United States
Environmental Protection
Agency

Air and Energy Engineering
Research Laboratory
Research Triangle Park NC 27711

Research and Development

EPA/600/S7-85/024 Aug. 1985

EPA

Project Summary

Coal Sampling and Analysis: Methods and Models

Alan Gleit, William Moran, and Arthur Jung

New Source Performance Standards for large coal-fired boilers and certain State Implementation Plans require operators to monitor SO_2 flue gas emissions. As an alternative to stack gas monitoring, sampling and analysis of feed coal has been proposed to estimate SO_2 emissions. This report provides information on coal sampling and analysis (CSA) techniques and procedures and presents a statistical model for estimating SO_2 emissions. In particular, this study assesses the various coal sampling techniques and equipment, the various sample preparation and analytic methods, and common practices for CSA. It describes the variables associated with the prediction of SO_2 emissions from CSA data; e.g., sulfur retention, variability, measuring errors, and auto-correlation. Finally, it presents a time series model for predicting emissions which takes into consideration the correlation of the sulfur content of the coal, the measuring errors, and the sampling procedures for coal collection. The model is used to fit 53 data sets with little evidence of non-fit.

This Project Summary was developed by EPA's Air and Energy Engineering Research Laboratory, Research Triangle Park, NC, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

The purpose of this study was to evaluate the use of CSA techniques for estimating SO_2 emissions from coal-fired boilers. Coal sampling, whether performed manually or automatically, must extract a quantity of coal much smaller than the original lot for laboratory analysis. The sample, to be representative, must have the same characteristic qualities and constituents as the entire coal lot. However, inherent variability in the coal parameters and the imprecision of the entire measurement process make sampling a rather technical subject. This report addresses the methodology, equipment, and modeling of the CSA procedures. The full report covers:

- Assessment and statistical evaluation of coal sampling techniques and equipment.
- Assessment and statistical evaluation of coal sample preparation and analytical techniques.
- Sulfur loss (not emitted as SO_2) in coal-fired utility boilers.
- Common industry practices in CSA.
- Important mathematical parameters used to describe coal sulfur variability.
- A statistical time series model for characterizing coal sample and emission data.
- The consideration of measurement errors.
- The consideration of sampler bias.
- The analysis of CSA and emission data with a time series model.

This Summary briefly discusses some of the topics covered in the full report.

Objectives of Coal Sampling

The sampling of coal, whether performed manually or automatically, must extract a quantity of coal much smaller than the original lot but with proportionately the same characteristic

qualities and quantities present in the entire lot. It has long been realized that the properties in coal are not distributed uniformly. The variability of coal makes it difficult to collect a sample that is representative of a large mass of coal. For instance, grab samples of coal from the same source may show different analytical values if tested in different laboratories or by different technicians in the same laboratory.

Besides coal's inherent variability, other factors (e.g., how the coal was handled, or how the samples were obtained) affect the collection of a representative sample. The coal may have become segregated during loading, transport, or unloading so that the particles are grouped together by size. When samples are taken from a stationary source (e.g., a coal storage pile or railroad car), it is difficult to obtain an accurate sample because the material in the center of the pile will be inaccessible when conventional sampling techniques are used. Generally, samples taken around the pile will be limited by the depth of penetration of the sampling device into the stationary source. Similarly, when samples are taken from a moving stream (e.g., a conveyor), they should be taken from the entire width of the belt to avoid biasing the sample.

To be effective, a sampling plan must employ measures to reduce the effect of segregated particles, minimize the effect of the variability of the coal properties, and identify any mechanical bias due to the sampling method. Sampling material from a conveyor or a chute through which the coal is flowing provides access to a cross-section cut of the entire stream. This cross-sectional cut will provide a characteristic sample even though the vertical distribution of material on the conveyor may be segregated by particle size.

Manual sampling bias may be reduced by the use of automatic equipment which is not dependent on human discretion for operation. These systems are generally elaborate and have been designed for a specific plant's application. Manual sampling methods can be used, but care must be taken to ensure that the sampling technique has been consistently applied.

Sampling personnel must consider the variability between discrete units of coal when attempting to collect a sample which is representative of a specific lot; e.g., a 10,000-ton lot of coal may consist of 100 discrete units or railroad cars of 100 tons each. If the variability between railroad cars is high because of differences in loading procedures and coal characteristics, a composite made up of one increment collected from every tenth car may be insufficient to represent the entire 10,000-ton lot. In this case, increasing the number of increments (or increasing the number of railroad cars sampled) will produce a composite which better characterizes the entire lot. On the other hand, if the variability is expected to be relatively low, then sampling from each car may be an extensive effort with little or no extra benefit.

