Statistical Science
1995, Vol. 10, No. 1, 104–117

A Conversation with Hirotugu Akaike

David F. Findley and Emanuel Parzen

Abstract. Hirotugu Akaike was born in Fujinomiya City, Shizuoka Prefecture, Japan on the fifth of November 1927. He studied at the Naval Academy of Japan, the First Higher School and the University of Tokyo, where he earned his B.S. degree and his external Doctor of Science degree, both in mathematics.

After receiving his bachelor's degree in 1952, he was hired by the Institute of Statistical Mathematics, which had been founded eight years earlier by the Japanese government. He was Director of the institute's Fifth Division, concerned with time series analysis and control, from 1973 until 1985. When the institute was reorganized as an interuniversity research institute in 1986, he became a Professor and Director of the Department of Prediction and Control. In 1987, he became Director General of the Institute, the position from which he retired on March 31, 1994. He was also Professor and Head of the Department of Statistical Science of the Graduate University for Advanced Studies, an independent university whose departments are distributed among the 11 interuniversity research institutes, from 1988 until 1994.

He has held visiting positions at a number of universities: Princeton (1966–1967), Stanford (1967, 1979), Hawaii (1972), the University of Manchester Institute of Science and Technology (1973), Harvard (Vinton Hayes Senior Fellow in Engineering and Applied Physics, 1976), Wisconsin–Madison (Mathematics Research Center, 1982) and several Japanese universities.

His honors include two major technology prizes, each shared with one or more collaborating engineers: with Toichiro Nakagawa, he was awarded the 1972 Ishikawa Prize for modernization of production management by the Ishikawa Prize Committee of the Japan Union of Scientists and Engineers; and, with Hideo Nakamura and others, he received the 1980 Okochi Prize of the Okochi Memorial Foundation for contributions to production engineering. In 1989, he was the recipient of two of Japan's most respected culture and science awards, the Purple Ribbon Medal given by the Emperor of Japan and the Asahi Prize of the Asahi Shimbun Foundation, awards which recognize writers and artists and other citizens as well as inventors and scientists for distinguished contributions to Japanese society. He was a member of the Science Council of Japan from 1988 to 1991.

He has published more than 140 papers and several monographs and textbooks. His 1972 monograph with T. Nakagawa on the statistical analysis and control of dynamic systems has been republished in English translation (Akaike and Nakagawa, 1988). To indicate the magnitude of the impact of the methods described in this book, Professor Genshiro Kitagawa kindly provided us with a table from an article published in Japan in February 1994 listing the outputs of electric power plants in Japan that were built to be controlled by statistical models based on these methods. The table shows these plants generated approximately 12% of Japan's electrical power obtained from nonnuclear and nonhydroelectric sources.

The initial conversation, in which David Findley and Emanuel Parzen spoke with Professor Akaike, took place in May 1992 at the University of Tennessee in Knoxville during the "First U.S.–Japan Conference on

the Frontiers of Statistical Modeling: An Information Approach." Findley later obtained clarifications and amplifications of some points from Professor Akaike during visits to the Institute of Statistical Mathematics in Tokyo in March 1993 and February 1994.

EDUCATION

Findley: Hiro, Manny Parzen and I are very pleased to have this opportunity to explore some aspects of your scientific development and career. The success of your methods in applications in many fields has helped us to understand more fully what statistical science and statisticians can contribute to other fields. To begin, could you tell us something about your education and the decisions that led to your becoming a statistician and time series analyst?

Akaike: The part of my education that is closely related to my present career started perhaps when I was a student at the Naval Academy during the wartime, because I learned something about statistics there, for example, fitting a straight line by using the method of least squares. After the war I was interested in going to the University of Tokyo to learn about atomic physics, but I got the impression that research on atomic energy was being discouraged in Japan. So I planned to approach the subject indirectly through electrical engineering. But in the year when I wanted to begin, age restrictions prevented me from taking the entrance examination to the university. So I followed a general science curriculum at the First Higher School (Daiichi Kōtō Gakkō) and afterwards went to the University of Tokyo to study mathematics. At that time I was very much interested in learning how to solve nonlinear differential equations and related problems so that I could help engineers. But the mathematicians were more interested in finding conditions for the existence of solutions and similar abstract issues, so I looked outside pure mathematics and started learning probability to prepare for statistics.

I came to the Institute of Statistical Mathematics in 1952. Since I was very much interested in radio receivers and the related electronics, it was natural that I developed an interest in the subject of time series. Consequently, from the very beginning of my career, I was doing research in time series.

David F. Findley is Principal Researcher at the Statistical Research Division of the Bureau of the Census, Washington, DC 20233-9100. Emanuel Parzen is Distinguished Professor of Statistics at Texas A&M University, College Station, Texas 77843-3143.

Findley: What led you to the institute?

Akaike: I was trying to get a job with the government as a statistician. Through an interview with the National Personnel Authority I almost went to the Ministry of International Trade and Industry, but then Professor Zyoiti Suetuna suggested that I go to the institute and meet Dr. Chikio Hayashi (who later became its director general). I knew a little bit about the institute through their publications. Also, Dr. Hayashi's group was making significant contributions to the field of social surveys and sample surveys. Since I saw they were doing practical applications, I thought I could learn something there.

Parzen: At that time, was anyone at the institute doing work in time series analysis?

Akaike: Not specifically, but Dr. Masami Sugawara, who later became Director of the Research Center for Disaster Prevention, was doing an analysis of river flooding using a kind of simulator with fluid. That study was very interesting to me. Probably he was the only person working with actual time series. Of course there were people who were interested in some theoretical aspects.

Findley: Did you have any training in time series analysis before you went to the institute?

Akaike: No. I had read Cramér's book on random variables and probability distributions in a seminar with Professor Suetuna during my last year at the university and had attended probability courses given by Professor Yukiyoshi Kawada in my first and last years. That was the extent of my education related to statistics.

Findley: So you learned the techniques of time series analysis on your own, perhaps through reading some of the available literature?

Akaike: I read everything I could find. There were chapters on time series analysis in some Japanese books on statistics, and I had a chance to look at Wold's book.

FIRST EXPERIENCES WITH MODELING

Findley: In your writings, you have argued forcefully for the important role played by models in the development of statistical solutions to applied problems. Could you tell us something about your early experiences with statistical models, experiences which might have brought you to this view?

Akaike: I'm not quite sure what the main reason is for my interest in modeling, but maybe I have some instinctive tendency to treat problems through modeling because I am very engineering-oriented. In the time series literature I read then, only such things as the estimation of autocovariances were discussed. I was more interested in developing statistical models which could be manipulated. Particularly when I did an analysis of stock price series, I felt ordinary time series analysis was almost useless. You need to have very detailed knowledge about the time series itself before you apply any established technique. I was really looking for the structure of individual time series.

