300
6-15-81
JAL

(2)

DR-2748

Bins - 200
N.T. 1-S-25

DOE/MC/14361-T1

MASTER

**EMPIRICAL SYSTEMS MODEL OF THE WESTINGHOUSE
FLUIDIZED BED PDU COAL GASIFIER**

Task Report No. 1

By
Nelson H. Kemp

December 1980

Work Performed Under Contract No. AC21-80MC14361

Physical Sciences Inc.
Woburn, Massachusetts

**FOSSIL
ENERGY**

U. S. DEPARTMENT OF ENERGY

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

DISCLAIMER

"This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof."

This report has been reproduced directly from the best available copy.

Available from the National Technical Information Service, U. S. Department of Commerce, Springfield, Virginia 22161.

Price: Printed Copy A02
Microfiche A01

TASK REPORT #1

EMPIRICAL SYSTEMS MODEL OF THE WESTINGHOUSE
FLUIDIZED BED PDU COAL GASIFIER

by
Nelson H. Kemp

December 1980

Prepared for

U. S. Department of Energy
Morgantown Energy Technology Center
Morgantown, WV 26505

Under Contract No. DE-AC21-80MC14361

Prepared by

Physical Sciences Inc.
30 Commerce Way
Woburn, MA 01801

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
INTRODUCTION	1
MODEL ALGORITHMS	2
REFERENCES	7
APPENDIX A	8
APPENDIX B	14
APPENDIX C	18

LIST OF TABLES

<u>Table</u>		<u>Page</u>
I	REGRESSION COEFFICIENTS	4/5/6
A-1	INPUT DATA	12
A-2	OUTPUT DATA	13
C-1	PERCENTAGE ERROR OF PREDICTIONS	22
C-2	PERCENTAGE ERROR OF TEST CASES	23

INTRODUCTION

A systems-level model of the Westinghouse PDU coal gasifier has been constructed. This model is based entirely on the use of the data gathered by testing this gasifier, as presented in quarterly reports from Westinghouse (Refs. 1-8). It represents a purely empirical attempt to correlate the published data by means of multiple linear regression analysis, leading to prediction equations for output variables as linear combinations of a number of input variables.

The prediction equations so derived do not have any physical laws incorporated into them explicitly. They rely entirely on measurements of the physical processes going on in the gasifier-cyclone combination. However, the choice of relevant input variables was made with the help of some intuition about the variables which affect these physical processes, based on work performed at PSI to model coal gasifiers. It should also be noted that the data comes from engineering tests which were not necessarily designed to produce systematic variation of all variables, so the data may contain some biases about which variables are varied.

The computer implementation of the model is very fast and simple. For each output variable a set of 11 coefficients are prescribed. To obtain a predicted value of that output variable, for a specified set of values of the 10 input variables, a linear combination (dot product) of the 10 input variables is formed, using the 11 coefficients.

The use of this empirical model is restricted to predictions for the Westinghouse PDU coal gasifier (and associated cyclone) within the range of operating variables represented in the data base, as described below. In particular, the data is all for operation on Pittsburgh Seam Coal. However, there is a range of coal feed rates included, and both air-blown and oxygen-blown runs were included.

In the next section the algorithms will be described. In Appendix A the data used to construct the model is presented. In Appendix B the method of linear regression used is given. In Appendix C the limitations of the model are described.

MODEL ALGORITHMS

An analysis of the Westinghouse gasifier data showed that there were ten input variables that defined the operating condition and input flows. These ten variables are:

P = Pressure level (psig)
FC = Coal flow rate (lb/hr)
TC = Coal flow temperature (°F)
FO = Oxygen flow rate (lb/hr)
FN = Nitrogen flow rate (lb/hr)
TON = Oxygen and nitrogen flow temperature (°F)
FS = Steam flow rate (lb/hr)
TS = Steam flow temperature (°F)
FT = Transport gas flow rate (lb/hr)
TT = Transport gas flow temperature (°F)

Prediction equations were constructed for 11 output variables:

TB = Bed temperature (°F)
TP = Product gas temperature (°F)
FNET = Net output gas flow rate (lb/hr)
FCO = CO output flow rate (lb/hr)
FCO2 = CO2 output flow rate (lb/hr)
FH2 = H2 output flow rate (lb/hr)
FCH4 = CH4 output flow rate (lb/hr)
FH2S = H2S output flow rate (lb/hr)
FN2 = N2 output flow rate (lb/hr)
FH2O = H2O output flow rate (lb/hr)
FSOL = Solid output flow rate (lb/hr)

For each output variable, say Z, the prediction equation has the form

$$Z = B + \sum_{i=1}^{10} A_i u_i \quad (1)$$

where u_i are the ten input variables. The 11 coefficients B and A_i are determined by a least squares fit of the data.