Sampling Guidelines

With all these factors contributing to the inaccuracy of coal sampling, there has been concern over the reliability of coal analysis data. Inaccurate data, whether due to an error in sampling, poor analytical techniques, or some other factor, can result in data misuse. Because of this, there are many opinions concerning guidelines for sampling coal and the establishment of standard methods.

The most widely regarded standards were established by the American Society for Testing and Materials (ASTM). The ASTM issued recommended procedures for a variety of sampling situations. ASTM Method D 2234 details the minimum number and weight of increments and the amount of the gross sample needed to provide a stated level of precision.

The ASTM also evaluates conditions under which an increment is collected; e.g., stopped-belt cut, full-stream cut, part-stream cut, and stationary sampling. ASTM points out that an automatic or mechanical sampling method, without human discretion, is more precise than manual sampling. They classify each condition under which an increment is collected. The highest classification, or most precise method, is the stopped-belt cut removed by a mechanical cutter with increments spaced systematically. The stopped-belt technique allows all of the particles in a cross-section of the belt to be collected, eliminating segregation due to particle size. ASTM specifications restrict the cutter speed to 18 in./sec and indicate a minimum cutter width (size of the opening of 2-1/2 to 3 times the topsize of the coal.

ASTM also considered the ratio of maximum particle size and the large variation of ash content among coa pieces of different size. Ash content wa used because it was believed to be th most sensitive measure of variations i coal quality. ASTM specifies that no les than the minimum increment weight b collected so as to reduce the bias due t particle size and ash content.

Table 1 summarizes sampling guide lines recommended by various expert and authorities.

Sampling Equipment

Although the recommended proce dure for collecting a coal sample is witl an automatic sampler, manual sam pling is sometimes the only alternative Manual samples can be taken from a stationary source or, more reliably from a moving stream.

In stationary source sampling (e.g. from railroad cars or storage piles) manual techniques which employ shov els, buckets, probes, augers, and othe instruments can be used. The simplest and perhaps most inaccurate, means would be sampling with shovels o scoops from a stationary source. Al though discouraged, there are severa guidelines for this type of sampling: al of them recommend taking several in crements from different points and at a consistent depth. This method will no always produce a representative sam ple of the entire lot because only the uppermost particles in the storage pile or transport vehicle are accessible to the sampling device.

An auger can be used to obtain a core sample of consistent depth into the pile. The augering technique may not accu rately represent fine particles because they tend to fall out of the sampler prior to discharge into the sample collection container. Although the use of an auger allows deeper penetration into a pile than shoveling, quite often the inner portions of the pile remain unsampled.

At least one manufacturer offers slot ted sampling probes of varying lengths and diameters. These devices may be inserted into a pile but are limited to materials with a nominal topsize of less than 0.5 in.

Samples taken from a moving stream by manual techniques are more precise than samples from a stationary source, but they are subject to human error. The human factor involved in accurately re peating the cutting process adds to the sampling error. For sampling material from a falling stream, the simplest de vices include shovels, scoops, and

Table 1. Summary of Sampling Guidelines

Author	Recommended Method	Number of Increments	Size of Increments, lb	Gross Sample, lb	Recommendations
USBM/G.S. Pope	From a moving stream	Spaced systematically over entire lot	10–30	1,000	Larger increments result in more accurate sample
USBM/B.A. Landry	From a moving stream	$N = \frac{S^2}{S^2_{(N)W}} W$	All increments must be same size	Not considered	Sampling is based on variability of coal and desired precision
U.S. Steel	Full cross-section, stopped belt	One every 30 minutes	6–8	1,000	Orderly spacing of increments to represent entire lot
USBM/N.H. Snyder	Full cross-section, moving stream	50	10–30	500–1,500	Increment size dependent on coal topsize
ASTM	Full cross-section, stopped belt	Minimum of 15	2–15	Dependent on increment number and size	Increment size dependent on coal topsize
A.A. Orning	Not specified	Number and size of increments depend on variability of coal's size and properties.			
W.M. Bertholf	Not specified	Minimum number based on coal's variability	Not specified	Not specified	Proposes collection of a primary increment, which is then resampled
P. Gy	Mechanical	40 to 50, based on estimated sample size	Dependent on speed and width of cutter	Dependent on size and shape of particles	Recommends a mechanical sampler with a large cutter width

buckets. The shovel or similar device should have raised sides and capacity sufficient to hold a cross-sectional cut without overflowing.