This was the state of mind or psychological situation in which my work on modeling began. My first opportunity to apply modeling to a practical problem came when Mr. Akinori Shimazaki of the Sericultural Experiment Station of the Ministry of Agriculture (now a professor at Shinshu University) visited the colleague I shared an office with. I overheard that he was having trouble applying ordinary control chart techniques to the process of winding filaments of silk from bunches of cocoons into a thread and onto a reel. I felt that a model which I had developed for the analysis of traffic flow in the street could be applied to this problem, and I went to Mr. Shimazaki to propose that he try this model. This was quite successful.

Findley: So modeling traffic flow was the first successful modeling experience that you had?

Akaike: In a sense. I did this just to check my idea that there must be some structure in the flow, which I treated as a very simple stationary time series of ones and zeroes: one car in the sampling interval or no cars. By assuming the independence of the lengths of time intervals between cars, I was able to derive the structure of the series, which I called a gap process (Akaike, 1956). It is a stationary renewal process in discrete time. Essentially the same structure was applicable in the silk reeling process to the dropping ends of exhausted silkworm cocoons. From the test reeling carried out to determine the proper boiling condition for the cocoons, the distribution of filament lengths and the structure of the gap process were estimated. The time series of the number of dropping ends under normal operating conditions is represented by the sum of as many gap processes as there are cocoons. The resulting process provides a basis or reference process to use for the detection of abnormalities in the actual reeling process (Akaike, 1959). Shimazaki tested and implemented this procedure with so much success that it enabled him to earn the first doctorate in agriculture awarded for work in the area of sericultural engineering by the University of Tokyo, and it eventually led to further research which changed the method of silk production in Japan. Afterward, he became Japan's leading researcher in the area of silk manufacturing and received several prizes. This important real proof of practical applicability gave me confidence in the power of models.

Findley: If I have understood you correctly, the work you did modeling traffic flow was a project you chose for yourself as an exercise in statistical modeling. This suggests that you had considerable freedom in the choice of projects you worked on.

Akaike: Definitely. I think this was a very fortunate situation.

CONTACTS WITH ENGINEERS VIA THE FREQUENCY DOMAIN

Findley: Did the institute have many contacts then with industrial researchers?

Akaike: No. When I entered the institute I was in the division mainly concerned with the analysis of social phenomena. But the director of the division, Dr. Hayashi, didn't force me to work in only that area. Then I was put in charge of a newly created section for the study of time series in another division. Afterward, I was able to develop contacts with people in industry, many of whom are now my friends. But it took almost 10 years for me to develop substantial contacts, and this happened largely by coincidence.

Actually, when I entered the institute, there were several people from outside the institute I would get together with to study statistics, at the house of a friend, Ichiro Kaneshige. After Kaneshige began to work for the Isuzu Motor Company, I was always suggesting to him that he try some statistical techniques. Then, at his request, Isuzu sent him to do research for a year at the Tokyo Institute of Technology, and he told me that the subject he had chosen was the analysis of random vibration by spectral techniques.

Findley: And you worked with him on this topic, and this led to a successful procedure for analyzing suspension systems?

Akaike: Yes. He got a prize from the Automobile Technology Association for his work on the analysis and design of a suspension system by statistical methods. This was the first practical application of a procedure for the estimation of frequency response functions of linear systems that I developed with Y. Yamanouchi of the Transportation Technical Research Institute of the Ministry of Transportation (Akaike and Yamanouchi, 1962; Akaike and Kaneshige, 1962). The original, nonstatistical approach to estimating this function requires using many sinusoidal inputs to the system at different frequencies. But this is often not practical, for example, when you are steering a car on the road! Then you need to work with the natural input, from normal driving. N. Goodman had published a statistical estimation approach in 1957, but in his paper little attention was given to the effect of the window chosen for the smoothing required for spectrum estimation. In a 1960 paper, I showed that this effect was the source of the severe underestimates of the gain function obtained by Darzell and Yamanouchi when they used Goodman's procedure, and I proposed a solution. The 1962 paper with Yamanouchi gives a full theoretical analysis.

Findley: Did you have to know anything about automotive engineering to participate in this research?

Akaike: No. But I thought I needed to have some kind of direct feeling for random vibration. So I bought a motor scooter and drove around Mount Fuji.

Findley: And that was useful?

Akaike: Yes. I could easily see how the surface of the road changes when it is used by heavy trucks, and, of course, I could see the interaction between the frequency characteristics of the suspension system and the wavy pattern of the road, because these roads are unpaved.

Findley: Do you feel that this was essential for your success with the project?

Akaike: Not in this case. But still, it's very important to have some direct feeling for the subject, as I did in the case of silk reeling, due to my father's raising silkworms.

FEEDBACK ANALYSIS REQUIRES BOTH FREQUENCY AND TIME DOMAINS

Findley: You continued to work on frequency domain estimation techniques for a while, but when you started to develop diagnostics for systems with feedback, you found it necessary to return to modeling in the time domain. This work on feedback systems was very fruitful. It ultimately led to your various FPE (final prediction error) order selection criteria, which in turn stimulated you to develop your general model selection criterion AIC (an information criterion). But this research first produced your relative spectral power feedback diagnostic and your impulse response estimates for feedback, powerful tools that are hardly known to statisticians outside Japan. The only place I have seen them in a western journal read by statisticians is in your recent partially tutorial paper with Takao Wada and others in *Computers and Mathematics with Applications* (Wada et al., 1988). So it would be useful if you could talk at some length about this work.

Akaike: We had a special research project on the practical application of frequency response function estimation techniques. I was able to get some support for this from the Ministry of Education, which funded over and oversaw the institute then as now. I asked engineers in various fields to come try our method of frequency response function estimation with their own data, which they could bring on an analog tape. I had developed a switching circuit to feed a two-variate time series into our computer from such a tape through an analog-to-digital converter.

The main results were reported in a supplement of the *Annals* of our institute in 1964. I think these reports convincingly showed the practical applicability of statistical techniques for the estimation of frequency response functions. But I recognized that there are many important problems where there is feedback from the output to the input. Then the basic assumption for the application of these techniques breaks down. I tried to find some resolution of this difficulty, but it took a long time.

I reported my first analysis of the failure of the frequency domain methods at the 1966 symposium at Madison, Wisconsin, on "Spectral analysis of time series" (Akaike, 1967). At that time I wasn't successful yet. It was early in the next year, after I finished my visit at Stanford, where you had invited me, Manny, that I finally developed an idea how to treat this problem (Akaike, 1968). I was forced to come back to the time domain because the condition of physical realizability, concerned with the fact that effects propagate forward but not backward in time, was not easy to implement with frequency domain techniques. This finally led me to the general use of multivariate autoregressive modeling for both analysis and control in this type of problem.

Findley: Did you have any practical experience with feedback?

Akaike: Yes. By the time I developed this analysis technique with multivariate time series models, I was already designing my own audio amplifier with multiple feedback, and I experienced much difficulty in designing the proper feedback.

Findley: Did you have a specific application in mind at the time you developed these techniques?