Table I presents the set of coefficients for each of the 11 output variables. The first coefficient in each column, opposite the input variable 1, is B . The remaining ten coefficients in each column are associated with the corresponding input variables listed in the left-hand column.

To make a prediction for any set of input variables, a linear combination is made of those variables with the coefficients for the desired output variable. To this is added the appropriate coefficient B , and the result is the predicted value of the output variable.

TABLE I
REGRESSION COEFFICIENTS

OUTPUT VARIABLE: INPUT VARIABLE	<u>TB</u>	<u>TP</u>	<u>FNET</u>	<u>FCO</u>
1	2.0543E3	1.6620E3	3.3199E3	6.7702E2
P	-2.2542E-1	9.8582E-2	6.0231E0	2.2115E0
FC	-6.4938E-2	-7.7966E-2	1.7145E0	9.0885E-1
TC	1.0979E0	-1.8359E-1	-5.0844E0	-7.7327E-1
FO	5.9134E-1	6.1107E-1	-1.2740E0	5.0557E-1
FN	-1.1905E-2	1.8167E-2	1.6478E0	-4.3281E-2
TON	5.8988E-2	1.1663E-1	-4.8080E-1	-5.2152E-1
FS	-3.0216E-1	-5.7759E-2	9.1149E-1	-2.3896E-1
TS	-5.1013E-1	-4.3659E-1	-9.2361E-2	-1.2302E0
FT	-1.6542E-1	-1.3423E-1	-7.9275E-1	-5.3638E-2
TT	-1.5978E-2	-3.1415E-1	-2.3108E0	3.0368E-1

TABLE I (Cont.)

OUTPUT VARIABLE: INPUT VARIABLE	<u>FCO2</u>	<u>FH2</u>	<u>FCH4</u>	<u>FH2S</u>
1	1.5565E3	-6.5615E0	-3.0671E1	1.1931E1
P	1.5801E0	1.2319E-1	1.7612E-1	6.5300E-2
FC	-2.1778E-1	5.7488E-2	2.5370E-2	6.8888E-2
TC	-3.7263E0	-2.9323E-2	-1.5217E-1	-1.3630E-3
FO	-1.7855E-1	-8.3009E-3	-1.3979E-2	-5.7668E-2
FN	3.8006E-1	-9.0005E-4	-1.6481E-3	-4.3004E-3
TON	-3.7506E-1	-6.9698E-3	-2.1557E-2	-1.2876E-2
FS	5.4065E-1	1.2396E-2	2.7555E-2	1.2776E-2
TS	1.5160E0	6.2377E-3	7.8439E-2	-5.9696E-2
FT	-3.3298E-1	-9.4446E-2	1.5911E-3	8.8790E-3
TT	-1.1990E0	-7.3401E-3	-5.5638E-4	-1.9924E-2

TABLE I (Cont.)

OUTPUT VARIABLE: INPUT VARIABLE	<u>FN2</u>	<u>FH2O</u>	<u>FSOL</u>
1	-2.6346E2	1.3636E3	-8.9453E1
P	1.4259E0	3.9357E-1	2.5912E-1
FC	7.2101E-1	1.4204E-1	-5.2381E-2
TC	-3.0476E-2	-4.7379E-1	1.5232E-1
FO	-9.8481E-1	-5.2153E-1	2.1497E-1
FN	1.0639E0	2.5132E-1	1.2660E-2
TON	3.9001E-1	8.4742E-2	-1.7644E-1
FS	1.5652E-1	4.0056E-1	-1.9949E-2
TS	-5.4560E-1	1.8476E-1	-2.7577E-1
DT	2.9850E-2	-4.3302E-1	8.1714E-2
TT	4.4994E-1	-1.8642E0	3.0269E-1

REFERENCES

1. "Advanced Coal Gasification System for Electric Power Generation," Second Quarterly Progress Report, Jan. 1 - March 31, 1978, FE-1514-84, Advanced Coal Conversion Department, Westinghouse Electric Corporation, Madison, PA, June 21, 1978.
2. Same series as Ref. 1, Quarterly Progress Report Third Quarter - Fiscal Year 1978, Period April 1 to June 30, 1978; FE-1514-88, Nov. 30, 1978.
3. Same series as Ref. 1, Quarterly Progress Report Fourth Quarter - Fiscal Year 1978, Period July 1 to September 30, 1978; FE-1514-93, Dec. 28, 1978.
4. Same series as Ref. 1, Quarterly Progress Report First Quarter - Fiscal Year 1979, Period October 1 to December 31, 1978; FE-1514-97, June 30, 1979.
5. Same series as Ref. 1, Quarterly Progress Report Second Quarter - Fiscal Year 1979, Period January 1 to March 31, 1979; FE-1514-101, August 31, 1979. ,
6. Same series as Ref. 1, Quarterly Progress Report Third Quarter - Fiscal Year 1979, Period April 1 to June 30, 1979; FE-1514-105, Sept. 30, 1979.
7. Same series as Ref. 1, Quarterly Progress Report Fourth Quarter - Fiscal Year 1979, Period July 1 to September 30, 1979; FE-1514-109, Dec. 21, 1979.
8. Same series as Ref. 1, Quarterly Progress Report First Quarter - Fiscal Year 1980, Period October 1 to December 31, 1979; FE-1514-113, Jan. 25, 1980.