The collection of coal samples by mechanical means is considered a much more precise and representative method than manual sampling. The ASTM bases this assumption on the fact that an automatic sampler will not select an increment on a discretionary basis. There is no human element involved.

Mechanical samplers are nearly always associated with coal conveyor belts on coal chutes. Most manufacturers of mechanical samplers have used designs based on ASTM standards. Some commercial mechanical samplers use primary, secondary, and tertiary automatic samplers (combined). The primary sampler collects a large sample which is subsequently sampled by a secondary sampler. In some cases the coal from the secondary sampler is crushed prior to division by an automatic tertiary sampler. In other cases size reduction and further subdivision of the primary or secondary sampler is performed manually. Samplers now in use include cross-stream cutters, rotating scoops, augers, and rotary arm samplers. The reliability and accuracy of different automatic systems is difficult to quantify and is a function of the installation of the equipment, the physical characteristics of the material to be sampled, and the operating techniques used by the facility personnel.

The many different approaches to mechanical sampling allow for many different applications and consideration of specific requirements. Although mechanical sampling is more precise than manual, not all mechanical methods collect equally representative samples. The design should include a means for taking a sample from the entire cross-section of the coal stream. Other important factors to consider are the speed of the cut and the size of the cutter opening.

In general, the cutter should move at a uniform speed to ensure that the entire cross-section is represented in equal proportions. The speed should also be slow enough to prevent segregation and rejection of particles due to disturbances of the coal stream. The cutter opening should be large enough, at least three times the size of the larger coal particles, to allow equal representation of all size particles.

The preferred method of sampling would be the method that obtained the most precise representative increment. Table 2 shows the order of preference for each sampling scheme discussed here.

Sample Preparation

The collection of a representative gross sample using manual or automatic methods often yields a bulk quantity of material which may weigh as much as several hundred pounds. Normally, 50 to 100 g of material are necessary to meet the analytical requirements of most coal characterization tests. The method used to reduce the gross sample to the analytical sample must maintain the integrity and representativeness of the sample while meeting the particle size and weight requirements as specified by the particular analytical methods.

The preparation of a gross sample for coal analysis requires several processing steps. Air drying is used to bring the moisture content of the sample to equilibrium with the air in the room where further preparation will take place. Sam-

Table 2. Order of Preference of Sampling Procedures and Methods

Collection Procedure	Collection Method	Order of Preference
Stopped belt cut	Mechanical, automatic system	1
Stopped belt cut	Manual sampling	2
Full stream cut	Mechanical, automatic system	3
Full stream cut	Manual sampling	4
Part stream cut	Mechanical, automatic system	5
Part stream cut	Manual sampling	6
Stationary sampling	Mechanical, automatic system	7
Stationary sampling	Manual sampling	8

ples are reduced or crushed from a nominal 3-in. topsize to minus No. 4 or No. 8 mesh used to optimize the reproducibility of subsequent processing procedures. Sample division or splitting results in a smaller quantity of material without a loss in the sample's representativeness. Another sample reduction process, pulverizing to minus 60 mesh, is generally required for specific analytical procedures. Thoroughly mixing and homogenizing the analytical sample is necessary so that the small aliquots which are taken for individual tests are representative of the sample.

Sample Analysis

Proximate and ultimate analyses are often used to characterize selected coal properties. Proximate analysis refers to the determination by prescribed methods of moisture, volatile matter, fixed carbon (by difference), and ash. Ultimate analysis includes: the determination of carbon and hydrogen in the material (as found in the gaseous products of its complete combustion); the determination of sulfur, nitrogen, and ash (in the materials as a whole); and the estimation of oxygen by difference (ASTM D121-78). Although it is common to find several different methods for measuring each parameter, it is generally through agreement between parties, such as the coal supplier and customer, that specific methods and operating conditions are selected.

The most important analyses for evaluating coal SO_2 emission are moisture, ash, total sulfur and calorific value. Standardized techniques for performing these analyses are summarized in Table 3. In addition to these standard methods, a number of automated techniques substantially reduce the analysis time and hence the costs of these analyses. A brief summary of some of these techniques is provided in the full report.

