John Tukey and the Correlation Coefficient

David R. Brillinger *
Statistics Department
University of California
Berkeley, CA, 94720-3860
E-mail: brill@stat.berkeley.edu

Abstract

At a young statistical age John Tukey joined Charlie Winsor's Society for the Suppression of Correlation Coefficients. Its guiding principle was that most correlation coefficients should never be calculated. In this article some of John Tukey's complaints are brought out by a succession of quotes from his writings. Also his career is reviewed and with some remarks about his mentor C. P. Winsor. Consideration turns to the case of partial correlation coefficients. Stationary time series extensions of these are defined and an analysis is carried out with Mississippi River daily flow rates at 10 dams along the river.

Keywords

Coherence, coherency, correlation coefficient, graphical model, history, network, partial coherency, partial correlation, J. W. Tukey, C. P. Winsor.

1 Introduction

John W. Tukey (JWT) was the Interface Keynote Speaker three times! - in 1972, 1982, 1986. His topics of the second and third talks were: "Another look at the future" and "The interface of computing: in the small or in the large?". The concern of this article is to address some aspects of his career related to the subject of this conference - On The Interface of Computing and Statistics.

The article begins with some description of JWT's career. There are then sections describing JWT's contributions to: Computing, Information Retrieval, Statistics, and the Correlation Coefficient. Next are sections: Some Details, Partial Correlation (including an application to river flow), Extensions, Some Stories and finally a Summary and Conclusions.

The focus is on correlation coefficients and their extensions. Correlation coefficients apparently peeved JWT from the start of his career in statistics.

There are two numbering systems employed for the references. One follows that of *The Collected Works of John W. Tukey* Vol. VIII, the second has classical form. JWT wrote so many papers a standard form has not yet been settled upon.

*The research was supported by NSF grants DMS-9704739 and DMS-9971309.

2 JWT's Career

John Wilder Tukey (JWT) was born 16 June 1915 in New Bedford, Massachusetts. He was educated at home entering Brown University on the basis of College Board exams. He was there 1933-37 obtaining a B.Sc. in 1936 and a M.Sc. in 1937, both in Chemistry. He moved on to graduate work at Princeton University, first in Chemistry, but quickly switched into Mathematics. He had taken some amount of mathematics at Brown so the switch was not totally dramatic. He obtained a M.A. in 1938 and Ph.D. in 1939, both in Mathematics. His thesis was titled *Convergence and Uniformity in Topology.* The greater part of it appeared as the second publication of Princeton's Annals of Mathematics series.

John stayed at Princeton for the rest of his career. For 1939-41 he was H. B. Fine Instructor of Mathematics. For 1941-48 he was Assistant Professor of Mathematics. When World War II came he worked, 1941-45, at the Fire Control Research Office in Princeton. In 1945 he moved to a joint position at Princeton and Bell Telephone Laboratories, being in succession at the latter: Member of Technical Staff, Associate Executive Director, Communications Research Principles, and finally Associate Executive Director, Research Information Sciences. At Princeton he was in turn: Associate Professor of Mathematics (1948-1950), Professor of Mathematics (1950-1966), Professor of Statistics (1966-1985). He retired from both these organizations in 1985, but did not stop working. He died 26 July 2000 in New Brunswick, New Jersey.

JWT was appointed to many positions of importance and received many honors. During the years 1960-64 he was a member of the President's Scientific Advisory Committee. He received the National Medal of Science in 1973, the James Madison Medal of Princeton University in 1984 and was elected a Foreign Member of the Royal Society of London in 1991. He received seven honorary doctorates. Beyond that he was renowned for the number of committees he served on.

To elaborate more on a few of the early details: JWT has described the period 1938-41 as one of his being "a mathematician in the narrow sense." In May 1941 things changed, dramatically, when he joined the Fire Control Research Office, at 20 Nassau St., Princeton, under Director Merrill M. Flood. His title there evolved from Consultant to Assistant Director and most importantly he came to work together with C. P. Winsor. There will be more on "Charlie Winsor" later.

We next will review parts of JWT's career concerned with computing, information retrieval and statistics.

