to the data.
FORWARD	forward selection. The method starts with no variables in the model and adds variables.
BACKWARD	backward elimination. The method starts with all variables in the model and deletes variables.
STEPWISE	stepwise regression. This is similar to FORWARD except that variables already in the model do not necessarily stay there.
MAXR	forward selection to fit the best one-variable model, the best two-variable model, and so on. Variables are switched so that R2 is maximized.
MINR	similar to MAXR, except that variables are switched so that the increase in R2 from adding a variable to the model is minimized.
RSQUARE	finds a specified number of models with the highest R2 in a range of model sizes.
ADJRSQ	finds a specified number of models with the highest adjusted R2 in a range of model sizes.
CP	finds a specified number of models with the lowest Cp in a range of model sizes.

Model Selection and Details of Selection

model selection	SELECTION =
specify maximum number of variables selected	BEST=
print summary statistics at each step	DETAI LS
provide names for groups of variables	GROUPNAMES =
include first n variables in the model	INCLUDE=
set criterion for entry into model	SLE=
set criterion for staying in model	SLS =
specify number of variables in model to begin the comparing and switching process	START=
stop selection criterion	STOP=

Options for RSQUARE, ADJRSQ, and CP Model Selection

compute adjusted RSQUARE	ADJ RSQ
compute AKAIKE's information criterion	AIC
compute parameter estimates for each model	B
compute Sawa's Bayesian information criterion	BIC
compute Mallow's CP statistic	CP
compute estimated MSE of prediction assuming multivariate normality	GMSEP
compute JP, the final prediction error	JP
compute MSE for each model	MSE
compute Amemiya's prediction criterion	PC
print root MSE for each model	RMSE
compute the SBC statistic	SBC
specify the true standard deviation of error term for computing CP and BIC	SIGMA=
compute SP statistic for each model	SP
compute error SS for each model	SSE

The following options are available in the **MODEL statement** after a slash (/):

…

SELECTION = name specifies the method used to select the model, where name can be FORWARD (or F), BACKWARD (or B), STEPWISE, MAXR, MINR, RSQUARE, ADJRSQ, CP, or NONE (use the full model). The default method is NONE. Only one method can be specified in a MODEL statement. If you want to use several methods, you must use separate model statements for each method. See **Model-Selection Methods** for a description of each method.

…

SLENTRY=value
SLE =value
specifies the significance level for entry into the model used in the FORWARD and STEPWISE methods. The defaults are 0.50 for FORWARD and 0.15 for STEPWISE.

SLSTAY=value
SLS =value
specifies the significance level for staying in the model used in the BACKWARD and STEPWISE methods. The defaults are 0.10 for BACKWARD and 0.15 for STEPWISE.

…

Options in the **MODEL statement** after a slash (/): **continued...**

BEST=

is used with the RSQUARE, ADJRSQ, and CP model-selection methods. If SELECTION= CP or SELECTION= ADJRSQ is specified, the BEST= option specifies the maximum number of subset models to be printed or output to the OUTEST= data set. For SELECTION= RSQUARE, the BEST= option requests the maximum number of subset models for each size.

If the BEST= option is used without the B option (printing estimated regression coefficients), the variables in each MODEL are listed in order of inclusion instead of the order in which they appear in the MODEL statement.

If the BEST= option is omitted and the number of regressors is fewer than 11, all possible subsets are evaluated. If the BEST= option is omitted and the number of regressors is greater than ten, the number of subsets seiected is at most equal to the number of regressors. A small value of the BEST= option greatly reduces the CPU time required for large problems.

CP

computes Mallows' Cp statistic for each model selected (Mallows 1973; Hocking 1976). This option is available in the RSQUARE, ADJRSQ, and CP model-selection methods only.

DETAILS

produces a table of statistics for entry and removal for each variable at each step in the model-building process. This option is available only in the BACKWARD, FORWARD, and STEPWiSE methods. The statistics produced include the tolerance, R^2, and F statistic that results if each variable is added to the model, or the partial and model R^2 that results if the variable is deleted from the model.