APPENDIX A

WESTINGHOUSE GASIFIER DATA

The input data was obtained from the information presented in the Westinghouse quarterly reports for 1978 and 1979 (Refs. 1-8). These reports contain information on 24 set points run with Pittsburgh Seam Coal.

For each set point, a table entitled "Gasifier Heat and Material Balances" is presented. This table gives the flow rate, temperature, composition and heat content of each of 18 streams which are input to or output from the gasifier-cyclone combination. Many of these streams have similar compositions and temperatures, and can be combined for a systems model.

This Appendix will describe how the data used to construct the present systems model was obtained from the data given in the Westinghouse reports, and will provide tabulations of the data used.

Input Data

- 1) Pressure level P: The system pressure is specified for each set point in tables entitled "Summary of Operating Data for Modified Gasifier Test TP-XXX-X."
- 2) Coal flow rate FC: This is taken from the solid part of Stream No. 11.
- 3) Coal flow temperature: This is taken from the temperature of Stream No. 11.
- 4) Oxygen flow rate FO: This is taken as the flow rate of Stream No. 12 or 23.2% of the flow rate of Stream No. 4, which is the weight fraction of oxygen in air.
- 5) Nitrogen flow rate FN: This is taken as 75.5% of the flow rate of Stream No. 4, which is the weight fraction of nitrogen in air.
- 6) Oxygen and nitrogen flow temperature TON: This is taken as the temperature of Stream Nos. 4 or 12, whichever one has a non-zero flow rate.

7) Steam flow rate FS: There are three streams through which steam may be put into the gasifier, Stream Nos. 1, 2 and 3, and each may have a different flow rate and temperature. The variable FS is the sum of the flow rates of these three streams.

8) Steam flow temperature TS: To account for the temperature differences in the three steam input streams, the temperature of the combined steam flow is taken as the weight-flow-weighted temperature of the three streams, with the temperature measured in degrees Rankine for weighting purposes, and the weighted temperature converted back to degrees F. The equation used is thus

$$TS = -460 + \frac{\sum_{i=1}^3 T_i (^{\circ}R) W_i}{\sum_{i=1}^3 W_i} \quad (A-1)$$

where W_i , T_i are the flow rate and temperature of the three steam streams.

9) Transport gas flow rate FT: There are eight streams which may contain transport gas. Stream Nos. 6, 7, 8, 9, 10, 11 and 14 are input streams, and Stream No. 16 is an output stream. However, Stream Nos. 9 and 10 are input to the cyclone, not the gasifier, so the transport gas input flow is taken as the sum of the gaseous components of Stream Nos. 6, 7, 8, 11 and 14. The transport gas is essentially product gas re-cycled, and one may at first think it unnecessary to include it in a systems model. However, it provides a substantial flow of gas through the gasifier, and so may be important to the fluid-mechanical behavior of the gasifier. For this reason, it is included as an input variable.

10) Transport gas flow temperature TT: The various transport gas streams have different temperatures, so again a weight-flow-weighted average temperature was used. It was calculated from the gaseous components of Stream Nos. 6, 7, 8, 11 and 14 as described under input variable number 8, the stream flow temperature TS, using Eq. (A-1), but this time summed over these five transport gas streams.

The resulting values of the ten input variables for the Pittsburgh Seam Coal set points used are presented in Table A-1. Each line is the data for one set point identified by the Westinghouse test and set point number. Although there are 24 such set points, only 22 of them were used, because two had missing data. For Test TP-017#4 the temperature of Stream No. 7 is listed as 1.0°F, which is obviously incorrect. For Test TP-018#3, the temperature of Stream No. 12 is listed as 0.0°F, also incorrect. Therefore, these two set points were not used in the data base.

Output Data

1) Bed temperature TB: This variable is specified for each set point in tables entitled "Summary of Operating Data for Modified Gasifier Test TP-XXX-X."

2) Product gas temperature TP: This is taken as the temperature of Stream No. 15.

3) Net output gas flow rate FNET: The actual amount of gas produced by the gasifier needs to be calculated, since some of the product is recycled as transport gas. A study of the tables "Gasifier Heat and Material Balances," suggests that Stream No. 15, Product Gas, is measured before the gas to be re-cycled is taken off, so that some of the Product Gas is really input gas. The net inflow of transport gas is found by subtracting the gaseous flow rate of Stream No. 16 from the sum of the gaseous flow rates of Stream Nos. 6, 7, 8, 9, 10, 11 and 14. The former is an output stream to the gasifier or cyclone. The net amount of product gas is then taken as the gaseous flow rate of Stream No. 15 minus the net inflow of transport gas.

