Airplane Model Structure Determination from Flight Data

V. Klein*

JIAFS/GWU Langley Research Center, Hampton, Virginia and

J. G. Batterson† and P. C. Murphy†
NASA Langley Research Center, Hampton, Virginia

A procedure for airplane model structure determination from flight data based on modified stepwise regression (MSR), several decision criteria, and postulated aerodynamic model equations is presented. The MSR is constructed to force a linear model for the aerodynamic coefficient first, then to select significant nonlinear terms and reject nonsignificant terms from the model. In addition to the statistical criteria in the stepwise regression, the prediction sum of squares criterion and analysis of residuals are examined for the selection of an adequate model. The procedure is used in examples with simulated and real flight data. It is shown that the MSR performs better than the ordinary stepwise regression and that the technique can be also applied to the large amplitude maneuvers.

Introduction

NTEREST in post-stall and spin flights has created a need to extend parameter estimation into flight regimes where nonlinear effects could become pronounced. Because of the uncertainty in the model form for these regimes, the model structure determination from flight data should be included in the airplane identification procedure. The first treatment of the airplane model structure determination based on the stepwise regression was presented in Ref. 1. The technique was applied to simulated data and in limited extent to the flight data. The extension of this research work is covered in Ref. 2, where the review of various criteria for the selection of the "best" model is also included. Further application of the technique mentioned can be found in Refs. 3 and 4.

The purpose of this paper is to re-examine the applicability of stepwise regression to the determination of airplane model structure from flight data. The emphasis is given to the development and interpretation of criteria, which would enable the researcher to select the "best" model for a given test run, and to the verification of the model selected.

Modified Stepwise Regression

The linear regression technique is employed to estimate the functional relationship of a dependent variable to one or more independent variables. In the case of airplane models, the resultant aerodynamic force and moment are expressed by means of the aerodynamic model equations, which may be written as

$$y(t) = \theta_0 + \theta_1 x_1(t) + \dots + \theta_{n-1} x_{n-1}(t)$$
 (1)

In this equation, y(t) represents the resultant coefficient of aerodynamic force or moment (the dependent variable); θ_1 through θ_{n-1} are the stability and control derivatives; θ_0 is the value of any particular coefficient corresponding to the initial steady flight conditions; and x_1 through x_{n-1} are the airplane state and control variables (the independent variables). The variables x_1 through x_{n-1} may also include any combination of the state and control variables.

When a sequence of N observations on both y and x has been made at times $t_1, t_2, ..., t_N$, then the measured data can be related by the following set of N linear equations:

$$y(i) = \theta_0 + \theta_1 x_1(i) + \dots + \theta_{n-1} x_{n-1}(i) + \epsilon(i)$$

$$i = 1, 2, \dots, N$$
(2)

Because Eq. (1) is only an approximation of the actual aerodynamic relations, the right-hand side of Eq. (2) includes an additional term, $\epsilon(i)$, often referred to as the equation error. For N > n, the unknown parameters can be estimated from the measurements by the least-squares technique.

The stepwise regression is a procedure which inserts independent variables into the regression model until the regression equation is satisfactory. The order of insertion is determined by using the partial correlation coefficient as a measure of the importance of variables not yet in the regression equation.

At every step of the regression the variables incorporated into the model in previous stages and a new variable entering the model are re-examined using the F statistic. The partial F_p value is given as

$$F_p = \hat{\theta}_i^2 / s^2 (\hat{\theta}_i)$$
 $j = 1, 2, ...$ (3)

where $\hat{\theta}_j$ is the estimate of the parameter θ_j , and s^2 ($\hat{\theta}_j$) is the variance of estimate $\hat{\theta}_i$.

The process of selecting and checking variables continues until no more variables will be admitted to the equation and no more are rejected. The complete computing scheme for the stepwise regression can be found in Ref. 5.

In this paper a model will be described by a model structure (analytical representation of a model) and model parameters (coefficients in the analytical representation). Because the correct model of an airplane is unknown, a major problem in system identification is the selection, from measured data, of an adequate model. An adequate model is a model which sufficiently fits the data, facilitates the successful estimation of unknown parameters, and has good prediction capabilities.