INCLUDE =*n*

forces the first *n* independent variables listed in the MODEL statement to be included in all models. The selection methods are performed on the other variables in the MODEL statement. The INCLUDE= option is not available with SELECTION=NONE.

STOP= *s*

causes REG to stop when it has found the ~'best" s-variable model, where s is the STOP value. For the RSQUARE, ADJRSQ, and CP methods, STOP=s specifies the largest number of regressors to be reported in a subset model. For the MAXR and MINR methods, STOP=s specifies the largest number of regressors to be included in the model.The default setting for the STOP= option is the number of variables in the MODEL statement. This option can only be used with the MAXR, MINR, RSQUARE, ADJRSQ and CP methods.

Model-Selection Methods

The nine methods of model selection implemented in PROC REG are specified with the SELECTION= option in the MODEL statement. Each method is discussed below.

Full Model Fitted (NONE)

This method is the default and provides no model selection capability. The complete model specified in the MODEL statement is used to fit the model. For many regression analyses, this may be the only method you need.

Forward Selection (FORWARD)

The forward-selection technique begins with no variables in the model. For each of the independent variables, FORWARD calculates F statistics that reflect the variable's contribution to the model if it is included. The p-values for these *F* statistics are compared to the SLENTRY= value that is specified in the MODEL statement (or to 0.50 if the SLENTRY= option is omitted). If no *F* statistic has a significance level greater than the SLENTRY= value, FORWARD stops. Otherwise, FORWARD adds the variable that has the largest *F* statistic to the model. FORWARD then calculates *F* statistics again for the variables still remaining outside the model, and the evaluation process is repeated. Thus, variables are added one by one to the model until no remaining variable produces a significant *F* statistic. Once a variable is in the model, it stays.

Backward Elimination (BACKWARD)

The backward-elimination technique begins by calculating statistics for a model, including all of the independent variables. Then the variables are deleted from the model one by one until all the variables remaining in the model produce *F* statistics significant at the SLSTAY= level specified in the MODEL statement (or at the 0.10 level if the SLSTAY= option is omitted). At each step, the variable showing the smallest contribution to the model is deleted.

Stepwise (STEPWISE)

The stepwise method is a modification of the forward-selection technique and differs in that variables already in the model do not necessarily stay there. As in the forward-selection method, variables are added one by one to the model, and the F statistic for a variable to be added must be significant at the SLENTRY= level. After a variable is added, however, the stepwise method looks at all the variables already included in the model and deletes any variable that does not produce an *F* statistic significant at the SLSTAY= level. Only after this check is made and the necessary deletions accomplished can another variable be added to the model. The stepwise process ends when none of the variables outside the model has an *F* statistic significant at the SLENTRY= level and every variable in the model is significant at the SLSTAY= level, or when the variable to be added to the model is the one just deleted from it.

Model-Selection Methods ... continued

Maximum R^2 Improvement (MAXR)

The maximum R^2 improvement technique does not settle on a single model. Instead, it tries to find the "best" one-variable model, the ~best~ two-variable model, and so forth, although it is not guaranteed to find the model with the largest R^2 for each size.

The MAXR method begins by finding the one-variable model producing the highest R^2. Then another variable, the one that yields the greatest increase in R^2, is added. Once the two-variable model is obtained, each of the variables in the model is compared to each variable not in the model. For each comparison, MAXR determines if removing one variable and replacing it with the other variable increases R^2. After comparing all possible switches, MAXR makes the switch that produces the largest increase in R^2. Comparisons begin again, and the process continues until MAXR finds that no switch could increase R^2. Thus, the two-variable model achieved is considered the "best" two-variable model the technique can find. Another variable is then added to the model, and the comparing-and-switching process is repeated to find the "best" three-variable model, and so forth.

The difference between the STEPWiSE method and the MAXR method is that all switches are evaluated before any switch is made in MAXR. In the STEPWISE method, the ~worst~ variable can be removed without considering what adding the ~best~ remaining variable might accomplish. MAXR may require much more computer time than STEPWISE.