Akaike: I had already begun working with Dr. Toichiro Nakagawa from Chichibu Cement Company to find a method for the control of an industrial cement rotary kiln. The basic model I used for a feedback system includes a model for each variable of the system in terms of its own past and the contemporary and past values of the other variables, together with a driving noise. Each variable's driving noise process is presumed to follow an autoregressive model and is assumed to be uncorrelated with the noise processes of the other variables. This is a crucial assumption which makes it possible to represent the spectrum of each system variable as a sum of contributions, one from each of the noise processes. This is the source of the diagnostic you mentioned. It enables you to see which variable plays the dominant role in generating the fluctuations of another variable in a particular frequency band.

This idea has been useful in developing an understanding of the behavior of a variety of feedback systems. One successful application was made by Dr. Kohyu Fukunishi at Hitachi Ltd.'s Atomic Energy Research Laboratory. He did an analysis of the abnormal behavior of a nuclear power plant by comparing the relative spectral power contributions for normal and abnormal operation. He was able to confirm the correctness of a guess about the source of the abnormal fluctuations. His paper was published in *Nuclear Science and Engineering* (Fukunishi, 1977), and I heard that it led to his receiving an honor as the author of the best paper published in this area in the journal in the preceding three years.

In a very recent application, Wada and other medical scientists (Wada, Sato and Matsuo, 1993) were able to use the spectrum diagnostic and the impulse response functions calculated from the feedback model to solve a long-standing problem of differentiating the roles of chloride and potassium in the development of metabolic alkalosis (above-normal concentrations of bicarbonate in the blood and in other body fluids). The analysis showed a significant contribution of chloride concentration to bicarbonate concentration but none in the opposite direction.

I think this technique can be applied with little effort once you have an appropriate record of observations. You then get good insight into the system. I keep insisting that analysis before control is very important. Some people think only of control and immediately apply control methods, but it's dangerous. Unless you really know the structure of the fluctuations and introduce appropriate improvements to the system itself before implementing the control, you might not get good results. Yes, before you take your medicine, you had better adjust your physiology by adhering to regular activity! I believe that these diagnostics could also be applied to determine some of the feedback structure of economic time series.

Parzen: When you did this work on feedback and control did you have a staff to help you? Did you have assistants to do numerical work for you and things like that?

Akaike: Yes. At the time we were always helped by assistant researchers who were quite capable at doing programming and statistical calculations. I was already being helped by Miss Arahata, who has assisted me for many years.

FINAL PREDICTION ERROR ORDER SELECTION CRITERIA

Findley: It was in the course of this work that you developed two of your final prediction error criteria, FPE and MFPE, for selecting the order of scalar and vector autoregressions. What motivated this development?

Akaike: In the feedback model, a model order, meaning the maximum time lag of past values used, must be specified both for the system variables and for the noise autoregressions; that is, two orders are needed to specify the basic model. (It became one order later with the use of multivariate autoregressive models.) I suggested the use of this basic model to the people at Chichibu Cement Company, where Dr. Nakagawa was working. They got some interesting results, but were always having trouble fixing these orders. They kept asking me by telephone how to handle this problem, and it was very difficult even for me to make clear decisions.

I decided that the simpler situation is the modeling of the noise itself. This is just an ordinary stationary autoregressive model. But if you have a procedure for the choice of the model order, then you get an estimate of the power spectrum automatically. That means you don't need a statistician to get an estimate of the power spectrum. This is quite serious, you know! (laughter) So I was a little bit dubious about the idea of developing a definite procedure for the choice of model order. But I remembered from my visit to Princeton that people in the U.S. are pragmatic in the sense that if they can get a reasonable result, they think it's okay.

So rather than sticking to a strict attitude requiring precise justifications, I decided if I could produce a fairly reasonable answer, then that would be sufficient. I tried various possible criteria and did an enormous amount of numerical computation. Eventually, I recognized that the final use of the model is for prediction, and I thought that the expected value of the mean square error of prediction of an independent realization would give me a kind of basic criterion for the choice of the order. So I produced an almost unbiased estimate of this criterion. By watching the behavior of this estimated value of expected one-step-ahead-prediction-error variance in simulations first, I could see that there was a dip at the proper order and after that the output became very stable. Other criteria I had considered before this one had worked well with some particular model but not well with others. This one worked well with various choices of models, so I thought I had finally found a good criterion.

When I extended the idea to multivariate time series, I recognized that there are several measures of one-step prediction error, and I chose as basic criterion the generalized variance, the value of the determinant of the covariance matrix of the one-step-ahead prediction error. When I thought more about the estimated value of this covariance matrix, I recognized that the log of the estimated generalized variance appeared in a paper by Whittle in the formula for the asymptotic maximized log-likelihood of the Gaussian model. That was when I first recognized that there was some connection with the maximum likelihood concept. After I developed the appropriate FPE criterion, the engineers at Chichibu used it for order selection without my help.

FIRST CONTACTS OUTSIDE JAPAN

Findley: A few moments ago you mentioned that in the United States you found a more practical attitude towards statistical methods, which helped you. Could you elaborate?

Akaike: Yes. This may not be the general attitude of American statisticians, but when I visited Princeton University in 1966 and 1967 at Professor John Tukey's invitation, I attended a lecture by him. He was talking about how to handle abnormal observations in a time series, and his attitude toward this type of problem was quite pragmatic. He wasn't following any particular model of outlier but he was developing reasonable procedures to eliminate irregular observations. Even though this was not completely theory-based, it was very reasonable. My position at the time was either the rather theoretical one that unless we have a definite reason to eliminate these observations, we shouldn't eliminate them, or an extremely practical one. I remember saying when asked how to handle this type of problem that you only have to use a digital-to-analog converter with an audio amplifier and listen to the sound: if there are abnormal observations you will hear a clicking sound. I recall that John Hartigan enjoyed this answer. So I was alternating between the two extremes, very practical or very strict, but Professor Tukey was in-between, developing reasonable analyses and also producing very reasonable ideas about how to handle such problems. I thought maybe this is a good attitude. Consequently, during my later work on autoregressive model selection, I thought that even if I couldn't really get a definitive solution to the problem of the choice of order, if I could find a solution which produced a reasonable answer in many situations, then that would be sufficient, even if I couldn't prove any kind of optimality. This attitude helped me very much.

Parzen: This story really begins in 1965. There was an event called the U.S.–Japan Joint Seminar on engineering applications of stochastic processes, which I want to ask about next. But first I want to mention that, as a consequence of that seminar, I invited you to come to the United States. Then John Tukey heard about your planned visit to the U.S. and asked you to spend the first part of the trip in Princeton, which you did. Now let's go back to the U.S.–Japan Joint Seminar. Could you tell us something about the origin of that? What I know is that Frank Kozin received a grant to conduct a U.S.–Japan joint seminar but didn't know what topic to select. Will Gersch is a friend of Frank Kozin, and Will had read a review I wrote in *Mathematical Reviews* of your work on spectral estimation. On the basis of that, they decided you were the person in Japan to contact to organize the Japanese side of the seminar, and you were contacted by Frank Kozin. Take it from there.