For the model structure determination procedure the following assumptions will be made:

- 1) The general equations of a rigid body motion adequately define the airplane motion.
- 2) The model for the aerodynamic force and moment coefficients can be represented by multivariable polynomials in response and control variables. The parameters in these

Presented as Paper 81-1866 at the AIAA Atmospheric Flight Mechanics Conference, Albuquerque, New Mex., Aug. 19-21, 1981; submitted Oct. 5, 1981; revision received June 7, 1982. This paper is declared a work of the U.S. Government and therefore is in the public domain.

^{*}Research Professor. Associate Fellow AIAA.

[†]Aerospace Technologist. Member AIAA.

equations are the coefficients of the Taylor series expansion around the values corresponding to the initial steady-state flight.

3) Linear terms in the Taylor series expansion make generally larger contributions to aerodynamic functions, followed by higher-order terms. The third assumption will result in a constraint on the selection of significant terms in the regression equation.

The determination of an adequate model using the stepwise regression includes these three steps; namely, the postulation of terms which might enter the final model, the selection of an adequate model, and finally the verification of the model selected. An example of a postulated model is given by the yawing moment coefficient, which was expressed as $C_n = C_{n_0}$ plus linear terms plus the variation of aerodynamic derivatives with the angle of attack α plus the nonlinear dependence of C_n on the sideslip angle β plus the variation of C_n with α .

The computing scheme for the selection of an adequate model is modified from that in Ref. 5. The linear terms in the model are examined first. They enter the regression according to their partial correlation coefficients and are kept in the model regardless of the value of F_p . This means that during this part of the procedure no hypothesis testing is applied. When all linear terms are included, the nonlinear terms postulated are searched and the null hpothesis concerning their significance and significance of all terms already included in the model is tested. The stepwise regression technique with the constraint mentioned will be further referred to as the modified stepwise regression (MSR). Because of the constraint, the MSR provides the information about the performance of a linear model. The constraint was substantiated by many examples, using both the simulated and real data. Its effect is also mentioned in examples presented.

Selection of an Adequate Model

Experience with several test runs showed that the model based only on the statistical significance of individual parameters in the regression equation can still include too many parameters. The selection of an adequate model is facilitated by invoking the "principle of parsimony," which states: Given two models fitted to the same data with residual variances σ_1^2 and σ_2^2 , which are close to one another, choose the model with the fewer parameters. It is therefore recommended that more quantities and their variations be examined as possible criteria. They include the following:

- 1) Compute the values of F_p for each parameter in the model. It should have the maximum value for an adequate model.
- 2) Compute the value of the square of the multiple correlation coefficient R^2 , which can be interpreted as a measure of the usefulness of the terms other than θ_0 in the model. However, the improvement in R^2 due to adding new terms to the model must have some real significance and should not only reflect the effect of the increased number of parameters. The square of the multiple correlation coefficient can be computed from the expression

$$R^2 = \frac{\theta^T X^T Y - N\bar{y}^2}{Y^T Y - N\bar{y}^2} \tag{4}$$

where θ^T is the $n \times 1$ vector of unknown parameters; X is the $N \times n$ matrix of measured independent variables; Y is the $N \times 1$ vector of measured values of y(i); and

$$\bar{y} = (1/N) \sum_{i=1}^{N} y(i)$$

3) Compute the value of F, which is given as the ratio of (regression mean square) to (residual mean square), or

$$F = \frac{\theta^T X^T Y - N \bar{y}^2}{(n-1)s^2} \tag{5}$$

where

$$s^2 = \frac{1}{N-n} \sum_{i=1}^{N} \hat{\epsilon}^2(i)$$

The model with the maximum F value has already been recommended in Ref. 1 as the "best" one for a given set of data.

4) Compute the prediction sum of squares (PRESS) criterion proposed in Ref. 7 and defined as

PRESS =
$$\sum_{i=1}^{N} [y(i) - y(i|x(1),...,$$

$$x(i-1),x(i+1),...,x(N))$$
² (6)

where y(i|...) is the estimate of $E\{y(i)\}$, using the investigated subset and excluding the *i*th observation. This criterion is based on the principle of the mean square prediction error (MSPE), which can be expressed as the variance of the response plus the variance of the prediction plus the squared bias of prediction.

It can be shown (see Ref. 8) that in a model with a redundant number of parameters, MSPE will increase owing to the increase of the variance of the prediction. For the incomplete model the MSPE will increase owing to the bias in prediction. For an effective computing of PRESS, Eq. (6) can be modified to

PRESS =
$$\sum_{i=1}^{N} \frac{[y(i) - \hat{y}(i)]^{2}}{1 - \text{var}\{y(i)\}/\sigma^{2}}$$
(7)

where σ^2 is the variance of the response and y(i) is now based on all data points.