Minimum R^2 Improvement (MINR)

The MINR method closely resembles MAXR, but the switch chosen is the one that produces the smallest increase in R^2. For a given number of variables in the model, MAXR and MINR usually produce the same ~best~ model, but MINR considers more models of each size.

R^2 Selection (RSQUARE)

The RSQUARE method finds subsets of independent variables that best predict a dependent variable by linear regression in the given sample. You can specify the largest and smallest number of independent variables to appear in a subset and the number of subsets of each size to be selected. The RSQUARE method can efficiently perform all possible subset regressions and print the models in decreasing order of R^2 magnitude within each subset size. Other statistics are available for comparing subsets of different sizes. These statistics, as well as estimated regression coefficients, can be printed or output to a SAS data set.

The subset models selected by RSQUARE are optimal in terms of R^2 for the given sample, but they are not necessarily optimal for the population from which the sample was drawn or for any other sample for which you may want to make predictions. If a subset model is selected on the basis of a large R^2 value or any other criterion commonly used for model selection, then all regression statistics computed for that model under the assumption that the model is given a priori, including all statistics computed by REG, are biased.

While the RSQUARE method is a useful tool for exploratory model building, no statistical method can be relied on to identify the ~true~ model. Effective model building requires substantive theory to suggest relevant predictors and plausible functional forms for the model. italics added by j.h . and a.n. .

The RSQUARE method differs from the other selection methods in that RSQUARE always identifies the model with the largest R^2 for each number of variables considered. The other selection methods are not guaranteed to find the model with the largest R^2 . RSQUARE requires much more computer time than the other selection methods, so a different selection method such as STEPWISE is a good choice when there are many independent variables to consider.

Adjusted R^2 Selection (ADJRSQ)

This method is similar to RSQUARE, except that the adjusted R^2 statistic is used as the criterion for selecting models, and the method finds the models with the highest adjusted R^2 within the range of sizes.

Mallows' Cp Selection (CP)

This method is similar to ADJRSQ, except that Mallow's Cp statistic is used as the criterion for model selection.

Additional Information on Model-Selection Methods

If the RSQUARE or STEPWISE procedure (as documented in SAS User's Guide: Statistics, Version 5 Edition) is requested, PROC REG with the appropriate model-selection method is actually used.

Reviews of model-selection methods by Hocking (1976) and Judge et al. (1980) describe these and other variable-selection methods.

Criteria Used in BACKWARD, FORWARD, and STEPWISE Model-Selection Methods

When many significance tests are performed, each at a level of, say 5 percent, the overall probability of rejecting at least one true null hypothesis is much larger than 5 percent. If you want to guard against including any variables that do not contribute to the predictive power of the model in the population, you should specify a very small significance level. In most applications many of the variables considered have some predictive power, however small. If you want to choose the model that provides the best prediction using the sample estimates, you need only guard against estimating more parameters than can be reliably estimated with the given sample size, so you should use a moderate significance level, perhaps in the range of 10 percent to 25 percent.

In addition to R^2, the Cp statistic is printed for each model generated in the model-selection methods. Cp was proposed by Mallows (1973) as a criterion for selecting a model. It is a measure of total squared error defined as

$$Cp = (SSE_p / s^2) - (N - 2*p)$$

where s^2 is the MSE for the full model, and SSE_p is the error sum of squares for a model with p parameters including the intercept, if any. **If Cp is plotted against p, Mallows recommends the model where Cp first approaches p**. When the right model is chosen, the parameter estimates are unbiased, and this is reflected in Cp near p. For further discussion, see Daniel and Wood (1980).

The adjusted R^2 statistic is an alternative to R^2 that is adjusted for the number of parameters in the model. The adjusted R^2 statistic is calculated as

$$ADJRSQ = 1 - [((n - i)(1 - R^2)) / (n - p)]$$

where n is the number of observations used in fitting the model, and i is an indicator variable that is 1 if the model includes an intercept, and 0 otherwise.