Sulfur Loss in Coal-Fired Utility Boilers

Coal feedstock is usually sampled as it is bunkered. When it is removed, it moves on to the pulverizers and then to the boilers. Sulfur is removed during pulverizing and combustion from pathways other than the emission of gaseous SO_2 in the flue gas. Hence the amount of gaseous SO_2 in the flue gas is not simply related to the measured sulfur in a CSA program. As SO_2 is regulated by EPA, guidelines need to be developed to estimate the SO_2 based on as-bunkered coal sulfur measurements. Several factors must be considered in this development: the coal characteristics (ash constituents, organic and pyritic sulfur content, and heating value), the combustion process (boiler size, firing type, and generating load), and possible fuel additives.

In general, insufficient knowledge is currently available concerning the effects and interrelationships of the various factors involved in sulfur retention in the pulverization and combustion process to allow a blanket recommendation for sulfur loss credits. However, some possibilities for giving credits for sulfur loss can be given:

a. A constant 5% rate seems a reasonable compromise between regulatory demands for conservancy and industrial demands for "reality."
b. Each utility can take a constant rate (say 5%) or can, using mass balance, prove it is entitled to more. The regulatory difficulty here is that the retention rate is so variable that past or current information may bear little relation to the future.
c. Allow no retention rate. This is the most conservative approach, forcing the sources to bear the entire burden of the retention if they wish to use CSA.
d. Allow no retention in general but allow each source, using mass balance, to prove it is entitled to one. The drawback stated in (b) applies.

CSA Practices in the Utility Industry

Several common CSA techniques are currently used by coal-fired utilities to evaluate their fuel. A brief study was conducted to determine the extent of use of these methods and any quality-control measures associated with them. This information is important to the study because it indicates how closely current industry practices match the proposed EPA Reference Method 19A requirements (48 FR 48960, October 21, 1983) and because it may show where alternative procedures need to be incorporated or approved for use somewhere within the proposed regulation.

Two major sources of information were used to evaluate CSA practices in the utility industry. The first source was a report entitled "Electric Utility Coal Sampling and Analysis Practice: A Comparison to Proposed EPA Reference Method 19A Requirements Based on Utility Responses to FERC Survey." This report tabulated and analyzed certain data collected from a survey of 190 utility plants for CSA information. The tabulations were set up so they could be compared with the proposed Reference Method requirements; therefore, the sampling method was not specifically defined beyond a group of ASTM methods. The sampling location was not specified other than "as-received" or "as-fired," and the method of sulfur analysis was not defined if any method other than the ASTM standards was used.

The other source of information was individual contacts at 24 utility plants with which Versar had previously worked. These plants were part of a nationwide sampling program and, as such, represent a variety of boiler types, coal types, and geographic locations.

Table 3. *Summary of Selected Standards for Coal Analysis*

Standard	Determination	Brief Description
ASTM D3173	*Moisture*	*A sample is air dried in an oven under controlled conditions. The moisture content is determined from the sample weight* loss.
ISO 1171[a]	*Moisture*	*A stream of preheated, oxygen-free dry nitrogen is passed through a retort containing the coal sample. Moisture from the nitrogen is collected in a weighing tube containing a desiccant. The moisture content is determined from the weight increase of the desiccant.*
ISO 388	*Moisture*	*The coal sample is distilled with toluene. The moisture is condensed from the coal/toluene mixture, and the coal moisture content is determined from the volume of condensed water.*
ASTM D3174	*Ash*	*A coal sample is completely combusted, and the ash content is determined by the weight of residue.*
ISO 1171	*Ash*	*Similar to ASTM method but with different heating rates and maximum temperature.*
ASTM D3377	*Sulfur*	*There are three alternative methods:* *a. Eschka Method—A weighed sample and Eschka mixture [2 parts MgO, 1 part $Na_2(CO)_3$] are ignited together, and the sulfur is precipitated from the resulting solution of barium sulfate ($BaSO_4$). The precipitate is filtered, ashed, and weighed.* *b. Bomb Washing Method—Sulfur is precipitated as $BaSO_4$ from oxygen-bomb calorimeter washings. The precipitate is filtered, ashed, and weighed.* *c. High Temperature Combustion Method—A sample is burned with oxygen in a tube furnace generating sulfur and chlorine oxides. The oxides are collected in absorption bottles and converted to acids. The acids are titrated to determine the equivalent amount of sulfur formed during combustion.*
ISO 351	*Sulfur*	*Similar to ASTM High Temperature Combustion Method, above.*
ASTM D3286 and ISO 1928	*Calorific Value*	*The gross calorific value is determined using an isothermal-jacket bomb calorimeter.*
ASTM D2015 and ISO 1928	*Calorific Value*	*The gross calorific value is determined using an adiabatic bomb calorimeter.*