2.1 JWT and Computing

Throughout his career JWT worked with pencil, paper and slide rule making use of a variety of computing tricks. Some of these are indicated in his book [1977a]. (References given in this form refer to the numbering system of the Bibliography in *The Collected Works of John W. Tukey* Vol. VIII.) Those who went on consulting expeditions with him saw him working hard to finish things by the end of the day and he usually did. This happened even in days when laptops and handheld calculators were in the future and despite some fairly complex analyses being at hand.

He was long interested in the construction of electronic computers. For example in Anscombe (1988) he remarked that in 1939 Brockway MacMillan

> *"and I were both interested in audio and one thing and another. We actually bought a set of parts with the intention of putting up one bit's*

worth of what would now be a central processor."

When John moved on to the Fire Control Research Office he got to use one of the first IBM multiplying punches, Fernholz and Morgenthaler (1997). He further talked about using a 403 (401?) accounting machine in 1942-43 to prepare scatter plots [1973c].

In the classic and semenal 1946 paper by Burks, Goldstine and von Neumann one finds: "The authors also wish to express their thanks to Dr. John Tukey, of Princeton University, for many valuable discussions and suggestions." When I contacted Burks to ask for some detail he wrote,

> *"John Tukey designed the electronic adding circuit we actually used in the IAS (Institute for Advanced Studies Computer). In this circuit, each binary adder fed its carry output directly into the next stage without delay. ... And this was the circuit actually used because it was reliable and much simpler than the alternative"*

JWT set up the Statistical Techniques Research Group (STRG) at the University in the mid-50's. In 1959 it had an IBM 650. I believe that this was the first electronic computer on campus. For a period he carried around an HP-45 and I believe gave Hewlett Packard advice on the statistical functions. David Andrews and he developed "6 line teletype plots", Andrews and Tukey (1973). JWT created some words that came into common use in the field: "bit", "linear programming" and perhaps "software". He made contributions to numerical analysis, approximations and Monte Carlo. He is referenced in Abramowitz and Stegun (1964) for having contributed material concerning the generation of random numbers.

His name will surely echo through history in connection with the Fast Fourier Transform (FFT). In 1963 in Mathematics 596 [1984j] he showed that a fourier transform of a series of length GH could be computed using $(H + 2 + G)GH$ multiplications instead of the naive number G^2H^2. He proceeded by writing $t = g + hG$, $n = k + jH$ with $0 <= g, j < G$, $0 <= h, k < H$. The Cooley-Sande-Tukey FFT algorithms then arrived on the scene.

He provides some history of the FFT in the 1970 Scott Lectures [1985gh].

2.2 JWT and Information Retrieval

JWT was interested in citation indices for many years. An early paper, "A citation index for statistics and probability." [1963f], described his enlisting an army of renowned statisticians as helpers in coding the articles in many journals.

In particular we list [1962c], [1966I], [1969g], [1973e], [1973f], [1973g], [1975d], [1975e], whose titles are given in the References section at the end of this article. The important components of his citation indices were taken over by the American Mathematical Society and are available at its web site.

After retirement in 1985 JWT became a consultant at Xerox PARC in California. There, working with others, he obtained a succession of patents. It is interesting to read the list of some of the titles:

1. Method and apparatus for information access employing overlapping clusters (1999)
2. Method of ordering document clusters given some knowledge of user interests (1999)
3. Automatic method of identifying drop words in a document image without performing character recognition (1999)

4. Automatic method of generating thematic summaries from a document image without performing character recognition (1998)

5. Method and apparatus for information access employing overlapping clusters (1998)

6. Method and apparatus for automatic document summarization (1997)

7. Detecting function words without converting a scanned document to character codes (1995)

8. Scatter-gather: a cluster-based method and apparatus for browsing large document collections (1995)

9. Itterative technique for phrase query formation and an information retrieval system employing same (1994)

2.3 JWT and Statistics

John Tukey's contributions to the field of statistics are massive. There is clear evidence of this in the eight volumes of *The Collected Works of John W. Tukey*, the wonderful series for which Bill Cleveland was the Editor-in-Chief. Yet still there are important subjects, like robustness, that escape that series. There are further papers written after a volume on a particular topic had appeared and there are countless unpublished pieces of work.

He analyzed what he did, perhaps more than any other statistician ever has. He was particularly concerned with the way to do statistics focussing often on CDA and EDA. Incidentally the first occurrence of the use of the terms "Exploration and Confirmation" that I find is in [1950a]. There he also refers to the *tabula rasa* fallacy that an objective scientific approach is one with an empty mind and no hypotheses or preconceptions.