Limitations in Model-Selection Methods

The use of model-selection methods can be time-consuming in some cases because there is no built-in limit on the number of independent variables, and the calculations for a large number of independent variables can be lengthy. The recommended limit on the number of independent variables for the MINR method is 20+i, where i is the value of the INCLUDE= option.

For the RSQUARE, ADJRSQ, or CP methods, with a large value of the BEST= option, adding one more variable to the list from which regressors are selected may significantly increase the CPU time. Also, the time required for the analysis is highly dependent on the data and on the values of the BEST=, START=, and STOP= options.

EXAMPLE

```
data a;
   INFILE 'kkm12_23.dat';

   INPUT Obsn  MONTH  DAY MAXST MINST  AVST
         MAXAT MINAT  AVAT MAXH  MINH   AVH
         WIND EVAP ;
```

The data set WORK.A has 46 observations and 14 variables.

```
proc reg;
 model EVAP=
           MAXST MINST  AVST MAXAT MINAT  AVAT
           MAXH  MINH   AVH    WIND /  DETAILS

   SELECTION = FORWARD SLENTRY = 0.10;
```
see page 5

```
proc reg;
 model EVAP=
           MAXST MINST  AVST  MAXAT MINAT  AVAT
           MAXH  MINH   AVH    WIND /  DETAILS

   SELECTION = BACKWARD  SLSTAY = 0.10 ;
```
see page 6

```
proc reg;
 model EVAP=
          MAXST MINST  AVST   MAXAT MINAT  AVAT
          MAXH  MINH   AVH    WIND /  DETAILS

   SELECTION = STEPWISE  SLENTRY = 0.10 SLSTAY = 0.15;
```
see page 7

```
proc reg;
    model EVAP=
          MAXST MINST  AVST
          MAXAT MINAT  AVAT
          MAXH  MINH   AVH    WIND  /

    SELECTION = CP BEST = 5 ;
```
see page 7 (bottom right)

Variable	N	Mean	Std Dev	Min	Max
MONTH	46	6.4	0.5	6.0	7.0
DAY	46	14.8	7.6	1.0	30.0
MAXST	46	87.5	6.0	73.0	96.0
MINST	46	71.2	3.2	65.0	76.0
AVST	46	173.5	20.0	131.0	202.0
MAXAT	46	90.7	5.0	77.0	97.0
MINAT	46	70.0	3.6	59.0	76.0
AVAT	46	190.5	20.9	147.0	215.0
MAXH	46	94.7	1.2	93.0	98.0
MINH	46	48.5	10.0	24.0	73.0
AVH	46	396.9	29.4	345.0	478.0
WIND	46	277.6	149.0	72.0	663.0
EVAP	46	34.60	14.6	1.0	54.0

Statistics for Entry: Step 1 DF = 1,44

Variable	Tolerance	R**2	F	Prob>F
MAXST	1.000000	0.5917	63.7741	0.0001
MINST	1.000000	0.2909	18.0470	0.0001
AVST	1.000000	0.4727	39.4466	0.0001
MAXAT	1.000000	0.5229	48.2291	0.0001
MINAT	1.000000	0.1106	5.4696	0.0240
AVAT	1.000000	0.5050	44.8827	0.0001
MAXH	1.000000	0.0354	1.6128	0.2108
MINH	1.000000	0.4517	36.2427	0.0001
AVH	1.000000	0.6815	94.1366	0.0001
WIND	1.000000	0.0025	0.1102	0.7415

Step 1 Variable **AVH** Entered R-sq = 0.681 C(p) = 30.5

	DF	Sum of Sq	Mean Sq	F	Prob>F
Regression	1	6570.8	6570.85	94.14	0.0001
Error	44	3071.2	69.80		
Total	45	9642.1			

Variable	Parameter Estimate	Standard Error	Type II Sum of Squares	F	Prob>F
INTERCEP	197.36	16.813	9618.18	137.79	0.0001
AVH	-0.40	0.042	6570.85	94.14	0.0001