Akaike: I received a letter from Frank Kozin. He was visiting Taiwan at the time and asked me to be something like the chairman for the seminar. But since I was too junior, I had to find someone else who could do this. I talked to my friends in engineering and they decided to ask Professor Takashi Isobe of the University of Tokyo to organize this activity. He and some other engineers proposed this seminar to the Japan Society for the Promotion of Science, JSPS. It was accepted, and the U.S.–Japan Joint Seminar took place in 1965. Manny came, and Professor Drennick from Brooklyn Polytechnic Institute and Professor Ho from Harvard.

This was the beginning of my contact with the outside world. Before that time I was only interested in developing statistical techniques based on the demand within Japan, because unless you have a real problem close at hand, you cannot develop any good ideas.

Parzen: When I invited you to come to the United States, what was your feeling about that?

Akaike: Well, we had already met and I had found that there were many good colleagues in the time series area. I felt this was an opportunity for me, because in Japan I was rather isolated. There were not many time series analysts, particularly time series analysts interested in engineering, an extremely difficult situation for me. Sometimes it is very stressful when you are exploring a new area and you feel that you are alone.

STATISTICAL CONTROL OF COMPLEX INDUSTRIAL PROCESSES

Findley: Your feedback diagnostics and FPE criteria evolved as powerful new tools around which you and Dr. Nakagawa developed a statistical-model-based approach for controlling industrial processes. This approach has been successfully applied to many processes too complex for classical, differential-equation-based control procedures. How did this work with Dr. Nakagawa begin?

Akaike: We met around 1961 at a meeting of the Society of Instrument and Control Engineers where I insisted to the speaker that he did not have enough data to resolve different peaks in a spectrum the way he wanted to. Doctor Nakagawa was sitting near me and decided to ask me to help him to clarify the characteristics of a measurement device he had developed which used a kind of nonlinear noise reduction technique. He then came to Tokyo from Chichibu to talk with me. After we began to work together I asked him to provide me with a problem related to control. I wanted very much to demonstrate the usefulness of statistical methods for controlling a modern industrial process. But it took several years until he brought me the record of operation of a cement kiln. The paper on feedback I presented in Wisconsin in 1966 was a report about our initial work with such data. This work finally gave me the complete idea of a multivariate system with feedback, which we implemented into a linear quadratic Gaussian control approach.

Findley: This is described in your 1972 book with Dr. Nakagawa, which is now available in English (Akaike and Nakagawa, 1988).

Akaike: Yes. By that time we had a complete success proving the effectiveness of this approach, because the control realized by this approach showed a significant reduction of the power spectrum, compared with human operator control, in the very low frequency range. This means that long-lasting movements away from the desired state of the system were reduced.

Findley: I believe that your work with Dr. Nakagawa is of historical significance for statistics. Some of the applications of your book's methods and software have been extraordinary. At the conference this week, we heard another of your collaborators, Dr. Nakamura, mention that already nineteen 500-megawatt and larger thermoelectric power plants, 16 in Japan, one in Canada and two in China, have been built to be controlled by such statistical models, with more plants still under construction in Japan.

Akaike: I was extremely pleased when I saw the operation of the first power plant that used a controller designed by this technique, which was built by the Kyushu Electric Power Company, the employer of Dr. Nakamura at that time. The plant engineer there told me that the quality of control achieved was very good, and they were quite appreciative of the controller's performance. I was particularly delighted to learn, after the plant had been in operation for about a year and they had to perform some maintenance work, that even after this work the controller still performed quite well without much adjustment. This meant that the whole system was quite robust. I think that this point was also mentioned by Dr. Nakamura in his talk (Nakamura, 1994). You could very easily see that when the control was on, the whole system became stable, and the plant's ability to track the changing electricity load demand was increased.

Findley: Doctor Nakamura also mentioned that the controller is easy enough to use and so effective that someone can successfully manage the operation of the plant without having a detailed engineering knowledge of the plant. Doctor Nakamura feels that this property of the controller insures that it can be used for many years.

Akaike: Doctor Nakamura spent perhaps four or five years just to implement this type of control on a power plant. He even refused a promotion in order to be able to continue the research on this.

Findley: In 1972, you and Dr. Nakagawa were honored for your joint work with the Ishikawa prize given by the Ishikawa Prize Committee of the Japanese Union of Scientists and Engineers. Could you tell us about this prize?

Akaike: I think the prize is normally given to a research group, or sometimes to institutions, who develop significant contributions to the modernization of production management. I believe that our work on the control of cement kiln production and particularly its general formulation for other possible applications was considered to be a significant contribution in this area.

Findley: Were you the first statistician to receive this honor?

Akaike: I'm not sure; but until that time not many individuals received the prize. Of course there were two of us, Dr. Nakagawa and myself.

Findley: And in 1980, Dr. Nakamura and you were awarded the Okochi prize of the Okochi Memorial Foundation for the successful work on power plant control described in your joint paper in *Automatica* (Nakamura and Akaike, 1981). What kind of prize is this?

Akaike: Someone connected to the awarding of this prize told me that it is something like a Japan Olympic medal for production engineers. So, in this case, I was considered an engineer.

THE PATH TO AIC

Findley: How were you led to discover AIC?

Akaike: I was interested in the factor analysis problem. I don't know why, but maybe this was because of some numerical aspect of the model. Factor analysis requires numerical optimization of the likelihood function to get the solution, and I have always been interested in numerical procedures. I have one paper in this area.

Parzen: In what year?

Akaike: In 1959.

Parzen: When you came to Stanford, Professor Forsythe knew your name very well. Is this the paper that he knew about?

Akaike: Yes. The paper gave a mathematical proof of the convergence of the error term of a gradient "hill climbing" optimization method, irrespective of the dimension of the function's argument, into a two-dimensional region, a peculiar phenomenon. Professor Forsythe had a paper in the *Pacific Journal of Mathematics* verifying this behavior for the special case of a three-dimensional optimization and commenting that even this special case required a very complicated mathematical treatment. I got interested, and one day the idea came to me that this method could be interpreted as a transformation of a probability distribution on the eigenvectors of the quadratic form associated with the locally quadratic approximation to the function. This is the kind of approximation used in my derivation of AIC also.

While I was thinking about factor analysis, I got an invitation to the Second International Symposium on Information Theory from Dr. Tsybakov, who was in Moscow. I recognized that there was a similarity between the choice of the number of factors in factor analysis and the determination of the order of an autoregression; but I couldn't really find any definite connection between the two. For example, in the case of autoregressive models, the idea of prediction is clear. You just compute the predicted value of the next time point. But what are you doing when you fit the factor analysis model to data? What is the prediction in that case? I wasn't sure about this analogy, and since there was a time limit for the submission of the manuscript to the symposium, my attention was split by two separate efforts: preparation of a manuscript for the symposium concerning my experience with autoregressive models for spectral estimation; and finding a good concept for the similarity between the autoregressive modeling and factor analysis problems. I was under very much pressure, psychological pressure, so during the nighttime I often woke up thinking and during the daytime I was almost sleeping.