The "big four" topics that he worked on are perhaps: multiple comparisons, spectrum analysis, robust/resistant techniques, and EDA. Other topics to mention include: graphics, quality control, biostatistics and ANOVA.

One can mention the 22 books that he wrote and edited.

To end consider the remark:

> *"...I know of no large machine installation whose operations are adapted to the basic step-by-step character of most data analysis, in which most answers coming out of the machine will, after human consideration, return to the machine for further processing."*
>
> [1961a]

Here we find JWT anticipating modern interactive statistical computing.

3 JWT and the Correlation Coefficient

3.1 Some Quotes

John Tukey spoke and wrote concerning the uses and limitations of correlation and regression coefficients. In particular he did not think highly of the former, except in limited circumstances and he recognized that there were substantial difficulties of interpretation going along with the latter.

> *"The correlation coefficient has two real excuses for preference over other measures of relation or association. First, it is independent of changes in*

scale of the two variables concerned, so that it can properly be used when the size of the units are devoid of meaning. Second, it is symmetrical in the two variables, so that it can properly be used when the causal relation of the variables is unclear or nonexistant. For either excuse to function, it is usually necessary for there to be a well- defined population in which we are interested.

[1950a]

"Does anyone know when the correlation coefficient is useful, as opposed to when it is used? If so, why not tell us? What substitutes are better for which purposes?"

[1954a]

"If Charlie Winsor were alive, he would perhaps be surprised to see the writer on a program with the word "correlation" in the title, for he would know me as an early member of his informal society for the suppression of correlation coefficients - whose guiding principle is that most correlation coefficients should never be calculated. ... For more and more perplexities and confusions seem to be related to one another and to the unnecessary use of correlation.

I frequently hold to the position that correlation coefficients are justified in two and only two circumstances, when they are regression coefficients, or when the measurement of one or both variables on a determinate scale is hopeless. ... All in all, I ascribe the success of the correlation coefficient in establishing the facts about human inheritance and in laying the foundation for the work of Fisher and Wright on qualitative inheritance to the fact that in this application, the coefficients were also regression coefficients, or nearly so.

... it seems to the writer to be far better to work with regression rather than correlation.

... analyses in terms of causation are usually more appropriately stated in terms of regression than of correlation.

One of the major arguments for regression instead of correlation is potential stability. We are very sure that the correlation cannot remain the same over a wide range of situations, but it is possible that the regression coefficient might.

We are seeking stability of our coefficients so that we can hope to give them theoretical significance.

We are thus led to a rank order of preference among coefficients: First, appropriate regression coefficients, ... Second, coefficients of (linear) determination. Third, and as a last resort, coefficients of correlation - used because we see nothing else to do."

[1954b]

"If our real question cannot be answered by a correlation coefficient it can be fatal to insist on using a correlation coefficient to NOT answer our question, whether or not some other question appears to be answered.

Like the late Charles P. Winsor, a statistician far ahead of his time, I find the use of a correlation coefficient a dangerous sympton. It is an enemy of generalization, a focusser on the "here-and-now" to the exclusion of the "there-and-then". Any influence that exerts selection on one variable and not the other will shift the correlation coefficient. What usually remains constant under such circumstances is one of the regression coefficients.

Why then are correlation coefficients so attractive? Only bad reasons seem to come to mind. Worst of all, probably, is the absence of any need to think about units for either variable. Given two perfectly meaningless variables, one is reminded of their meaninglessness when a regression coefficient is given, since one wonders how to interpret its value. A correlation coefficient is less likely to bring up the unpleasant truth - we THINK we know what $r = -.7$ *means. DO WE? How often? Sweeping things under the rug is the enemy of good data analysis. Often, using the correlation coefficient is "sweeping under the run" with a vengeance. Being so disinterested in our variables that we do not care about their units can hardly be desirable.*

In statistical THEORY, the correlation coefficient seems to be mildly convenient in discussing the behavior of estimates and other statistics. This is very different from calculating a correlation coefficient from the data.

When we are calculating from data, regression formulas, covariances, and even variance components answer meaningful questions better."