Statistics for Entry: Step 2 DF = 1,43

Variable	Tolerance	R**2	F	Prob>F
MAXST	0.426871	0.7302	7.7743	0.0079
MINST	0.768023	0.7076	3.8457	0.0564
AVST	0.538289	0.7113	4.4350	0.0411
MAXAT	0.570359	0.7396	9.5930	0.0034
MINAT	0.995732	0.7594	13.9303	0.0006
AVAT	0.712105	0.7821	19.8543	0.0001
MAXH	0.925976	0.6829	0.1957	0.6604
MINH	0.169767	0.7193	5.7931	0.0205
WIND	0.950307	0.7391	9.4974	0.0036

Step 2 Variable **AVAT** Entered R-sq = 0.782 C(p) = 9.6

	DF	Sum of Sq	Mean Squ	F	Prob>F
Regression	2	7540.995	3770.497	77.16	0.0001
Error	43	2101.113	48.863		
Total	45	9642.108			

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	107.634	24.565	938.055	19.20	0.0001
AVAT	0.262	0.058	970.141	19.85	0.0001
AVH	-0.309	0.041	2672.060	54.68	0.0001

Statistics for Entry: Step 3 DF = 1,42

Variable	Tolerance	R**2	F	Prob>F
MAXST	0.180754	0.7830	0.1831	0.6709
MINST	0.327889	0.7965	2.9728	0.0920
AVST	0.179863	0.8044	4.7811	0.0344
MAXAT	0.187592	0.7832	0.2085	0.6503
MINAT	0.210278	0.7821	0.0059	0.9391
MAXH	0.910728	0.7821	0.0014	0.9699
MINH	0.120726	0.7829	0.1565	0.6944
WIND	0.863893	0.8050	4.9410	0.0317

Step 3 Variable **WIND** Entered R-sq = 0.805 C(p) = 6.3

	DF	Sum of Sq	Mean Sq	F	Prob>F
Regression	3	7762.159	2587.386	57.80	0.0001
Error	42	1879.949	44.760		
Total	45	9642.108			

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	123.9017	24.6244	1133.236	25.32	0.0001
AVAT	0.2227	0.0591	635.678	14.20	0.0005
AVH	-0.3429	0.0427	2876.508	64.26	0.0001
WIND	0.0159	0.0071	221.164	4.94	0.0317

Statistics for Entry: Step 4 DF = 1,41

Variable	Tolerance	R**2	F	Prob>F
MAXST	0.175268	0.8050	0.0041	0.9490
MINST	0.324093	0.8159	2.4241	0.1272
AVST	0.163746	0.8169	2.6548	0.1109
MAXAT	0.169303	0.8053	0.0489	0.8260
MINAT	0.178793	0.8099	1.0434	0.3130
MAXH	0.841911	0.8067	0.3471	0.5590
MINH	0.113772	0.8051	0.0140	0.9065

No other variable met the 0.10 significance level for entry into the model.

Summary: Forward Selection Procedure for Dep. Var. EVAP

Step	Variable Entered	Number In	Partial R**2	Model R**2	C(p)	F	Prob>F
1	AVH	1	0.6815	0.6815	30.51	94.1	0.0001
2	AVAT	2	0.1006	0.7821	9.61	19.8	0.0001
3	WIND	3	0.0229	0.8050	6.39	4.9	0.0317

Step 0 **All** Variables **Entered** R-sq = 0.846 C(p) = 11.0

	DF	Sum of Sq	Mean Sq	F	Prob>F
Regression	10	8159.83	815.98	19.27	0.0001
Error	35	1482.27	42.35		
Total	45	9642.10			