It follows from Eq. (7) that for $N\rightarrow\infty$, the PRESS approaches the sum of squares of residuals. Experience with flights containing in excess of about 100 data points indicated that even that number of points led to asymptotic behavior. It was found that from about 30 to 40 points gave a good PRESS result. Therefore a reduced number of data points (e.g., every tenth point) was used in computing the PRESS. Then the sufficient sensitivity of PRESS criterion to the changes in the number of parameters included in the model was maintained.

5) For an adequate model, the time history of the residuals $\epsilon(i)$ should be close to a random sequence which is uncorrelated and Gaussian.

Model Verification

Checks on the accuracy of estimated parameters and the prediction qualities of the selected model are considered as the verification of the model. The parameter estimates can be compared with the results from repeated measurements under the same conditions, that is, the same flight conditions and input forms. Further, the least-squares estimates can be compared with the estimates using different techniques but the same data and model. For this comparison the maximum likelihood method is recommended because of its optimal asymptotical properties. Finally, the parameter estimates must have physical values and should be compared with wind tunnel results and theoretical predictions.

Examples

In the following three examples the modified stepwise regression was applied to various sets of simulated and measured data of a general aviation airplane. In all cases the general equations of the airplane motion were used.

Table 1	Effect of measurement noise on model structure and	d narameter estimates for simulated data
1 able 1	Effect of measurement hoise on model structure and	a parameter estimates for simulated dat

			Estimate, $\hat{\theta}$	
Parameter	Time value and noise-free estimate	Noise on dependent variable	Noise on independent and dependent variables	
			Low noise level	High noise level
$\frac{C_{Y_0}}{C}$	0.0069	0.0088	0.0088	0.0084
$C_{V_0}^{I_0}$	-0.555	-0.557	-0.556	-0.553
$C_{Y_{oldsymbol{eta}}}^{V_{oldsymbol{G}}}$	-0.103	-0.103	-0.102	-0.101
L - **	0.88	0.795	0.710	0.640
$C_{Y_{\delta a}}$	-0.075	-0.077	-0.074	-0.069
	0.05	0.050	0.056	0.050
$C_{Y_{p\alpha}}^{Y_{\delta r}}$	1.34	1.44	1.60	1.53
$C_{Yr\alpha}^{'p\alpha}$ $C_{Yr\alpha}^{'p\alpha}$ $C_{Y^2}^{\alpha}$ C_{l_0} C_{l_p} $C_{l_r}^{l_r}$ $C_{l_{\delta\sigma}}^{l_r}$	$-51 \\ 0.47$	• • • • • • • • • • • • • • • • • • • •	•••	•••
$R^{2}{}^{\alpha}$ 0%	•••	98.9	98.7	98.4
C_{i}	-0.00042	-0.00027	-0.0011	0.0012
$C_{l}^{'0}$	-0.11	-0.108	-0.107	-0.105
$C_{i}^{'\beta}$	-0.15	-0.145	-0.141	-0.139
$C_{i}^{'p}$	0.21	0.197	0.255	0.228
$C_{t}^{\prime r}$	-0.09	-0.092	-0.094	-0.093
$C_{i}^{t\delta a}$	0	0 .	0	-0.003
$C_{l}^{'\delta r}$	1.0	1.05	0.92	0.87
$C^{ou}_{l_{\delta r}} \ C_{l_{plpha}} \ R^2$, %	•••	95.2	95.2	94.6
C	0.00099	0.00109	0.00102	0.00105
$C_{-}^{n_0}$	0.03	0.0296	0.0270	0.0260
$C^{n_{\beta}}$	-0.063	-0.063	-0.064	- 0.065
$C_{n_0} \ C_{n_{eta}} \ C_{n_p} \ C_{n_r} \ C_{n_r} \ C_{n_{\delta r}}$	-0.084	-0.064	-0.086	-0.090
$C_{-}^{n_r}$	0.013	0.013	0.016	0.016
$C_{-}^{n_{\delta r}}$	-0.033	-0.032	-0.031	-0.031
$C_{n_{\delta a}}^{n_{\delta r}}$	0.77	0.359	0.856	0.838
$C_{n_{p\alpha}}^{\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	-1.33	-1.34	3.330	
R^2 , %		85.0	84.1	83.0

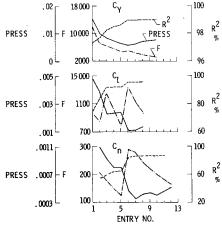


Fig. 1 Values of F, PRESS, and R^2 at different entries of modified stepwise regression; simulated data, dependent variables in error.