[1969f]

3.2 Charlie Winsor (1985-1951)

C. P. Winsor has been mentioned several times above. It is clear that JWT had tremendous respect for him. The EDA book[1977a] has at the outset,

> *"Dedicated to the memory of CHARLIE WINSOR, biometrician, and EDGAR ANDERSON, botanist, data analysts both, from whom the author learned much that could not have been learned elsewhere"*

Speaking in 1961 about data analysis JWT [1961u3 in 1986d] said that "The core of insight around which these attitudes developed came from working alongside Charles P. Winsor during World War II.

In the video, Gnanadesikan and Hoaglin (1993), JWT made the following remarks (somewhat paraphrased),

> *"It was Charlie and the experience of working on the analysis of real data, that converted me to statistics.*
>
> *I learned an awful lot from Charlie Winsor ... statistics, physiology ... a lot that was not in the literature."*

Of Winsor's life he says

> *"He was overseas in World War I, underage, as an ambulance driver. He returned and enlisted when of age. He was a Harvard graduate in engineering after which he worked briefly for the telephone company.*

Then he went to Johns Hopkins and worked with Ray Pearl, after which he obtained a Harvard Ph.D. in physiology."

Winsor was referred to as an "engineer-turned-physiologist-turned-statistician."

In 1941 Winsor went to the Fire Control Research Office in Princeton where JWT then "spent 4 years learning from him." Winsor would have been about 46 and JWT about 26 in 1941. There were lunches, dinners, and walks, Gnanadesikan and Hoaglin (1993).

JWT has pointed out that Winsor wrote the second paper on variance components (Winsor and Clark (1940)) JWT wrote three scientific papers with Winsor: "Industrial statistics" [1946d], "Low moments for small samples: a comparative study of order statistics" [1947b] and "The education of a scientific generalist" [1949d].

John named some principles after Winsor.

Winsor's Principles.

1. "distributions of ... very large samples are, except for discreteness, reasonably normal in the middle" [1977b] This was also phrased as "All actual distributions are Gaussian in the MIDDLE!" [1982m].

2. "... no one ever had more than 100-200 degrees of freedom in practice, because by then the heterogeneity of variance is beginning to show." [1986a]

He also baptized Winsor's robustizing technique, of replacing outliers by the nearest observation not seriously suspect, as *Winsorizing*.

4 Some Details

By way of interpretation and summary of the quotes of Section 3 we add some remarks.

1. We start with a JWT quote

"Any influence that exerts selection on one variable and not the other will shift the correlation coefficient. What usually remains constant is one of the regression coefficients."

To formalize this slightly suppose

$$Y \;=\; \alpha + \beta X + \epsilon,\; \sigma^2 \;=\; var\ \epsilon$$

σ may change with experiment, but not β. For example better controlling the experiment reduces σ.

2. JWT refers to use of the correlation coefficient as an indicator of uncertainty. Perhaps he had in mind that for the simple model of linear regression, $Y \;=\; \alpha+\beta x+\epsilon$ one has

$$var\ \hat{\beta} \doteq s_Y^2(1 \;-\; r^2)/ns_X^2$$

3. He refers to the coefficient of determination

$$r^2 \;=\; 1 \;-\; variance\{residuals\}/variance\{response\}$$

being of greater use in particular as a measure of the quality of fit.

4. Abelson and Tukey [1959h] use r^2 as a criterion for the success of a choice of scores. In [1963c] they refer to r as a particular noncentrality parameter in a distribution.

5. Basford and Tukey (1999), p. 222 "... use the correlation coefficient as the similarity measure among the attributes."

6. JWT and his fifth cousin Paul Tukey suggest locating potentially interesting structure, as in projection pursuit, by working with the index

$$r - r_{robust}$$

[1985k]. Here r_{robust} denotes a robust correlation coefficient, see Section 6.1 below.

It is clear that JWT was unhappy about many of the uses correlation coefficients were put to. He was also concerned about the interpretation of regression coefficients.

> *"It is not enough to know what a regression coefficient multiplies; we must know what other carriers (explanatories) are offered."*
>
> Mosteller and Tukey [1977b]

> *"Another quite distinct general purpose (of regression) is the removal of what parts of y seem to be associated with the given x's in order to use the clarified values – the residuals from the regression – in further analyses. The removal of seasonal effects before studying the interrelation of two meteorological parameters would be a natural example. (Since almost any two such will both show seasonal effects, the appearance of a spurious relationship can almost be guaranteed if we do not clarify our variables.)"*
>
> Tukey (1979)

A correlation coefficient becomes a partial correlation coefficient when computed between the residuals of two variates having removed the effects of a third by linear regression.