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	-54.074	130.720	7.24	0.17	0.6816
MAXST	2.231	1.003	209.31	4.94	0.0328
MINST	0.204	1.104	1.45	0.03	0.8539
AVST	-0.742	0.349	191.06	4.51	0.0408
MAXAT	0.501	0.568	32.84	0.78	0.3845
MINAT	0.304	0.788	6.29	0.15	0.7022
AVAT	0.092	0.218	7.56	0.18	0.6750
MAXH	1.109	1.133	40.62	0.96	0.3341
MINH	0.751	0.487	100.51	2.37	0.1324
AVH	-0.556	0.161	501.85	11.85	0.0015
WIND	0.008	0.009	40.07	0.95	0.3373

Statistics for Removal: Step 1 DF = 1,35

Variable	Partial R**2	Model R**2
MAXST	0.0217	0.8246
MINST	0.0002	0.8461
AVST	0.0198	0.8265
MAXAT	0.0034	0.8429
MINAT	0.0007	0.8456
AVAT	0.0008	0.8455
MAXH	0.0042	0.8421
MINH	0.0104	0.8358
AVH	0.0520	0.7942
WIND	0.0042	0.8421

Step 1 **MINST** Removed R-sq = 0.846 C(p) = 9.0

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	-45.673	120.965	5.87	0.14	0.7080
MAXST	2.194	0.970	210.71	5.11	0.0299
AVST	-0.696	0.242	339.33	8.23	0.0068
MAXAT	0.509	0.559	34.09	0.83	0.3691
MINAT	0.316	0.775	6.86	0.17	0.6855
AVAT	0.089	0.214	7.22	0.18	0.6780
MAXH	1.080	1.106	39.26	0.95	0.3355
MINH	0.735	0.473	99.34	2.41	0.1293
AVH	-0.546	0.151	537.81	13.05	0.0009
WIND	0.009	0.008	47.18	1.14	0.2917

Statistics for Removal: Step 2 DF = 1,36

Variable	Partial R**2	Model R**2
MAXST	0.0219	0.8243
AVST	0.0352	0.8109
MAXAT	0.0035	0.8426
MINAT	0.0007	0.8454
AVAT	0.0007	0.8454
MAXH	0.0041	0.8420
MINH	0.0103	0.8358
AVH	0.0558	0.7903
WIND	0.0049	0.8412

Step 2 **MINAT** Removed R-sq = 0.845 C(p) = 7.1

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	-18.178	99.342	1.34	0.03	0.8558
MAXST	2.046	0.890	212.97	5.29	0.0272
AVST	-0.653	0.216	368.44	9.15	0.0045
MAXAT	0.506	0.553	33.81	0.84	0.3655
AVAT	0.144	0.165	30.62	0.76	0.3889
MAXH	0.889	0.991	32.39	0.80	0.3756
MINH	0.716	0.466	95.14	2.36	0.1328
AVH	-0.525	0.140	564.39	14.01	0.0006
WIND	0.010	0.007	76.14	1.89	0.1775

Statistics for Removal: Step 3 DF = 1,37

Variable	Partial R**2	Model R**2
MAXST	0.0221	0.8233
AVST	0.0382	0.8072
MAXAT	0.0035	0.8419
AVAT	0.0032	0.8422
MAXH	0.0034	0.8420
MINH	0.0099	0.8355
AVH	0.0585	0.7869
WIND	0.0079	0.8375

Step 3 **AVAT** Removed R-sq = 0.842 C(p) = 5.9

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	-59.989	86.726	19.15	0.48	0.4933
MAXST	2.380	0.801	352.98	8.82	0.0051
AVST	-0.665	0.214	384.06	9.59	0.0037
MAXAT	0.720	0.494	85.01	2.12	0.1533
MAXH	1.314	0.861	93.29	2.33	0.1351
MINH	0.999	0.332	361.31	9.03	0.0047
AVH	-0.605	0.106	1295.30	32.36	0.0001
WIND	0.012	0.007	110.57	2.76	0.1047

Statistics for Removal: Step 4 DF = 1,38

Variable	Partial R**2	Model R**2
MAXST	0.0366	0.8056
AVST	0.0398	0.8024
MAXAT	0.0088	0.8334
MAXH	0.0097	0.8326
MINH	0.0375	0.8048
AVH	0.1343	0.7079
WIND	0.0115	0.8308