One day I recognized that the factor analysis people are maximizing the likelihood and this maximization is trying to get a good distributional model for the purpose of prediction. However, in this case the prediction is not a value, but is the fitted distribution itself, which is applied to understand the next similar observation. For the next similar problem you use the present model, and the accuracy criterion for this prediction is given by the log-likelihood function. Then once I got this far, it was just one step to recognize that the expected log-likelihood is related to the Kullback information. This idea came when I was standing on the train from my home to the institute. I still have the page of my notebook where I wrote down one or two lines to explain this. That was the solution.

Findley: This was in 1971?

Akaike: In March of 1971. This experience is described in the historical note that the Institute for Scientific Information asked me to produce (This Week's Citation Classic, 1981).

Parzen: Let's go over those references at this point. I see in your vita that you have two short papers in the proceedings of the Fifth Hawaii International Conference on System Sciences, one entitled "Uses of an information theoretic quantity for statistical model identification" and the other "Automatic data structure search by the maximum likelihood." So you announced some of these ideas in Hawaii?

Akaike: Yes. Actually I first announced the idea in the annual meeting of Japan Statistical Society in July 1971. Then I went to Armenia to the Second International Symposium on Information Theory in September and gave a talk on this subject. Professor Tom Kailath from Stanford was there, and he became interested in this idea and asked me to contribute to a special volume on time series analysis of the *IEEE Transactions on Automatic Control* (Akaike, 1974a), where you also have a paper, Manny. After coming back from the symposium, I visited Hawaii for a year, so I prepared a talk on this subject and presented the results of several applications there. The paper that I presented at Tsahkadsor in Armenia was eventually published in 1973, so there was a delay in the publication.

Parzen: This paper was previously very inaccessible but is now available with an introduction by Jan de Leeuw in the first volume of *Breakthroughs in Statistics* (Johnson and Kotz, 1991, pages 599–624). I should also point out that the *IEEE*

Transactions paper was listed as one of the most frequently cited papers in the area of engineering technology and applied sciences, and, of course, it's cited and applied in almost every area of physical and social science. How did you become aware of all the different fields in which people were applying AIC?

Akaike: I received papers from many people reporting the results of their applications of AIC in various areas. After I published the paper in the *IEEE Transactions*, I also wrote a paper for the popular Japanese magazine *Mathematical Science*. This explained how the idea was developed and how one can apply it to practical problems. These two articles were the source for the spread of the criterion. People developed their own applications just from reading these papers. The idea is so simple.

Findley: From looking at citation indexes, it appears that the use of AIC by statisticians developed more slowly than its use by other researchers. How did you feel about the negative reactions of some statisticians to AIC?

Akaike: My main concern was the fact that there were so few statisticians with sufficient experience in handling real problems. Lacking such experience, they could not check the validity of the basic idea of AIC on their own problems. Many had such a dogmatic attitude that they did not even question the use of a maximum likelihood estimate when nothing was known about the "true" form of the distribution, yet they criticized the use of a minimum AIC estimate. People with serious statistical problems, like applied engineers, could easily appreciate the contribution of AIC simply by getting useful answers to problems which could not be handled by a conventional statistical approach.

Findley: The first paper I know of in a statistical journal to use AIC is the one Dick Jones published in 1975 in the *Journal of Statistical Computation and Simulation*. This paper is concerned with an application to biostatistics (Jones, 1975a). He also applied AIC to meteorology (Jones, 1975b). He must have learned about it from you during your visit to Hawaii.

Akaike: Yes, in the first version of his paper, he called it something like ASC for "Akaike selection criterion." I thought this was not very appropriate. I just used IC, information criterion, in the paper on Markovian representations of time series (Akaike, 1974b), but when I was preparing my manuscript for the *IEEE Transactions* my assistant Miss Arahata was doing some programming for me. I asked her to calculate some values of the IC criterion. She was programming in FORTRAN and needed to put a different letter at the beginning of IC since it has a noninteger value, so she put the A in front of it. I thought this might be a good idea as it still suggested an information criterion. So in the paper for the *IEEE Transactions* I used AIC to denote this criterion, and also suggested the name minimum AIC estimate, MAICE. Of course, I was aware that there would be a succession of criteria, AIC, BIC, . . . , and minimum BIC would be MBICE, etc.

OTHER CRITERIA

Findley: Around 1976, you developed a criterion for regressor selection you called BIC which gives consistent estimates of the correct regressor in overparameterized linear regression situations.

Akaike: That BIC criterion, which I presented in Dayton, Ohio, in 1976 (Akaike, 1977b), was derived by a Bayesian argument with Gaussian likelihoods. For regression, it is asymptotically equivalent to the criterion independently obtained by Gideon Schwarz. He used a Bayesian argument in the more general situation of models from a Koopman–Darmois exponential family to derive his criterion, in which the term added to minus two times the log maximum likelihood is the number of parameters multiplied by the logarithm of the sample size, instead of by the number 2 as in AIC. Earlier, I had considered a modification of FPE to obtain a consistent estimator of the order of a finite-order autoregression, the consistency property that Schwarz's criterion has (Akaike, 1970).

Parzen: What about Mallow's C_P? Could you discuss the connection with C_P?

Akaike: I must confess that I was not a good reader of journals, so I didn't know at first about Mallow's C_P criterion. Only after I wrote the original paper on AIC and was preparing the manuscript for the *IEEE Transactions* did I suddenly realize that there were two papers whose ideas were very close to my own. Mallow's 1973 paper was one, and the other was a 1966 paper by Davisson (Davisson, 1966). Davisson discussed the problem of order selection, but proposed no definite procedure. And, of course, Mallow's criterion is just for scalar regression problems.

Parzen: Davisson was the first person to give the formula for the effect of parameter estimation on one-step prediction mean square error.

Akaike: Yes. Since I was only interested in practical applications, when forced to do some theoretical thinking I usually developed my own ideas from scratch. Of course, there are similarities or connections with the ideas in these two papers, but the idea of using log-likelihood as a general criterion was quite new.

STATE SPACE AND SOFTWARE

Parzen: I'd like to bring in one other visit that seems to have been significant. In 1973, just after the development of AIC, you visited the University of Manchester Institute of Science and Technology, invited by Professor Maurice Priestley, and you wrote two very important papers there. In them you used state space ideas to solve the identifiability problem for representing general vector ARMA processes, and you developed a minimal representation. These papers are cited in leading textbooks and monographs. Please talk about this period.