Example 1

For the first example, a simulated data set was created using a fourth-order Runge-Kutta integration computer program with a step size of 0.0001 s. The aerodynamic model in the integration was that estimated by applying the MSR to a high-angle-of-attack lateral maneuver which exhibited longitudinal oscillations due to coupling effects. When applied to the simulated data, the MSR selected the correct model structure and parameter estimates, thus verifying the MSR in a noise-free environment. Next, zero mean Gaussian noise was added to the lateral dependent variables C_{γ} , C_{l} , C_{n} . The standard deviation for this measurement noise was that estimated from real flight data. The results of this application of the MSR are

summarized in Table 1. In the side force equation, the selection of a model consisting of the linear terms plus the $p\alpha$ term is based on the maximum F value after the MSR was allowed to consider all candidate variables. This model also corresponds to that indicated by the minimum PRESS value (Fig. 1). The $r\alpha$ and α^2 terms were two of the next three terms to enter the model. None of the parameter estimates was statistically different from its true value. In its application to the rolling moment equation, the MSR was tested for its ability to delete a linear term (that it was earlier constrained to include) once that term was shown to be insignificant. The control term δ_r was eliminated after the MSR was allowed to search the nonlinear candidate variables. The true model corresponded to both the maximum F value and the minimum PRESS. Again the parameter estimates did not deviate statistically from their true values.

When applied to the yawing moment equation, the MSR yields a six term model by maximum F value. In this model all linear parameter estimates are within 2σ of their true value. However, $C_{n_{p\alpha}}$, the nonlinear parameter, did deviate more than 2σ from its true value. When applying the PRESS criterion, one finds the seven parameter model (corresponding to the true model) to be best. When the $r\alpha$ term is added under this criterion, the $C_{n_{p\alpha}}$ estimate regresses to within 1σ of its true value.

As a measure of the robustness of the MSR, it was next applied to two cases in which both the dependent variables C_Y , C_I , C_n and the linear model variables β , p, r were corrupted by zero mean Gaussian noise. Two cases were examined. The standard deviation of the model variable noise in the first case was that estimated from the ground calibration of an instrumentation system. In the second case, five times higher noise levels were applied to the same model variables.

Table 2	Comparison of yawing moment parameter estimates for the maximum F-value model
	and the F_n selected model

		Estimate, $\hat{ heta}$			
		Low noise level		High noise level	
Т	rue value	From maximum F value	From F_p and and maximum F value	From maximum F value	From F_p and maximum F value
$\overline{C_{no}}$	0.00099	0.00102	0.00109	0.00106	0.00114
$C_{n_0}^{n_0}$	0.03	0.027	0.030	0.026	0.029
$C_n^{"\beta}$	-0.063	0.064	-0.063	-0.062	-0.062
$C_n^{"p}$	-0.084	0.086	-0.099	-0.090	-0.103
$C_{n}^{"r}$	0.013	0.016	0.014	0.016	0.015
$C_{n}^{''\delta a}$	-0.033	-0.031	-0.034	-0.031	-0.034
$C_n^{"or}$	0.77	0.86	0.79	0.838	0.78
C_{n_0} $C_{n_{eta}}$ C_{n_p} C_{n_r} $C_{n_{\delta a}}$ $C_{n_{\delta r}}$ $C_{n_{\delta r}}$ $C_{n_{p \alpha}}$ $C_{n_{r \alpha}}$	-1.33	***	-1.23	•••	-1.31

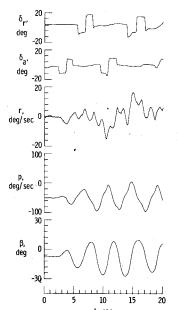


Fig. 2 Measured time histories of input and output variables.

In these last two cases one would expect the possibility of biased parameter estimates from an equation error method.