Step 4 **MAXAT** Removed R-sq = 0.833 C(p) = 5.9

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	-52.989	87.831	14.99	0.36	0.5498
MAXST	2.761	0.768	531.99	12.92	0.0009
AVST	-0.608	0.214	332.17	8.07	0.0071
MAXH	1.539	0.859	132.17	3.21	0.0810
MINH	1.095	0.330	451.54	10.96	0.0020
AVH	-0.631	0.106	1455.46	35.34	0.0001
WIND	0.011	0.007	96.78	2.35	0.1334

Statistics for Removal: Step 5 DF = 1,39

Variable	Partial R**2	Model R**2
MAXST	0.0552	0.7782
AVST	0.0345	0.7990
MAXH	0.0137	0.8197
MINH	0.0468	0.7866
AVH	0.1509	0.6825
WIND	0.0100	0.8234

Step 5 **WIND** Removed R-sq = 0.823 C(p) = 6.2

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	-46.405	89.194	11.52	0.27	0.6057
MAXST	3.190	0.727	818.20	19.22	0.0001
AVST	-0.722	0.204	532.35	12.50	0.0010
MAXH	1.377	0.867	107.42	2.52	0.1201
MINH	1.326	0.299	835.39	19.62	0.0001
AVH	-0.674	0.104	1780.19	41.81	0.0001

Statistics for Removal: Step 6 DF = 1,40

Variable	Partial R**2	Model R**2
MAXST	0.0849	0.7385
AVST	0.0552	0.7682
MAXH	0.0111	0.8122
MINH	0.0866	0.7367
AVH	0.1846	0.6387

Step 6 **MAXH** Removed R-sq = 0.812 C(p) = 6.7

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	74.345	47.516	108.09	2.45	0.1254
MAXST	3.028	0.733	752.01	17.03	0.0002
AVST	-0.669	0.205	469.66	10.64	0.0022
MINH	1.207	0.295	738.42	16.72	0.0002
AVH	-0.622	0.100	1681.21	38.07	0.0001

Statistics for Removal: Step 7 DF = 1,41

Variable	Partial R**2	Model R**2
MAXST	0.0780	0.7342
AVST	0.0487	0.7635
MINH	0.0766	0.7357
AVH	0.1744	0.6379

All variables left in the model are significant at the 0.10 level.

Summary: Backward Elimination Procedure for EVAP

Step	Variable Removed	Number In	Partial R**2	Model R**2	C(p)	F	Prob>F
1	MINST	9	0.0002	0.846	9.0	0.0344	0.8539
2	MINAT	8	0.0007	0.845	7.1	0.1667	0.6855
3	AVAT	7	0.0032	0.842	5.9	0.7602	0.3889
4	MAXAT	6	0.0088	0.833	5.9	2.1236	0.1533
5	WIND	5	0.0100	0.823	6.2	2.3500	0.1334
6	MAXH	4	0.0111	0.812	6.7	2.5232	0.1201

Stepwise Procedure for Dependent Variable EVAP

Statistics for Entry: Step 1 DF = 1,44

Variable	Tolerance	Model R**2	F	Prob>F
MAXST	1.000000	0.5917	63.7741	0.0001
MINST	1.000000	0.2909	18.0470	0.0001
AVST	1.000000	0.4727	39.4466	0.0001
MAXAT	1.000000	0.5229	48.2291	0.0001
MINAT	1.000000	0.1106	5.4696	0.0240
AVAT	1.000000	0.5050	44.8827	0.0001
MAXH	1.000000	0.0354	1.6128	0.2108
MINH	1.000000	0.4517	36.2427	0.0001
AVH	1.000000	0.6815	94.1366	0.0001
WIND	1.000000	0.0025	0.1102	0.7415

Step 1 **AVH** Entered R-square = 0.681 C(p) = 30.5

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	197.369	16.813	9618.18	137.79	0.0001
AVH	-0.409	0.042	6570.85	94.14	0.0001