[a]*ISO = International Organization for Standardization*

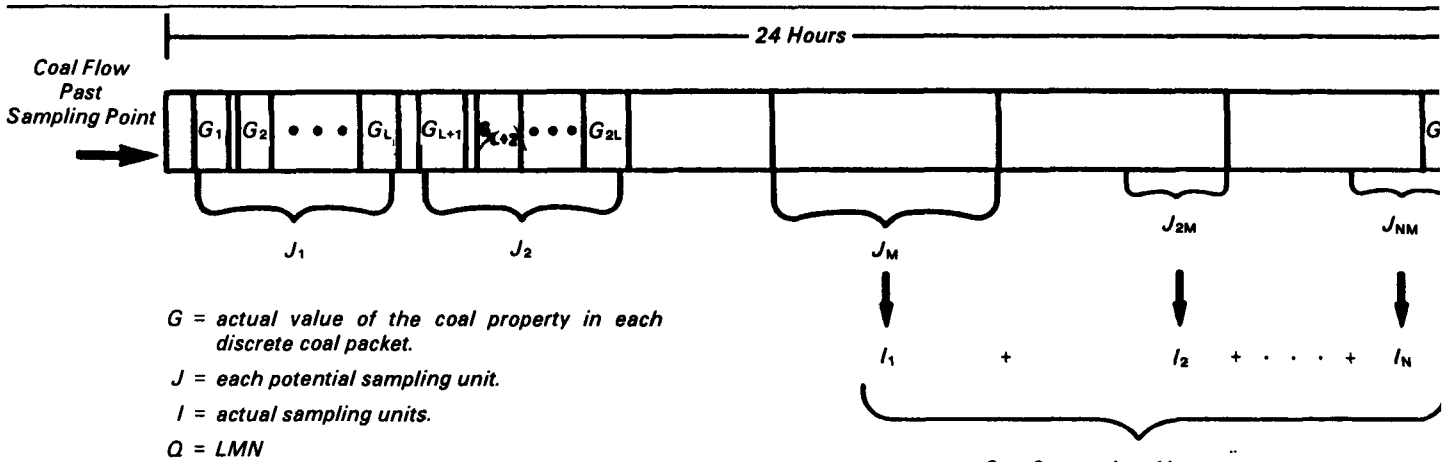
An analysis of these sources of information indicates that the following CSA procedures are common in current utility practice:

1. Sampling from conveyor belts using automated full-stream cut equipment is used by about 50% of all plants considered by Versar, and probably 67% of all the plants considered in this survey.
2. Sampling from a conveyor by taking random manual grab samples is used by 38% of the plants contacted by Versar, and probably 36% of all the plants considered.
3. Sampling coal to define as-received quality is used by 43% of all plants considered as opposed to 4% taking as-fired samples; however, no location was specified for 48% of the plants.
4. Determining heat content with an adiabatic calorimeter is used by 74% of all plants considered.
5. Analyzing for sulfur by the bomb-washing method is used by 45% of those plants which report that wet chemical procedures are used.
6. Analyzing for sulfur using automated infrared-detector-equipped analyzers may be more prevalent than use of standard wet-chemical procedures.
7. Using ASTM standard methods for the analysis of residual moisture, total moisture, and ash is universal.
8. Calibrating equipment with standard samples, duplicates, or blanks, is done most frequently on a daily, every-other-day, or weekly basis.
9. Analyzing standard samples is done for each batch (usually 5 to 10 samples).

Modeling SO_2 Emission with CSA DATA

The practical application of CSA data for estimating SO_2 emission is influenced by several factors. With current CSA techniques it is impossible to measure the sulfur and Btu content of all coal being fired in a boiler. For this reason it is important to adopt a statistical theory that will allow the modeling of a large population from a relatively limited data set. This model should provide statistical information on the coal population being burned and errors which are attributed to the CSA techniques employed.

An appropriate statistical theory for developing a model relating the CSA data to SO_2 emissions is time series analysis. In developing the CSA emission model it was assumed that the time-dependent nature of a coal or emission stream could be represented by an auto-regressive model of order one [AR(1,1)]. An auto-regressive moving average [ARMA(1,1)] model was then used to relate the CSA data to the underlying coal properties. By using the



G = actual value of the coal property in each discrete coal packet.