With the lower level of noise, the MSR reaches a maximum F value at six variables for the side force equation. However, the PRESS selects two additional variables that were not in the simulation model. This emphasizes a third piece of information available to the MSR user. The user can examine the F_p 's for each of the variables in the regression at a given point. If newly added variables have significantly lower F_p 's than those already in the model, one should pick the less complex model equal to or greater than that corresponding to the maximum F value. Using this method of examining the F_p 's for those models greater in complexity than that corresponding to $F_{\rm max}$, one extracts the correct model structure for the rolling and yawing moment equations at both levels of noise. For the side force equation, one still does not extract the correct model in that the two terms of least significance, $r\alpha$ and α^2 , never enter the regression.

The parameter estimates for both the maximum F value model and the model extracted by examination of F_p 's are presented for the yawing moment equation in Table 2. The modification which constrains the MSR to first fit the linear model is an important feature. For the five cases in which noise was added to the model variables, an unconstrained stepwise regression was inconsistent between the PRESS and maximum F criteria as to the best model structure. Also, terms that were not in the simulated model were accepted in certain best models for an unconstrained stepwise regression.

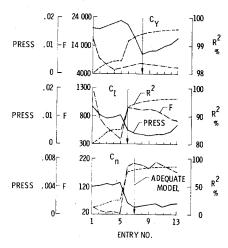


Fig. 3 Values of F, PRESS, and R^2 at different entries of modified stepwise regression; flight data, example 2.

Therefore this example substantiates the use of the MSR rather than the stepwise regression without constraint.

Example 2

In this example the MSR technique for model structure determination was applied to the measured data. The data, sampled at 0.05 s, represent a lateral response of the airplane at $\alpha \pm 20$ deg. The time histories of the input and some response variables are plotted in Fig. 2. The response variables indicate that the airplane exhibits a periodic lateral motion which is also strongly coupled with the short period longitudinal mode. In Fig. 3 the F, PRESS, and R^2 values for the lateral coefficients examined are plotted against the number of entry into the MSR.

An adequate model for the side force coefficient was selected at the eighth entry, where PRESS has its minimum and F the second maximum. For the coefficient C_i the F criterion indicates an adequate model at the sixth entry, the PRESS at the ninth. The difference in \mathbb{R}^2 at these two entries is only 2%. Therefore, considering the principle of parsimony, the model with the smaller number of parameters was selected. For the coefficient C_n the changes in the F, PRESS, and R^2 values after the fifth entry are apparent. These changes indicate that the linear model (first five entries) is completely inadequate and that some nonlinear terms must be included. An adequate model was selected at the seventh entry, where the PRESS values have their minimum and F values their first maximum. The comparison between measured time histories of C_n and those computed by using the linear and an adequate model is presented in Figs. 4 and 5. Also included are the autocorrelation functions of residuals.

Fig. 4 Time histories of measured and computed yawing moment coefficient and autocorrelation function of residuals; linear model.

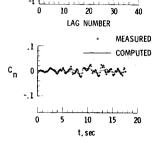
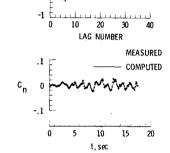


Fig. 5 Time histories of measured and computed yawing moment coefficient and autocorrelation function of residuals; adequate model.



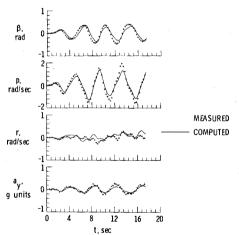


Fig. 6 Measured lateral flight data time histories and those computed using parameters obtained from modified stepwise regression.

For the linear model the fit to the data is poor. By adding two nonlinear terms, $p\alpha$ and $r\alpha$, the fit was improved substantially and the autocorrelation function of residuals was close to that for the uncorrelated random variable.

The variables included in adequate models for the three coefficients are summarized as follows in the order as they entered the model: for C_Y : β , δ_r , r, δ_a , p, $p\alpha$, $r\alpha^2$, α^2 ; for C_l : β , p, δ_a , r, $p\alpha$; and for C_n : δ_α , p, β , δ_r , r, $p\alpha$, $r\alpha$. In Fig. 6 the measured output time histories are compared

In Fig. 6 the measured output time histories are compared with those predicted by using the model for C_{γ} , C_{l} , and C_{n} determined by the MSR. The agreement in these time histories is good except for the yawing velocity.