Akaike: Actually, Professor Priestley and his group were interested in my work on feedback system analysis. He already had a paper in *Automatica* on this subject. When I visited UMIST, they were surprised to find me working in the time domain. They expected a frequency domain person. While I was in Honolulu, I had almost completed the basic framework for the state space Markovian representation of time series based on the concept of predictor space. I had already finished my work on the canonical correlation basis interpretation, but I had to write up the final results for the multivariate ARMA model. When I was in Manchester, I produced two technical reports, one on Markovian representations of stationary time series and the other on stochastic realizations. These were motivated by an earlier contact with Professor Rudolf Kalman (Akaike, 1977c).

Findley: Your TIMSAC 74 programs implementing these procedures were the first widely distributed programs for multivariate autoregressive-moving-average modeling. It must be said that the large number of programs published by you and your collaborators is a very impressive body of research software. There are the four TIMSAC, or "time series analysis and control," packages, and there are programs on other topics such as density estimation and contingency table analysis. The University of Tulsa, where I once taught, has distributed over 700 copies of the TIMSAC software. The development of this software clearly took a great deal of your time.

Akaike: Right. It also involved many younger members of the institute as coauthors. The programs were developed initially to check the applicability of our ideas to actual problems. I recognized that such programs might be useful to people who read the papers describing our new methods and wished to try them. Finally, I decided to distribute them to as many people as possible, because I thought that through this process people can understand how powerful statistics is.

There is another area of research and software development at the institute which has been very fruitful. That is point process modeling. This was stimulated by a visit by Professor David Vere-Jones. I had asked the Japan Society for the Promotion of Science to invite him. This led to contacts between the institute and researchers in the area of seismology and to his collaboration with Yosihiko Ogata, who has become the institute's main researcher in point-process modeling and its application to earthquake modeling and prediction. Professor Vere-Jones was giving a series of lectures on point processes at the institute when I came back from Harvard, which Ogata and Tohru Ozaki were attending. I asked them what is the most basic component of point process models, and they answered that it is intensity function, which describes the intensity of occurrence of an event at a particular time point. I then encouraged them to develop and fit intensity function models for earthquake data. This has led to very interesting modeling work for various kinds of earthquakes and for their distribution over the Earth's surface. Ogata has concentrated mainly on earthquake occurrences under or near the Japan mainland. He has used smoothness priors to obtain a smooth spatial image of the intensity distribution.

BAYESIAN MODELING

Findley: You have published more than a dozen papers concerned with Bayesian modeling and Bayesian philosophical issues, and most of your modeling examples and applications use the smoothness priors you just mentioned. Please tell us how your Bayesian modeling ideas developed.

Akaike: The idea of using Bayesian modeling was originally motivated by the desire to use the full information that a set of competing models can provide. I also wanted, in this way, to overcome the inadmissibility problems identified so clearly by Professor Stanley Sclove in situations in which a criterion such as AIC is used, for example, to select the order of a polynomial. This led to the discussion of combining models given in my 1979 *Biometrika* paper (Akaike, 1979). When I visited Harvard, I concentrated on Bayesian statistics. The most bothersome aspect was the split between the objectivist and subjectivist statisticians. If subjective information is so different from objective information, how can you combine both kinds? I spent a lot of time researching the historical development of this type of discussion and eventually ended up with the resolution that a Bayesian model is just another type of statistical model for extracting the information provided by the data, in order to obtain a good distribution applicable to a future observation. Then everything became quite transparent. Since this is just a statistical model, you can use the concept of likelihood, and also expected log-likelihood, as the basic criterion for the evaluation of the model. So, if there is an undetermined hyperparameter in the prior, for example, then you can adjust it by the method of maximum likelihood applied to the marginal density. As I discussed in my paper on the objective use of Bayesian models (Akaike, 1977a), this is formally the same as I. J. Good's type II maximum likelihood procedure. But he, by contrast, regards the procedure as a compromise between Bayesian and frequentist approaches with no particular justification.

Once I got this far, it was just one more step to the idea that we need a systematic approach to the construction of priors. It seemed to me it is only our lack of understanding of the nature of the prior that impedes our developing ideas for practical applications of Bayesian models. I was familiar with procedures for ill-posed problems, as described in Tikhonov's paper, for example (Tikhonov, 1965). Usually, somewhat artificial smoothness constraints are introduced. In the seasonal adjustment situation I was considering, this led to constrained least squares. I put an additive quadratic term into an exponent with a minus sign in front and, with the proper normalizing constant, this became a Gaussian prior density. Estimation of the variance ratios in this prior by maximum likelihood led immediately to a seasonal adjustment procedure whose simplicity seemed remarkable to me since it's just a minor modification of generalized least squares. The first output, for an artificial time series published by the Economic Planning Agency, was very encouraging. So I brought this to the First International Bayesian Meeting in Valencia, Spain. That was the beginning of my practical application of Bayesian models. In the paper I presented there (Akaike, 1980), I also showed how parameters for Stein-type shrinkage estimators, ridge regression and Shiller's smoothness-prior-based distributed-lag estimator can be determined by this approach, as well as O'Hagan's localized regression. Of course, in applications, you might need some criterion like the ABIC (A Bayesian information criterion) that I obtained for these models in order to compare competing models and priors.

I don't say that a smoothness prior is necessarily reasonable as a prior. Only that it is a useful prior if it helps you construct a model that extracts information of interest to you from the data.

Findley: Smoothness prior models with hyperparameters estimated by maximum likelihood have been used in a variety of applications by Kitagawa and Gersch, who are writing a book on this topic, and by others, particularly Andrew Harvey and his collaborators, with what they call structural time series models, which seem to have been stimulated to an extent by a Kitagawa and Gersch paper they reference.

Akaike: Actually those later models you mention are based on state space models, which are usually a kind of product of the imagination. Once you have the Bayesian interpretation, the motivation for such models becomes very clear.

Findley: We find your BAYSEA smoothness prior seasonal adjustment program useful at the Census Bureau, especially with time series which are too short for other seasonal adjustment methods. How did you become interested in seasonal adjustment?

Akaike: Because I have expertise in spectral analysis, people in the Economic Planning Agency thought I would be an appropriate person to comment on an extensive study they did of seasonal adjustment procedures. Actually the person in charge of producing this report was attending a course given by our institute for the public. I think I was talking on time series analysis, particularly frequency domain analysis, at that time in the 1960's, so I only gave comments from the point of view of the frequency domain properties of the filters used for seasonal adjustments.

Findley: This was a course offered by your institute for any person from the public who wished to attend?

Akaike: Right. My contact with Dr. Nakamura came through a course of this type.

PREDICTION AND ENTROPY

Parzen: You have given a very rich summary of your ideas on prediction and entropy in a paper with this title in the 1985 ISI Centenary volume *A Celebration of Statistics*, edited by Atkinson and Fienberg and published by Springer-Verlag. Please give us a quick introduction to what you think the importance is to statistics of the ideas of entropy and prediction.