J = each potential sampling unit.

I = actual sampling units.

Q = LMN

Q = 1,008,000 discrete coal packets per 24 hour period.

L = # of coal packets collected in each potential increment (determined by coal sampler cutter speed).

M = difference between increment numbers for actual sample increments.

N = # of increments actually collected to form a composite.

If L = 24 and N = 35 then: M = 1,008,000/(24 x 35) = 1200.

***Figure 1.** Increment collection and composite in a CSA program.*

models to analyze CSA data it is possible to develop estimates of CSA errors and statistical properties of the time dependent coal emission stream. The statistical properties and error estimates can in turn be used to obtain the following information.

1. The average SO_2 emission level that is needed to comply with a given emission standard.
2. The probability of exceeding or complying with a specified emission standard.
3. The effect of sampling and analysis frequency on the sampling and analysis error confidence interval.

Model Description

The model simulated the statistical properties present in either a moving stream of coal or stack emissions. In either case, the coal or stack gas passing the sampling point can be thought of as being made up of Q discrete packets of homogeneous material per day (Figure 1). The computer implementation is based on the assumption that Q = 1,008,000 discrete packets pass the sampling point each 24-hour period. When applied to a CSA program the numerical results are based on the assumption that the action of the sampler cutting through the coal stream takes about 2 seconds. Therefore, roughly 24 discrete packets would be collected with each pass of the sampler.

The actual value of the coal property in each discrete coal packet is G; e.g., G_i can be the total sulfur contained in the ith discrete packet of coal. Each potential sampling unit is J. Thus, for coal sampling, every collection of 24 coal packets is a potential sampling unit.

The increments, I, are the actual sampling units, which correspond to the potential sampling units (J) that were actually collected. The collected increments are composited to form C as shown in Figure 1.

A slightly different approach is used when describing emissions data. By the time flue gases reach the stack they have become so well mixed that each unit of gas has part of the products of combustion from hundreds of the coal packets. For a continuous emission monitor (CEM) system, hourly emission values are based on four equally spaced data points. Daily averages are the average of the 24 hourly averages. In terms of our model, N = 96 increments make up the daily composites. We assume that a sampling unit J is made each minute so M = 15. Hence, L = Q/MN = 700 coal units make up each sampling unit.

A continuous bubbler (CB)* system is modeled in a similar manner. Since the CB system is continuously withdrawing well-mixed stack gases, it receives part of the products of combustion from each coal packet. Thus, L = Q = 1,008,000 and M = N = 1.

The variability associated with the coal property G is subject to correlated and random elements. This stochastic process is modeled by an autoregressive model of order one [i.e., an AR(1) process].

Since the variability associated with the discrete coal packets is modeled by an AR(1) model, J(J = J_1, J_2...J_{NM}) is modeled as an average of an AR(1) model. This latter time series model is termed an auto-regressive moving average model, ARMA(1,1). Since I is a "thinned" version of J, it too is an ARMA(1,1) model. Next, consider that

one composite sample is formed each day from the N increments:

$$C = \frac{1}{N}(I_1 + I_2 + \ldots + I_N)$$

$$= \frac{1}{N}(J_M + J_{2M} + \ldots + J_{NM}) .$$

The composites are then used to determine the quantities that are of interest; e.g., total sulfur (weight %), since the actual measurements in the laboratory represent the composites (C) and not the potential sampling units (J), nor the increments (I), nor the packets (G).

Again, C is an averaged value derived from the I values, so C is also modeled by an ARMA(1,1) model.

In all cases the daily composite value, C, is defined as the actual value of the coal or stack emission. It is provided by the model. In actuality the estimation of emissions involves using measured values (X) of coal properties as determined by the analysis of composited laboratory samples. The measured value X does not equal C because X includes some unavoidable measurement error. The measurement error consists of variations in the sampling and analysis procedures.

The ARMA(1,1) model that describes the actual coal supply (C) can be extended to a model for the measured value (X) of the daily composites. Using the parameters that describe X and the properties of the model, the parameters that describe the actual coal supply can be derived.

To apply the ARMA(1,1) model which describes C there must be some relation between its parameters and the parameters that describe X. The relationship between C and X is:

$$X = C + ME$$

where ME is the measurement error of the sampling and analysis procedures. ME is modeled as a normally distributed random variable that is independent of all past history and of C. The measurement error, ME, and the actual value of the composites, C, can be determined from the model once the parameters that estimate the value of the measured daily composites (X) are known. The auto-regressive model can also be used to provide estimates for all the parameters describing the processes G, J, and I.