4)-10) Species output flow rates FCO, FCO₂, FH₂, FCH₄, FH₂S, FN₂, FH₂O: These represent the flow rates of the various species contained in the product gas. They are found from the volumetric composition of the product gas tabulated in Stream No. 15 by first calculating the fractional weight flow of each species, and then multiplying by the net product gas flow FNET. The fractional weight flow of species j is

$$W_j = V_j M_j / \sum_j V_i M_i \quad (A-2)$$

where V_i and M_i are the volume % and molecular weight of species i . (The sum in the denominator of Eq. (A-2) is actually 100 times the molecular weight of the gaseous part of Stream No. 15, which is also tabulated in the Westinghouse tables.) When W_j is known, then the output flow rate F_j of species j is

$$F_j = W_j \text{ FNET} \quad (A-3)$$

11) Solid output flow rate FSOL: This is the solid component of the product flow rate from Stream No. 15.

The resulting values of the output variables are given in Table A-2 for the 22 cases used in construction of the prediction algorithms.

The regressions were performed by fitting each column of Table A-2 by a linear combination of the data given in Table A-1. The method of accomplishing this is described in Appendix B.

TABLE A-1
INPUT DATA

Set Point	P psig	FC lb/hr.	TC °F	FO lb/hr	FN lb/hr	TON °F	FS lb/hr	TS °F	FT lb/hr	TT °F
016 # 1	199	763	161	679	2211	1003	248	492	2122	381
# 2	229	670	173	619	2014	1003	246	501	2395	399
# 3	229	803	182	802	2611	992	244	511	2452	385
# 4	229	770	127	712	2317	1003	536	519	1991	411
017-1# 1	230	529	210	565	1841	1005	254	526	2557	325
# 2	230	676	170	724	2358	1005	246	553	2895	335
# 3	232	731	109	813	2646	995	133	411	2785	343
# 5	231	984	163	1028	3344	855	528	541	2714	425
018-2# 4	129	520	209	517	0	500	797	509	2010	198
# 5	129	505	206	502	0	500	794	512	2093	201
# 6A	129	995	166	823	0	500	795	516	1865	202
018-3# 1	130	418	200	481	0	519	794	517	1416	336
# 2	130	555	188	660	0	579	794	520	1715	365
018-4# 1	132	457	213	500	0	515	769	506	1519	370
# 2	132	493	194	554	0	516	773	502	1580	363
# 3	132	690	154	640	0	518	710	498	1571	362
018-5# 1	129	705	138	630	0	513	765	486	1255	424
# 2	129	692	152	648	0	515	578	498	1517	372
# 3A	129	692	160	656	0	513	499	485	1543	385
# 3B	130	670	136	653	0	511	498	458	1523	363
# 4	130	690	121	656	0	514	353	496	1465	345
022-1# 4	130.5	695	102	656	0	506	652	455	1266	362

TABLE A-2
OUTPUT DATA

Set Point	TB OF	TP OF	FNET lb/hr	FCO lb/hr	FCO2 lb/hr	FH2 lb/hr	FCH4 lb/hr	FH2S lb/hr	FN2 lb/hr	FH2O lb/hr	FSOL lb/hr
016 # 1	1871	1533	4872	747	1167	26.3	15.3	0	2594	319	105
# 2	1860	1522	4633	724	1019	24.6	15.9	0	2634	213	210
# 3	1930	1580	5335	875	1120	27.3	11.4	0	3101	199	125
# 4	1760	1495	5686	691	1627	29.7	28.5	0	2858	448	43
017-1# 1	1759	1384	3606	419	988	13.1	15.0	0	2040	130	74
# 2	1768	1458	4549	647	1083	19.9	23.6	0	2681	92.4	130
# 3	1864	1553	5257	967	1144	17.4	13.4	0	3014	100	198
# 5	1835	1611	6658	1069	1594	38.1	35.0	0	3680	239	270
018-2# 4	1725	1377	1685	401	976	17.8	20.4	15.4	0	253	35
# 5	1695	1362	1672	368	889	16.6	19.0	13.6	0	366	27
# 6A	1844	1528	2368	1043	882	45.3	36.0	26.2	0	334	76
018-3# 1	1816	1376	1701	389	882	18.0	19.3	0	0	391	66
# 2	1841	1478	1564	538	872	21.1	20.3	0	0	111	97
018-4# 1	1808	1375	1474	437	778	16.6	17.5	0	0	196	30
# 2	1774	1402	1577	505	837	18.1	18.4	0	0	198	67
# 3	1803	1439	1758	733	824	24.0	21.3	0	0	155	90
018-5# 1	1848	1429	2100	829	1002	35.0	31.2	14.9	0	187	26
# 2	1885	1475	1926	873	859	31.0	25.0	11.2	0	116	36
# 3A	1899	1473	1727	848	792	28.2	21.7	10.2	0	26.0	50
# 3B	1880	1484	1769	730	900	23.1	18.9	7.88	0	87.5	49
# 4	1926	1493	1883	868	874	23.2	16.6	0	0	101	86
022-1# 4	1892	1674	2637	1048	1139	45.5	38.6	12.9	6.36	345	70.6