The next step in the airplane identification included the parameter estimation by using the maximum likelihood method with the model structure determined by the MSR. In this estimation process the nonlinear parameters were kept fixed on the least-squares estimates. Any attempt to estimate the whole set of aerodynamic parameters failed because of a divergence in the ML algorithm. The resulting ML and MSR estimates of parameters in the yawing moment equation are presented in Table 3. Some differences in the estimated parameters from both methods exist, mainly in the cross-derivative C_{n_p} . All these differences might be caused by undetected modeling error and by the correlation between linear and nonlinear parameters. Simulated studies of the flight regime analyzed also showed that the data were very sensitive to even small changes in certain parameters.

The model structures for the three coefficients C_{γ} , C_{l} , C_{n} were also determined by the stepwise regression without constraint. The resulting models included the following variables: for C_{γ} : β , $\beta\alpha^{2}$, δ_{r} , β^{3} ; for C_{l} : β , β^{3} , $\beta\alpha^{2}$, p, δ_{a} , $p\alpha$, and for C_{n} : $p\alpha$, β^{3} , p, δ_{r} , β , $r\alpha$.

As in the previous example with simulated data, these models are different from those determined by the MSR. In the second and third set, for example, the linear parameter C_{l_r} and C_{n_r} are missing. To obtain the output variables, a numerical integration was performed using the control time histories from the flight in which the models were identified. Where the new aerodynamic model equations were used for prediction of the output variables, a divergent motion of the airplane resulted. The first set based on the MSR gave the model which very well described the motion of the airplane, whereas the second set only fit equally well the time histories of C_Y , C_l , C_n but failed to predict the airplane motion correctly.

The physical meaning of some of the estimated nonlinear parameters can be assessed from Fig. 7, where the C_{n_p} estimates from five test runs are plotted against the angle of attack. The values of parameters $C_{n_{p\alpha}}$ and $C_{n_{p\alpha}^2}$ (slope and curvature of solid lines) agree well with the trend in the change of C_{n_p} with α . Also plotted in Fig. 7 are the ML estimates of C_{n_p} using adequate models as determined by the MSR.

Table 3 Parameters and their standard errors estimated from measurements using two estimation methods

	MSR		ML		
Parameter	Estimate, $\hat{\theta}$	Standard error, $s(\hat{\theta})$	Estimate, $\hat{ heta}$	Standard error, $s(\hat{\theta})$	
C_{n_0}	-0.00086		-0.00042	0.000061	
$C_{n_{\beta}}$	0.0316	0.00097	0.0300	0.000068	
$C_{n_p}^{\rho}$	-0.0616	0.0019	-0.0392	0.00093	
$C_{n_r}^{n_p}$	-0.071	0.017	-0.094	0.0050	
$C_{n_{\delta a}}^{r}$	0.013	0.0027	0.0225	0.00082	
$C_{n_{\delta r}}^{"oa}$	-0.033	0.0016	-0.027	0.0012	
$C_{n_{p\alpha}}^{n_{or}}$	0.77	0.015	0.77a	•••	
$C_{n_{r\alpha}}^{''p\alpha}$	-1.3	0.14	-1.3^{a}	•••	

^aFixed values. ^bCramer-Rao lower bound.

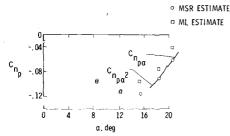


Fig. 7 Comparison of lateral parameter estimated from flight data using modified stepwise regression and maximum likelihood method.

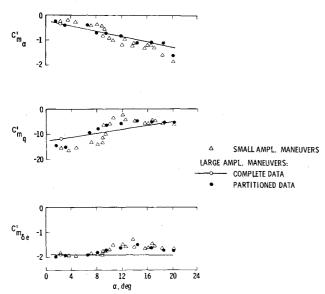


Fig. 8 Estimated longitudinal parameters from flight data using modified stepwise regression.

Example 3

In the last example the data from a longitudinal large amplitude maneuver were analyzed. The MSR selected a form of an adequate model for the coefficient C_m which included the terms α , α^2 , q, $q\alpha$, δ_e . The resulting parameters and their variations with the angle of attack are plotted in Fig. 8. These results are compared with the parameters obtained from 21 transient maneuvers initiated from pre- and post-stall steadystate flight regimes (triangle symbols). In these 21 maneuvers the excitation of the motion was considerably smaller than that in the large amplitude maneuver analyzed.