Evaluation of Models

The statistical models were evaluated using 53 data sets: 19 for CSA, 25 for CEM, and 9 for Method 6B (continuous bubbler). Each of the data sets was analyzed using the appropriate ARMA(1,1) model. Each data set was then subjected to goodness-of-fit analysis to determine if the model provided an adequate description of the original data set. A few analyses were performed to determine if the models adequately describe the coal variability and measurement error phenomena.

A computer routine was developed and implemented to analyze the data using the ARMA(1,1) and other appropriate time series models. Important statistical parameters developed for each data set with the ARMA(1,1) model included the mean value, the estimated variance of the underlying coal structure [$\hat{Var}$ (C)], the estimated variance of the measurement error [$\hat{Var}$ (ME)], and relative standard deviation of the measurement error [RSD (ME)]. The variance is a measure of the variability or spread of data points around the mean value. The RSD is defined as the square root of the variance divided by the mean value, and the ratio is multiplied by 100 to convert the ratio to a percentage.

Table 4 summarizes the analysis results for the CSA data sets.

One of the longer CSA data sets evaluated was the 226 point Republic Steel clean coal, Set 2. The coal variance for this data set (0.0080) represented 65.6% of the total variance, and the measurement error variance (0.0035) represented 30.5% of the total variance. The relative standard deviation of the measurement error [RSD (ME)] was 44%.

To determine if the fitted ARMA(1,1) model adequately fit the data, three indicators were considered: residual analysis, overfit, and R^2.

The residuals are estimated by the difference between the forecast of the next data point at each time and the actual observed data point:

$$\hat{\text{residual}} = (X_{t+1} - \bar{X}) - \hat{AR}(X)(X_t - \bar{X}) .$$

If the model fits, then these residuals should form an independent sequence of random variables. If they do, then 95% of their correlations should be in the interval $\pm 1.96/\sqrt{n-1}$ and

Q = (n)(sum of squared correlations)

should be small. For the Republic set, the computer calculated the first 20 correlations for the residuals. It was found that 2 of these 20 were not in the interval ± 0.12 and that Q = 21.31. Neither of these tests shows that the ARMA(1,1) model is inadequate (but, on the other hand, neither is very compelling either).

The second method used was "overfitting" of the model. The data was fit with ARMA(1,2) and ARMA(2,1) models to see if either was significantly better. For the Republic set analyzed above, neither the ARMA(1,2) nor ARMA(2,1) model provided a fit which was significantly better than the ARMA(1,1) model.

The third indicator of fit is the overall R^2 defined by:

$$R^2 = 1 - \frac{\text{residual variance}}{\text{total variance}} =$$

"% of variance explained by model."

In the above example for the Republic set, $R^2 = 0.748$.

In general, the CSA results indicated that VAR (ME) increased as VAR (C) increased. The analysis of variance from the R&F data sets, in which the laboratory analytical samples were split for duplicate analyses, indicated that the principal component of the measurement error was the sampling and sample preparation error (the analytical error was nearly constant for three of the four sets analyzed). From these observations it is postulated that the major component of the measurement error is the sampling error and the sampling error increases with the increased variability of the underlying coal population. This is to be expected since the probability that any given sample will be representative of the mean value decreases as the variability of the underlying coal population increases. This sampling representational error can be reduced by increasing the sampling frequency.

Conclusions and Recommendations

Coal sampling and analysis (CSA) procedures cannot guarantee "correct" results. Inherent coal variability and the representativeness of the sampling and preparation procedures lead us to conclude that the resulting 50–100 g of coal per day analyzed in the laboratory may not have coal parameters exactly equal to the daily average of the coal. The laboratory results rely on fallible humans, who may introduce additional inaccura-

cies. Thus, any discussion of CSA must address the statistical issues of estimating the true emissions from the scattered sulfur data.

This study presents a theoretical model for coal sulfur data involving time series modeling. Extensive verification of the model using real coal data does not suggest any widespread lack of fit. On the contrary, the model seems to do a very credible job in fitting the data. The model can be used by EPA to evaluate the impact of various stack emission standards and/or averaging periods on the mean level of compliance coal. Alternatively, it can be used by EPA to offer guidelines to the industry as to how it could meet proposed stack emission standards.