APPENDIX B

MULTIPLE LINEAR REGRESSION

Multiple linear regression is a standard curve-fitting technique which minimizes the square error of the fit. To accomplish such a fit one assumes there are N input variables for each of M experiments, $M > N + 1$. This means there is an $M \times N$ matrix of input variables $U = u_{ki}$, $k = 1, 2, \dots, M$; $i = 1, 2, \dots, N$. An element of the matrix is the value of the i -th input variable for the k -th experiment.

There are also M values of an output variable Z_k , one for each experiment. The regression model then expresses the predicted value of Z , say $\hat{Z}$, as a linear combination of the input variables u_i .

$$\hat{Z} = b + \sum_{i=1}^N a_i u_i \quad (B-1)$$

If this is applied to each experiment for which we know the output variables we have the M equations

$$\hat{Z}_k = b + \sum_{i=1}^N a_i u_{ki}, \quad k = 1, 2, \dots, M \quad (B-2)$$

as the predicted values of Z . This leads to an error

$$\epsilon_k = \hat{Z}_k - Z_k = b + \sum_{i=1}^N u_{ki} a_i - Z_k \quad (B-3)$$

The fit is obtained by minimizing the sum of square errors with respect to b and the a_i .

$$E = \sum_{k=1}^M \epsilon_k^2 = \sum_{i=1}^M \left[b + \sum_{i=1}^N u_{ki} a_i - z_k \right]^2 .$$

The derivative with respect to b , when set to zero, yields

$$b = \bar{z} - \sum_{i=1}^N \bar{u}_i a_i \quad (B-4)$$

where the averages over the M experiments are defined as

$$\bar{z} = \sum_{k=1}^M z_k / M , \quad \bar{u}_i = \sum_{k=1}^M u_{ki} / M . \quad (B-5)$$

The derivative with respect to a_i , when set equal to zero, yields

$$\sum_{i=1}^M z_k u_{kj} - \sum_{i=1}^N \sum_{j=1}^M u_{ki} u_{kj} a_i - b \sum_{k=1}^M u_{kj} = 0 .$$

If b is eliminated using Eq. (B-4), we find

$$\sum_{k=1}^M z_k u_{kj} - M \bar{z} \bar{u}_j = \sum_{i=1}^N a_i \left(\sum_{k=1}^M u_{ki} u_{kj} - M \bar{u}_j \bar{u}_i \right) .$$

It is easy to see that a useful alternate form of this equation is

$$\sum_{k=1}^M (z_k - \bar{z}) (u_{kj} - \bar{u}_j) = \sum_{i=1}^N a_i \left(\sum_{k=1}^M (u_{ki} - \bar{u}_i) (u_{kj} - \bar{u}_j) \right) . \quad (B-6)$$

This represents $j = 1, 2, \dots, N$ equations in the N unknown coefficients a_i , and shows that these coefficients depend only on the deviations of u and Z from their mean values over the M experiments. Once the a_i are known, then the coefficient b is found from Eq. (B-4).

As an aside, if we put Eq. (B-4) into the prediction equation (B-1), we find

$$\hat{Z} - \bar{Z} = \sum_{i=1}^N (u_i - \bar{u}) a_i$$

which shows that the average point $\bar{Z}, \bar{u}$ is on the regression curve, and also that the average value over the M experiments of the predictions $\hat{Z}_k$ is $\bar{Z}$, the same as the average of the actual values.

The solution of the N equations (B-6) can be easily expressed in matrix terms. Let U^* be the $M \times N$ matrix whose elements are $u_{ki} - \bar{u}_i$ and Z^* be the $M \times 1$ column vector whose elements are $Z_k - \bar{Z}$. Then Eqs. (B-4) become

$$U^{*T} Z^* = U^{*T} U^* a \quad (B-7)$$

where a is the $N \times 1$ column vector whose elements are the a_i . (The superscript T denotes the transpose.) The solution of this equation is obtained from the inverse of the $N \times N$ symmetric matrix $U^{*T} U^*$ as

$$a = \left(U^{*T} U^* \right)^{-1} U^{*T} Z^* \quad (B-8)$$

We now have all the coefficients of the regression curve (B-1). N values of a_i are found from Eq. (B-8) and the value of b from Eq. (B-4).