Akaike: Unless we have some reasonable prediction for the future, we cannot choose appropriate actions for the present. So this uncertainty about the future is the source, I guess, for our activities directed toward statistical understanding and analysis of data. In that sense, the predictive point of view is a prototypical point of view to explain the basic activity of statistical analysis. Within statistics, the entropy criterion is quite general and also quite natural. Of course, I don't have any mathematical proof that this is a uniquely natural criterion, but if you look through the history of statistics, you can see, as I discussed in this paper, that when people were coming close to the concept of entropy, they were producing very successful results and when they were far away from this concept, they were not so productive, at least from my point of view. So this is some kind of historical evidence for the productivity of the concept of information or entropy.

Findley: The cost of the entropy or information-theoretic point of view seems to be that it's necessary to construct a model from which a likelihood function can be obtained. If you can do that, then you have criteria with which you can compare competing models and this is, of course, very important. Science is, in part, an arena in which models compete, and the comparison of models is fundamental. But a significant part of the current development of statistics is focused on nonparametric methods which are often not very closely related to likelihood functions. Do you see any possibilities for modifying your conceptual framework so that it can also be used with more nonparametric models?

Akaike: In the past history of science people were always producing results by using some parameterization. For example, Newton's work on mechanics. So this is a very popular and useful point of view. If we can find a very powerful procedure which does not use this kind of structure, then, of course, I can easily change my point of view any time. I am, in that sense, very flexible. When I was young, I was interested in things like nonparametric or distribution-free statistics and the nonparametric approach. But when I looked through the development of science, particularly natural science, I found that many significant developments were related to the putting forward of hypotheses which clearly depend on parameters. So I thought perhaps this is the right way. It's just a subjective judgment.

Also, through the development of specific, appropriate types of models (and by insisting on their development) the accumulation of experience can be accomplished very quickly, and this accumulation becomes a kind of common resource for future development. Otherwise someone develops some arbitrary idea which cannot be used by anyone else. This communicability is very important. So is the portability and the compact parameterization when you are implementing the model for something like on-line control.

FISHER, INDUCTION AND ABDUCTION

Findley: In the first paper you presented at this conference (Akaike, 1994), you compared your views of likelihood with those of R. A. Fisher. It would be interesting for our readers if you could summarize your understanding of the differences.

Akaike: I didn't have much time to study historical developments, but when I read through some of the work of Fisher I recognized that his definition of likelihood was for a fixed model, a distribution whose functional form was given, known, and only the parameters were unknown. As I began to do autoregressive modeling, especially multivariate modeling, I became very aware of the problem of comparing different models. Also I had an impression that the division of theory into estimation and testing was somewhat unproductive in the area of practical applications. If you come to the choice of one particular model from among several alternatives by the conventional approach, you probably have to depend on some kind of test procedure rather than estimation. But to me the problem of determining the best model for the data from a collection of models was quite a natural extension of estimation.

When I was preparing my manuscript for this conference, I remembered a paper by the American scientist and philosopher Charles Saunders Peirce on the topic of induction and "abduction," meaning the generation of hypothesis (Peirce, 1955). As I reread it, I saw clearly that in scientific explorations it is the generation and comparison of hypotheses which is most important. Then it became quite clear to me why people outside of statistics showed such interest in applying AIC as soon as it was published. Their main interest was developing hypotheses based on data. So in that sense, model selection is more closely connected to what Peirce calls abductive logic. Fisher emphasized the use of likelihood in the inductive phase of inference, obtaining information about the population from the available sample under the assumption of a hypothesis, most specifically information about the parameters of the model that determines the likelihood (Fisher, 1935). In actual situations, however, you have to develop various possible hypotheses and compare them based on observations; AIC is directed toward comparison of different models. So I recognized that Fisher and I have different ways of interpreting the likelihood function. But as everyone knows, Fisher had very deep experience and insight into practical aspects of statistics. From that common experience, many of my attitudes are similar to Fisher's.

Findley: Some of your experience as a time series analyst is a kind of experience that Fisher did not have. For example, he did not work with complex dynamic physical systems as far as I know. Do you think that there is a special perspective that comes from time series analysis which has helped you find new approaches to some fundamental questions of statistics?

Akaike: If the models are clearly specified by your own past experience, then there's no need to develop a particular selection criterion or evaluation criterion. However, in time series analysis, there are many applications of linear system theory which are in a sense quite nonparametric. Of course the models are parametric, but, through different choices of the model order or dimension of the system, you get broad flexibility from the models, just like polynomial fitting. So this is a typical situation where you have to make some kind of decision, where you're going to choose one model in a practical application. I think this forced me in the direction of model selection, so I was lucky. Originally I was not very interested in this type of model. It's too general. If it's too general you cannot easily bring your own ideas or insight into the modeling. But it had very wide applicability, so I changed my mind.

DIRECTOR GENERAL OF THE INSTITUTE WITH NEW RESPONSIBILITY FOR GRADUATE STATISTICAL EDUCATION

Findley: You became Director General of the Institute of Statistical Mathematics in 1987, just after it changed from being almost exclusively a research institute to also being an institution of graduate education, the first nonuniversity institution in Japan to grant a Ph.D. in statistics. Could you please share with us some of your ideas concerning the training of statisticians?

Akaike: Statistics is a very difficult subject in the sense that it is essentially related to information, and information has no physical form. This means statistics is related to a subject lacking form, so it's very difficult to explain to society how it is important. From that viewpoint, I think the only solution is to get people from various disciplines to obtain some training in statistics and bring the knowledge back to their own fields. We need to train some professional statisticians, of course, but a very important part of our activity and mission should be directed toward the dissemination of statistical knowledge to other disciplines.

Our applicants come with a Master's degree, or equivalent work experience, from engineering and mathematics and medicine and physics and so forth, so we have to provide some basic training in statistics. But they can start their research in a particular area from the very beginning if they wish. The staff of our institute has such a broad spectrum of backgrounds and interests that we can accommodate the interests of almost any applicant. Some already have previous contacts with the institute. This year two students finished their Ph.D.'s. One got a job at the University of Tokyo in the Department of Mathematical Engineering and Information Physics. Another student, a woman, got a job at the Central Research Laboratory of Hitachi, Ltd. So they both have very respectable places where they can continue their research. So I think we are very fortunate.

Findley: In other words, your students will not receive general training in statistics as much as training that's relevant to the area of their interest. And they will not go back to a university necessarily, certainly not to a statistics department as they might in the U.S.A., because there aren't any statistics departments in Japan. Rather they will go back to departments which have some connection with the background that they had, or to departments or companies that value the background that they've gotten from your institute.

Akaike: This is exactly what we hoped for and it seems that our hope is being realized.

Findley: So you are training statistical scientists rather than statisticians! (Laughter.)

Akaike: Maybe that's a good expression. Yes, of course.

THE INFLUENCE OF THE SECOND WORLD WAR

Findley: Before we close this interview, could you tell us if there is some important way in which the experiences you had during the Second World War influenced your career.