Table of Conversion Factors

Multiply English Unit	*by*	*To Obtain SI Unit*
pound (lb)	*453.59*	*gram*
ton (2000 lb)	*0.907*	*megagram (Mg) = metric ton*
inch (in.)	*0.0254*	*meter (m)*

Tyler Screen Size Mesh Openings

Mesh Size	*mm*
14	*1.18*
24	*0.60*
48	*0.30*
100	*0.15*
200	*0.075*
270	*0.053*
325	*0.045*

Table 4. *Summary of Results for CSA*

Data Set	*Average*	*$\widehat{Var}(C)$*	*$Var(\widehat{ME})$*	*RSD(ME)*	*R^2*	*AC (Residuals) out of Range[a]*	*Q^b*	*Overfit Improvement*
1. R&F, A, ROM, Even Splits	*2.93*	*0.0352*	*0.0109*	*28*	*0.675*	*none*	*21.5*	*No*
2. R&F, A, ROM, Odd Splits	*2.92*	*0.0358*	*0.0096*	*29*	*0.705*	*3*	*29.9*	*Yes (Both)*
3. R&F, BC, ROM, Even Splits	*3.32*	*0.0610*	*0.0668*	*13*	*0.902*	*none*	*11.6*	*No*
4. R&F, BC, ROM, Odd Splits	*3.32*	*0.0578*	*0.0594*	*14*	*0.911*	*none*	*10.7*	*No*
5. R&F, A, Clean, Even Splits	*2.53*	*0.0039*	*0.00945*	*26*	*0.797*	*none*	*11.7*	*No*
6. R&F, A, Clean, Odd Splits	*2.50*	*0.0041*	*0.0087*	*27*	*0.805*	*1*	*22.2*	*Yes (1,2)*
7. R&F, BC, Clean, Even Splits	*2.82*	*0.0557*	*0.0314*	*16*	*0.852*	*none*	*11.2*	*No*
8. R&F, BC, Clean, Odd Splits	*2.83*	*0.0591*	*0.0316*	*16*	*0.869*	*none*	*11.9*	*No*
9. Republic, Clean, Set 1	*1.42*	*0.0033*	*0.0050*	*20*	*0.849*	*1*	*24.7*	*No*
10. Republic, Clean, Set 2	*1.38*	*0.0080*	*0.0035*	*23*	*0.748*	*2*	*21.3*	*No*
11. Republic, ROM, Set 1	*2.70*	*0.0643*	*0.00375*	*44*	*0.452*	*none*	*8.2*	*No*
12. Republic, ROM, Set 2	*2.69*	*negative split*				*1*	*9.8*	*Yes (both)*
13. Iowa P.S.	*0.54*	*negative split*				*1*	*24.2*	*Yes (1,2)*
14. Homer City, Unit 1 (Time 1)	*2.57*	*0.0107*	*0.0095*	*26*	*0.619*	*none*	*18.4*	*No*
15. Homer City, Unit 2, Set 1 (Time 1)	*2.49*	*negative split*				*none*	*7.3*	*No*
16. Homer City, Unit 2, Set 2 (Time 1)	*2.76*	*0.0326*	*0.0294*	*16*	*0.583*	*none*	*8.3*	*No*
17. Homer City, Unit 3, Set 1 (Time 1)	*2.66*	*negative split*				*none*	*11.9*	*No*
18. Homer City, Unit 3, Set 2 (Time 1)	*2.57*	*0.0091*	*0.0147*	*21*	*0.794*	*none*	*28.5*	*No*
19. Homer City, Unit 3, Set 3 (Time 1)	*2.47*	*0.0282*	*0.0440*	*12*	*0.968*	*none*	*15.0*	*No*

[a]*Expected number out of range is 1; i.e., 5% of 20 is 1.*
[b]*Significance levels for 10% = 26.0, for 5% = 28.9, for 1% = 34.8.*
[c]*Yes means white noise variance may be reduced by 10% by ARMA (1,2) or ARMA (2,1)*

A. Gleit, W. Moran, and A. Jung are with Versar, Inc., Springfield, VA 22151.
***James D. Kilgroe** is the EPA Project Officer (see below).*
The complete report, entitled "Coal Sampling and Analysis: Methods and Models," (Order No. PB 85-216 604/AS; Cost: $17.50, subject to change) will be available only from:
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650
The EPA Project Officer can be contacted at:
Air and Energy Engineering Research Laboratory
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

☆ U S GOVERNMENT PRINTING OFFICE: 1985 559-111/20642