A frequently used measure of the goodness of the fit which has been achieved is the correlation coefficient R , where

$$R^2 = \frac{\sum_{k=1}^M (z_k - \bar{z}) (\hat{z}_k - \bar{z})}{\sum_{k=1}^M (z_k - \bar{z})^2}$$

A value of $R^2 = 1$ is a perfect fit, $\hat{z}_k = z_k$. Low values of R^2 indicate poor fits. This measure of the goodness of fit depends on the overall fit to all M values of Z , not on the errors of the individual values of the prediction.

It can be shown that the minimized sum of square errors E is related to R^2 by

$$\frac{E}{\underline{Z}^{*T} \underline{Z}^*} = 1 - R^2.$$

This equation is often expressed by saying that R^2 shows how much of the variation in the data has been explained by the fit.

Notice that the only time-consuming algebraic operation needed to find the coefficients is the inversion of the matrix $U^{*T} U^*$. This operation depends only on the input variables u_{ki} , not the output variables. If there are several different output variables, the coefficients for each one are obtained from Eqs. (B-8) and (B-4). But the inverse matrix is the same for all of them, and, in fact, the $N \times M$ matrix $(U^{*T} U^*)^{-1} U^{*T}$ needs to be calculated only once for all output variables. Then the simple matrix multiplication of Eq. (B-8) provides the a_i for each different output variable z_k , and Eq. (B-4) yields b , using the a_i and the average $\bar{z}$ for each output variable.

APPENDIX C

ALGORITHM LIMITATIONS

The prediction algorithms described in this report are based entirely on the empirical data from 22 set points of the Westinghouse gasifier. They represent multiple linear regression fits to the data, using the input and output variables described in Appendix A.

The input variables were chosen as those which provided the physical information which define the operation of this particular device. Since all the experiments were on the same hardware, only the quantity, temperature and type of the input streams was varied, not the geometrical parameters. In addition, all the runs were for the same coal type, Pittsburgh Seam Coal. It was assumed that the coal preparation was also the same for all runs, so no variables giving coal size, moisture content, etc. were used.

For some of the input variables, there was not much range of variation over the 22 set points. The pressure had only three levels, as shown in Table A-1. There was one run at 200 psig, 7 at 230 psig and 14 at 130 psig. The eight high pressure runs had air as the oxidizing gas, while the 14 low pressure runs used oxygen. The coal feed rate varied considerably, from 418 to 995 lb/hr, with temperatures from 102 to 213°F. The oxygen feed rate ranged from 481 to 1028 lb/hr, with temperatures from 500 to 1005°F. The air-blown runs had the higher temperatures, while the oxygen-blown runs had temperatures between 500 and 600°F. The steam flow rates varied from 133 to 797 lb/hr at temperatures in the narrow range of 411 to 553°F. The transport gas input flow rate was between 1255 and 2895 lb/hr at 198 to 425°F of weight flow weighted temperature.

The output variables chosen for prediction were those felt to be of interest in systems analysis, within the limitations of the variables provided by the Westinghouse reports. They included the bed and product gas temperatures, the net flow of product gas and gas species, and the solid carryover flow.

The range of temperatures was quite narrow, from 1696 to 1930°F for the bed and from 1362 to 1674°F for the product gas. The net product gas

flow had to be inferred from the data, as described in Appendix A, and ranged from 1048 to 6658 lb/hr. The major species were CO, CO₂ and H₂O, as well as N₂ for the air-blown cases. Their flow rates were hundreds to thousands of lb/hr. The minor species were H₂, CH₄ and H₂S, with rates of a few tens of lb/hr. The solid component of the product gas stream had flow rates from 26 to 270 lb/hr.

When the prediction algorithms are constructed for each output variable, there are several ways to assess the accuracy of the predictions. One measure is the correlation coefficient of the predictions for each variable. Let an output variable be denoted by Z, with its mean value over the 22 set points denoted by $\bar{Z}$, and the predicted value obtained from the prediction algorithms as $\hat{Z}$. Then the correlation coefficient is defined as

$$R = \left| \frac{\sum_{i=1}^{22} (Z_i - \bar{Z}) (\hat{Z}_i - \bar{Z})}{\sum_{i=1}^{22} (Z_i - \bar{Z})^2} \right|^{1/2} \quad (C-1)$$

A perfect prediction would yield a value of unity, and poor predictions would give a value near zero. The square of the correlation coefficient can be interpreted as the fraction of variation of the variable Z which is accounted for by the regression predictions.