Akaike: I think there are actually two aspects, one quite personal. I developed an interest in mathematics because of an uncle who was killed during the war. He was a Navy pilot and was interested in mathematics. He sometimes talked to me about calculus and other topics in mathematics. The other aspect is that after the end of the war, as the Naval Academy was being closed, the person in charge of the academy stressed the importance of our role in postwar society, that when Japan was recovering from this damage we must do our best to rebuild the country.

I think this was accepted by students at that time without any question. After I started my study of mathematics, eventually I thought I could serve society or help people more directly by doing statistics. Later, when my friends from the Naval Academy would meet occasionally, there was a spread of political attitudes from right to left, but we would just get together and discuss each one's opinion and develop understanding and there was no antagonistic feeling there. They just wanted to be of service to the society I guess. This was a kind of generation-dependent attitude that maybe the people who are students today have moved away from and might call an obsolete attitude. But still I think it gave us a quite stable feeling and strength. There is no serious split among us, because we have some common concern for the society that we can discuss with each other. Maybe this explanation is quite incomplete.

Parzen: I am very happy that I had the opportunity to meet you in 1965 and that we have been in contact all these years. I never thought of us as representing the U.S. and Japan, just two good friends who are working together. I think this is important. International cooperation is based on the personal friendship of the scientists, and we learn from each other in many ways besides our common interest in science. Thank you for this conversation.

Akaike: Thank you. Maybe I have to add just one more point, a point we discussed partially. One reason I could develop this kind of contact in my own research is the type of protection provided by the institute in maintaining my complete freedom of choice of subject and way of developing research. I think this is very important and sometimes very difficult to maintain, this freedom of research, without any concern about promotion and so forth. This is a very unique environment.

Findley: I also want to thank you for this conversation.

Akaike: Thank you very much. Thank you.

REFERENCES

AKAIKE, H. (1956). On a zero–one process and some of its applications. *Ann. Inst. Statist. Math.* **8** 87–94.

AKAIKE, H. (1959). On the statistical control of the gap process. *Ann. Inst. Statist. Math.* **10** 233–259.

AKAIKE, H. (1967). Some problems in the application of the cross spectral method. In *Spectral Analysis of Time Series* (B. Harris, ed.) 81–107. Wiley, New York.

AKAIKE, H. (1968). On the use of a linear model for the identification of feedback systems. *Ann. Inst. Statist. Math.* **20** 425–429.

AKAIKE, H. (1970). On a semi-automatic power spectrum estimation procedure. In *Proceedings of the Third Hawaii International Conference on System Sciences* 974–977. Western Periodicals Company, Honolulu.

AKAIKE, H. (1974a). A new look at the statistical model identification. *IEEE Trans. Automat. Control* **AC-19** 716–723.

AKAIKE, H. (1974b). Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes. *Ann. Inst. Statist. Math.* **26** 363–387.

AKAIKE, H. (1977a). An objective use of Bayesian models. *Ann. Inst. Statist. Math.* **29** 9–20.

AKAIKE, H. (1977b). On entropy maximization principle. In *Applications of Statistics* (P. R. Krishnaiah, ed.) 27–41. North-Holland, Amsterdam.

AKAIKE, H. (1977c). Canonical correlation analysis of time series and the use of an information criterion. In *System Identification: Advances and Case Studies* (R. K. Mehra and D. G. Laniotis, eds.) 29–76. Academic, New York.

AKAIKE, H. (1978). A Bayesian analysis of the minimum AIC procedure. *Ann. Inst. Statist. Math.* **30** 9–14.

AKAIKE, H. (1979). A Bayesian extension of the minimum AIC procedure of autoregressive model fitting. *Biometrika* **66** 237–242.

AKAIKE, H. (1980). Likelihood and the Bayes procedure. In *Bayesian Statistics* (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.) 143–166. Univ. Press Valencia.

AKAIKE, H. (1994). Implications of informational point of view on the development of statistical science. In *Proceedings of the First US/JAPAN Conference on The Frontiers of Statistical Modeling: An Informational Approach* **3** (H. Bozdogan, ed.) 27–38. Kluwer, Dordrecht.

AKAIKE, H. and KANESHIGE, I. (1962). Some estimation of vehicle suspension system's frequency response by cross-spectral method. In *Proceedings of the 12th Japan National Congress for Applied Mechanics* 241–244. Univ. Tokyo Press.

AKAIKE, H. and NAKAGAWA, T. (1988). *Statistical Analysis and Control of Dynamic Systems.* Kluwer, Dordrecht.

AKAIKE, H. and YAMANOUCHI, Y. (1962). On the statistical estimation of frequency response function. *Ann. Inst. Statist. Math.* **14** 23–56.

DAVISSON, L. D. (1966). The prediction error of stationary Gaussian time series of unknown covariance. *IEEE Trans. Inform. Theory* **IT-11** 527–532.

FISHER, R. A. (1935). The logic of inductive inference. *J. Roy. Statist. Soc. Ser. A* **98** 39–54.

FUKUNISHI, K. (1977). Diagnostic analyses of a nuclear power plant using multivariate autoregressive processes. *Nuclear Science and Engineering* **62** 215–225.

JOHNSON, N. L. and KOTZ, S., eds. (1991). *Breakthroughs in Statistics. Volume I: Foundations and Basic Theory.* Springer, New York.

JONES, R. H. (1975a). Probability estimation using a multinomial logistic function. *J. Statist. Comput. Simulation* **3** 315–329.

JONES, R. H. (1975b). Estimation of spatial wave number spectra and falloff rate with unequally spaced observations. *J. Atmospheric Sci.* **32** 260–268.

NAKAMURA, H. (1994). Statistical identification and optimal control of thermal power plants. In *Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach 3* (H. Bozdogan, ed.) 57–79. Kluwer, Dordrecht.

NAKAMURA, H. and AKAIKE, H. (1981). Statistical identification for optimal control of supercritical thermal power plants. *Automatica* **17** 143–155.

PEIRCE, C. S. (1955). Abduction and induction. In *Philosphical Writings of Peirce* (J. Buchler, ed.), 150–156. Dover, New York.

THIS WEEK'S CITATION CLASSIC (1981). *Current Contents* **51** 22 (December 21). [Also included in (1986) *Contemporary Classics in Engineering and Applied Science* 42 (A. Thackary, ed.) ISI Press, Philadelphia.

TIKHNOV, A. N. (1965). Incorrect problems of linear algebra and a stable method for their solution *Soviet Math. Dokl.* **6** 989–991.

WADA, T., AKAIKE, H., YAMADA, Y. and UDAGAWA, E. (1988). Application of multivariate autoregressive modelling for analysis of immunologic networks in man. *Comput. Math. Appl.* 15 713–722.

WADA, T., SATA, S. and MATSUO, N. (1993). Application of multivariate autoregressive modeling for analyzing chloride/potassium/bicarbonate relationship in the body. *Medical and Biological Engineering and Computers* **31** S99–S107.