Statistics for Entry: Step 2 DF = 1,43

Variable	Tolerance	Model R**2	F	Prob>F
MAXST	0.426871	0.7302	7.7743	0.0079
MINST	0.768023	0.7076	3.8457	0.0564
AVST	0.538289	0.7113	4.4350	0.0411
MAXAT	0.570359	0.7396	9.5930	0.0034
MINAT	0.995732	0.7594	13.9303	0.0006
AVAT	0.712105	0.7821	19.8543	0.0001
MAXH	0.925976	0.6829	0.1957	0.6604
MINH	0.169767	0.7193	5.7931	0.0205
WIND	0.950307	0.7391	9.4974	0.0036

Step 2 **AVAT** Entered R-sq = 0.782 C(p) = 9.6

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	107.634	24.565	938.05	19.20	0.0001
AVAT	0.262	0.058	970.14	19.85	0.0001
AVH	-0.309	0.041	2672.06	54.68	0.0001

Statistics for Removal: Step 3 DF = 1,43

Variable	Partial R**2	Model R**2
AVAT	0.1006	0.6815
AVH	0.2771	0.5050

Statistics for Entry: Step 3 DF = 1,42

Variable	Tolerance	Model R**2	F	Prob>F
MAXST	0.180754	0.7830	0.1831	0.6709
MINST	0.327889	0.7965	2.9728	0.0920
AVST	0.179863	0.8044	4.7811	0.0344
MAXAT	0.187592	0.7832	0.2085	0.6503
MINAT	0.210278	0.7821	0.0059	0.9391
MAXH	0.910728	0.7821	0.0014	0.9699
MINH	0.120726	0.7829	0.1565	0.6944
WIND	0.863893	0.8050	4.9410	0.0317

Step 3 **WIND** Entered R-sq = 0.805 C(p) = 6.3

Variable	Parameter Estimate	Standard Error	Type II Sum of Sq	F	Prob>F
INTERCEP	123.901	24.624	1133.23	25.32	0.0001
AVAT	0.222	0.059	635.67	14.20	0.0005
AVH	-0.342	0.042	2876.50	64.26	0.0001
WIND	0.015	0.007	221.16	4.94	0.0317

Statistics for Removal: Step 4 DF = 1,42

Variable	Partial R**2	Model R**2
AVAT	0.0659	0.7391
AVH	0.2983	0.5067
WIND	0.0229	0.7821

Statistics for Entry: Step 4 DF = 1,41

Variable	Tolerance	Model R**2	F	Prob>F
MAXST	0.175268	0.8050	0.0041	0.9490
MINST	0.324093	0.8159	2.4241	0.1272
AVST	0.163746	0.8169	2.6548	0.1109
MAXAT	0.169303	0.8053	0.0489	0.8260
MINAT	0.178793	0.8099	1.0434	0.3130
MAXH	0.841911	0.8067	0.3471	0.5590
MINH	0.113772	0.8051	0.0140	0.9065

All variables left in the model are significant at the 0.15 level.
No other variable met the 0.10 significance level for entry into the model.

Summary of Stepwise Procedure for Dependent Variable EVAP

Step	Variable Entered	Removed	Number In	Partial R**2	Model R**2	C(p)	F	Prob>F
1	AVH		1	0.6815	0.6815	30.5	94.1	0.0001
2	AVAT		2	0.1006	0.7821	9.6	19.8	0.0001
3	WIND		3	0.0229	0.8050	6.3	4.9	0.0317

==========

N = 46 Regression Models for Dependent Variable: EVAP

C(p)	R-sq	Number In	Variables in Model
3.75	**0.834**	**5**	**MAXST AVST AVAT MINH AVH**
3.88	**0.833**	**5**	**MAXST AVST AVAT AVH WIND**
4.06	**0.824**	**4**	**MAXST AVST AVAT AVH**
4.64	**0.839**	**6**	**MAXST AVST AVAT MINH AVH WIND**
5.12	**0.836**	**6**	**MAXST AVST MINAT MAXH MINH AVH**