A more detailed measure of the prediction accuracy is the percentage error in the prediction of each of the 22 values of Z, defined as

$$PE = \left[\frac{(\hat{Z} - Z)}{Z} \right] * 100 \quad (C-2)$$

In case the actual value of Z is zero, this definition is not useful of course. An alternative measure which can be used then is the

$$PE^* = \left(\frac{\hat{Z}}{\bar{Z}} \right) * 100 \quad (C-3)$$

Table C-1 shows the correlation coefficient and the percentage error for each value of the 11 output variables, using the regression coefficients presented in Table I.

The regression coefficients, on the line labeled R at the top of the table, are all above 0.85, which is generally accepted as a sign of a good fit. The percentage errors of the various predictions varies widely. The temperatures are better than 10%. The net product gas flow rate has a maximum error of 12%. The major product species flow rates have errors mostly under 10%, with a few higher values, up to 61%. The minor species have somewhat higher errors than the major species, as a whole. The variable with the largest consistent set of percentage errors is the solid carry-over flow rate.

When considering the percentage errors as a measure of the accuracy prediction equations, one important fact should be kept in mind. The percentage error depends strongly on the value of the variables being predicted. If one were to add a constant to all 22 values of one of the output variables, the regression scheme would merely add that same constant to all the predicted values. Therefore, the difference between $\hat{Z}$ and Z would not change, but since Z is larger, the percentage error would become smaller. For example, if the bed temperature were measured in °R instead of °F, their percentage error would be smaller. In general, the variables with small values have larger percentage errors, while those with larger values have smaller percentage errors, for this reason.

On the other hand, the correlation coefficient would not be affected, since $\hat{Z}$, Z and $\bar{Z}$ would all increase by the same amount, and the correlation coefficient in Eq. (C-1) depends only on differences of these quantities, which would remain unchanged.

One check on the prediction equations which can be easily made is to try to predict an output variable consisting of an arbitrary set of numbers. For that purpose, a set of 22 numbers varying between 0 and 79 were chosen at random, and a prediction algorithm for this variable was constructed. The correlation coefficient for that variable was 0.478, much lower than any

of the values in Table C-1. So we have not constructed a prediction scheme which can handle random values of an output variable. The input and output variables are related in a causal way.

The regression coefficients in Table I and the resulting predictions shown in Table C-1 were obtained by using all 22 set points. This leaves no "test data" not used in the construction of the algorithm. One way to test the validity of a prediction algorithm is to hold out some data to use for testing. However, with only 22 cases, and ten input variables, there is really not enough data to hold some out for test data. However, there is a method which can provide test data. Prediction algorithms can be constructed using only 21 set points, reserving the 22nd one for testing. In this manner, 22 different sets of prediction equations can be obtained, each set omitting one set point. Then that omitted set point can be used to test the algorithm, comparing the actual value of each output variable with the predictions made using the actual input variables.

This procedure has been followed for the present data. The resulting percentage errors are given in Table C-2. Each line gives the percentage errors for that set point, when it is used as a test case for algorithms constructed using the other 21 set points.

The percentage errors shown are somewhat larger than the ones in Table C-1, where the predictions are all for data used to construct the algorithms. However, for the most part, they are not significantly worse, except for the solid carryover flow rate, where the errors are even larger than the considerable values in Table C-1.

TABLE C-1
PERCENTAGE ERROR OF PREDICTIONS

Set Point	TB	TP	FNET	FCO	FCO2	FH2	FCH4	FH2S	FN2	FH2O	FSOL
R	.940	.870	.998	.972	.972	.915	.890	.900	.999	.950	.856
016 # 1	.91	-.13	1.03	.12	-.24	-1.96	-14.59	-14.41*	1.98	.84	-12.62
# 2	-2.26	-4.56	-5.60	-6.23	.40	-11.47	3.17	40.73*	-5.80	-31.7	-44.69
# 3	-2.20	-2.38	-2.07	-3.58	1.86	-5.01	41.01	-15.41*	-2.62	-10.93	31.56
# 4	1.26	3.11	.51	11.89	-3.17	-18.74	15.42	22.94*	-1.52	7.40	103.7
017-1# 1	2.16	1.96	4.82	4.93	-5.10	-4.06	-27.66	-44.33*	8.28	29.82	33.87
# 2	-.23	-.29	.77	-4.82	4.78	-13.52	-15.16	-46.75*	1.01	-.01	14.89
# 3	.12	3.03	1.68	-.19	5.01	15.93	12.81	24.91*	.45	16.23	-.46
# 5	.52	-.54	-.02	.75	-2.21	-1.97	-5.92	-2.91*	.78	-.53	-8.57
018-2# 4	-.35	-.53	-1.77	2.15	-9.40	.77	.03	-13.40	-2.82*	32.71	12.58
# 5	-.09	-.95	-5.02	4.40	-1.80	-.83	10.40	.239	-2.26*	-17.84	60.85
# 6A	-.01	-.97	1.42	-1.23	6.50	.05	-2.82	-.75	.66*	-4.24	-19.44
018-3# 1	-1.47	1.22	.96	-6.78	11.53	-4.24	-1.65	-6.27*	-1.79*	-11.37	-64.20
# 2	-.86	-1.90	-7.16	1.24	-6.62	-2.65	4.35	-1.24*	-5.04*	-9.52	-22.78
018-4# 1	-.05	.39	3.52	-1.96	5.75	6.10	-4.24	39.37*	1.20*	10.42	70.38
# 2	1.85	.55	.62	-.76	2.73	6.54	6.98	46.23*	-1.21*	-4.06	-7.98
# 3	1.14	1.56	12.05	5.56	10.92	26.42	28.92	203.26*	3.75*	10.97	-28.11
018-5# 1	-.42	3.65	6.31	-.11	.97	-1.27	-3.74	-37.91	8.47*	21.07	109.96
# 2	-.52	.17	-3.63	-6.81	-.98	-6.78	-5.96	-32.75	.87*	3.02	86.14
# 3A	.32	.71	-2.45	.22	-7.40	-3.51	-12.16	-33.11	.13*	60.88	63.04
# 3B	1.34	1.68	4.47	20.28	-8.66	16.80	7.16	-5.28	-.59*	9.88	59.51
# 4	-.12	1.65	1.18	2.92	2.07	19.16	31.56	85.68*	-4.50*	11.28	-33.53
022-1# 4	-1.4	-7.90	-8.53	-12.92	-2.46	-25.69	-22.76	-29.92	315.1	-16.44	-29.49