The models for the large amplitude maneuver include the linear variation of some parameter values with α . The variations agree with the trend given by the results from the small amplitude maneuvers. This agreement was improved upon by partitioning the data from the large amplitude maneuver into six subsets according to the values of α . The first subset included the data with α varying from its minimum value to 4 deg. The second subset consisted of data corresponding to α between 4 and 8 deg, and so forth, until the sixth subset was filled with data corresponding to α between 20 and 24 deg. This partitioning was then repeated starting with the subset of data values for α between 22 and 26 deg. An adequate model was determined for each data subset by applying the MSR. The resulting parameters are plotted in Fig. 8 (closed symbols). The parameters from the partitioned data agree well with the results from 21 maneuvers. They therefore better describe the variations of the parameters with α than the estimates from the complete set, thus indicating a preferable way of analyzing large amplitude maneuvers.

Concluding Remarks

A procedure for the determination of airplane model structure and parameters from pre- and post-stall flight data has been presented. The procedure was demonstrated on both simulated data and real flight test data. Designated MSR (modified stepwise regression), the method postulates the aerodynamic coefficients as multivariable polynomials in input and output variables. The MSR consists of a stepwise regression algorithm that has been constrained to consider the linear model before considering any nonlinear terms. At each step, the MSR is complemented by the prediction sum of squares (PRESS) criterion, total F-value calculation, and partial F-value calculations for all variables currently in the model.

The following points can be drawn from the work reported herein:

- 1) The MSR can determine from simulated data the true airplane model better than a stepwise regression without constraint. Also the MSR chooses a model with better prediction capabilities as demonstrated on flight data.
- 2) A combination of criteria (viz., PRESS, F value, partial F values) facilitates the selection of a parsimonious model. In applying the PRESS criterion, a limited subset of the data string should be employed (e.g., every tenth point of a data string consisting of 300 points).
- 3) At the output variable noise levels considered for the simulated data, MSR is not adversely affected in its selection of an adequate model.
- 4) The nonlinear terms selected by the MSR in the pre- and post-stall regimes significantly enhance the fit of the aerodynamic coefficients and "whitens" the residual sequence.
- 5) The MSR can be applied to large amplitude maneuvers by applying it to the entire data length or by partitioning the data string as a function of variables, e.g. the angle of attack, the influence of which might lead to the existence of nonlinear

The procedure presented represents the first step toward the determination of an overall model of an airplane from flight data. When properly used it can provide results for better understanding of airplane aerodynamics at high angles of attack and for global stability and control analysis of an airplane at these flight conditions.

References

¹Hall, W.E. Jr., Gupta, N.K., and Tyler, J.S. Jr., "Model Structure Determination and Parameter Identification for Nonlinear Aerodynamic Flight Regimes," Methods for Aircraft State and Parameter Identification, AGARD-CP-172, May 1975, pp. 21-1-21-

21.
²Gupta, N.K., Hall, W.W. Jr., and Trankle, T.L., "Advanced Methods of Model Structure Determination from Test Data," AIAA

Paper 77-1170, Aug. 1977.

³Stalford, H.S., "Application of the Estimation-Before-Modeling (EBM) System Identification Method to the High Angle of Attack/Sideslip Flight of the T-2C Jet Trainer Aircraft," Vol. III, Identification of T-2C Aerodynamic Stability and Control Characteristics from Actual Flight Test Data, Dynamic Research Corporation, Rept NADC-76097-30, April 1979.

⁴Vincent, J.H., Hall, W.E. Jr., and Bohn, J.G., "Analysis of T-2C High Angle of Attack Flight Test Data with Nonlinear System Identification Methodology," Office of Naval Research, Rept. No.

ONR-CR212-259-1F, 1980.
⁵ Draper, N.R. and Smith, H., Applied Regression Analysis, John Wiley and Sons, Inc., New York, 1966.

⁶Kashyap, R.L., "A Bayesian Comparison of Different Classes of Dynamic Models Using Empirical Data," IEEE Transactions on Automatic Control, Vol. AC-22, No. 5, Oct. 1977, pp. 715-727.

⁷Allen, D.M., "The Prediction Sum of Squares as a Criterion for Selecting Predictor Variables," Tech. Rept. 23, University of Kentucky, Lexington, Ky., Aug. 1971.

⁸ Walls, R.C. and Weeks, D.L., "A Note on the Variance of a Predicted Response in Regression," *The American Statistician*, Vol. 23, June 1969, pp. 24-26.