* Percentage error using Eq. (C-3) for variables whose actual value is zero.

TABLE C-2
PERCENTAGE ERROR OF TEST CASES

Set Point	TB	TP	FNET	FCO	FCO2	FH2	FCH4	FH2S	FN2	FH2O	FSOL
016 # 1	9.9	-1.4	11.3	1.3	-2.5	-21.6	-159.	147.*	21.6	9.2	-136
# 2	-4.3	-8.6	-10.5	-11.8	.80	-21.6	5.8	73.*	-10.8	-59.2	-83.9
# 3	-3.0	-3.6	-3.1	-5.7	2.7	-7.4	61.3	-22.*	-3.9	46.9	46.9
# 4	4.0	9.8	1.6	37.7	-10.1	59.6	49.1	70.*	-4.8	23.6	329.
017-1 # 1	4.7	4.3	10.4	10.8	-11.1	-8.7	-60.0	92.*	17.9	64.4	73.4
# 2	-.69	-.87	2.3	-14.8	14.7	-41.7	-46.9	138.*	3.1	-.53	46.1
# 3	.75	20.0	11.0	-1.2	-33.0	105	84.9	157.*	2.9	107.	-3.2
# 5	4.4	-4.6	-.11	6.3	-18.6	-16.6	-49.8	-2.2*	6.6	-4.3	-72.1
018-2 # 4	-.56	-.86	-2.8	3.4	-15.2	1.2	0	-21.7	-4.3*	53.0	20.6
# 5	-.19	-1.9	-9.9	8.6	-3.5	-1.8	20.5	.42	-4.2*	-35.4	122.
# 6A	-.05	7.6	11.0	-9.6	50.7	.50	-21.8	13.8	4.7*	-33.3	-153.
018-3 # 1	-2.6	2.1	1.6	-11.8	20.1	-7.3	-2.8	-10.3*	-3.0*	-19.8	-112.
# 2	-4.5	-9.7	-36.7	6.4	-33.9	-13.1	23.0	-3.1*	24.8*	-48.5	-117.
018-4 # 1	-.06	.56	5.0	-2.8	8.2	8.7	-6.1	53.5*	1.6*	14.9	100.
# 2	2.3	.68	.75	-92	3.4	8.1	-8.6	54.4*	-1.4*	-5.0	-9.9
# 3	1.4	1.9	14.9	6.9	13.5	32.7	35.8	240.*	4.4*	13.6	-34.7
018-5 # 1	-.65	5.9	10.1	-.17	1.6	-2.1	-6.1	-61.2	13.0*	33.8	178.
# 2	-.62	.20	-4.3	-8.0	-1.1	-8.0	-7.1	-38.7	1.0*	3.6	102.
# 3A	1.1	.95	-3.3	.29	-10.0	-4.7	-16.4	-44.6	0*	82.5	85.0
# 3B	1.7	2.1	5.6	25.3	-10.8	21.0	9.0	-6.6	-.71*	12.3	74.3
# 4	-.36	4.7	3.4	8.4	6.0	55.3	91.2	237.*	-12.4*	32.8	-97.0
022-1 # 4	-2.2	-12.3	13.6	-20.2	-3.9	-40.1	-35.5	-46.8	480.	-25.7	-46.0

* Percentage error using Eq. (C-3) for variables where actual value is zero.