4.1 Definitions

For a trivariate random variable (or "chance quantity" as JWT would say) $\{Y_1, Y_2, Y_3\}$ with correlations ρ_{ij}, the partial correlation of Y_1 with Y_3 given Y_2 is

$$\rho_{13|2} = (\rho_{13} - \rho_{12}\rho_{23})/\sqrt{(1-\rho_{12}^2)(1-\rho_{23}^2)}$$

It has the interpretation

$$\rho_{13|2} = corr\{Y_1 - \beta_1 Y_2, Y_3 - \beta_3 Y_2\}$$

β_1, β_3 being the linear regression coefficients.

In the case of a stationary time series, $\mathbf{Y}(t) = (Y_1(t), Y_2(t), Y_3(t))$, the partial coherency at frequency λ of the processes Y_1 and Y_3 given Y_2 is

$$R_{13|2}(\lambda) = (R_{13}(\lambda) - R_{12}(\lambda)R_{23}(\lambda))/\sqrt{(1-|R_{12}(\lambda)|^2)(1-|R_{23}(\lambda)|^2)}$$

paralleling the trivariate definition. Its mod-squared represents the coherence of the processes Y_1 and Y_3 having removed the linear time invariant effects of Y_2. The coherence, $|R_{13|2}(\lambda)|^2$, has the interpretation as the proportion of Y_1's power at frequency λ explained by Y_3 having removed the linear time invariant effects of Y_2 from both Y_1 and Y_3. An estimate may be computed once the second-order spectra of the series have been estimated.

The coherence and the partial coherence parameters and their estimates are discussed extensively in Brillinger (1975).

4.2 Flow of the Mississippi River

The waters of the Mississippi River flow from Minnesota to the Gulf of Mexico. Flooding has long been a concern. Through the years the U.S. Army Corps of Engineers has constructed a series of locks along the river for flood control and as an aid to navigation. The waters of the river may be viewed as a system added to by precipitation and by the flow from entering streams and runoff and reduced by evaporation, absorption and diversion. This situation appears a useful one for studying the effectiveness of the partial coherency and partial coherence parameters as tools for system identification. Consider for example the rates of water flow, $Y_8(t), Y_9(t), Y_{10}(t)$ at Dams 8, 9, 10 along the river. Their locations may be seen in Figure 1. One sees the bulk of the waters passing from Dam 8 to Dam 9 and then onto Dam 10. One sees the Upper Iowa River entering between Dams 8 and 9 and the Wisconsin River entering between Dam 9 and Dam 10.

Following the connection of the dams in series one can envisage the following models for the flows at Dams 8 and 10, in terms of that at Dam 9,

$$Y_{10}(t) \;=\; \mu \;+\; \int_0^\infty a(t-u)Y_9(u)du \;+\; \epsilon(t)$$

$$Y_8(t) \;=\; \nu \;+\; \int_{-\infty}^0 b(t-u)Y_9(u)du \;+\; \eta(t)$$

where $\epsilon()$ and $\nu(t)$ represent approximately independent error series, e.g. coresponding to the water entering from the Upper Iowa and Wisconsin Rivers. If the situation is stationary and the errors series are independent, then the partial coherence of $Y_8()$ and $Y_{10}(t)$ given $Y_9(t)$ will be identically 0. This possibility will now be investigated.

4.3 Results

The basic data employed are the daily water flow rates as recorded along the Mississippi River in Wisconsin at Dams 2 to 10. The data were obtained from the WWW site:

www.mvp-wc.usace.army.mil/projects/lock_dam.html

Figure 1 was taken from that web site.

Figure 2 provides estimates of the partial coherences, $|\hat{R}_{i,i+2|i+1}(\lambda)|^2$, for the triples of successive dams along the river as well as of the simple coherences, $|\hat{R}_{i,i+2}(\lambda)|^2$. The horizontal lines in the figures give the approximate upper 95 % point of the distribution of the partial coherence estimates assuming the series $\epsilon()$ and $\eta()$ are independent, see Brillinger (1975).

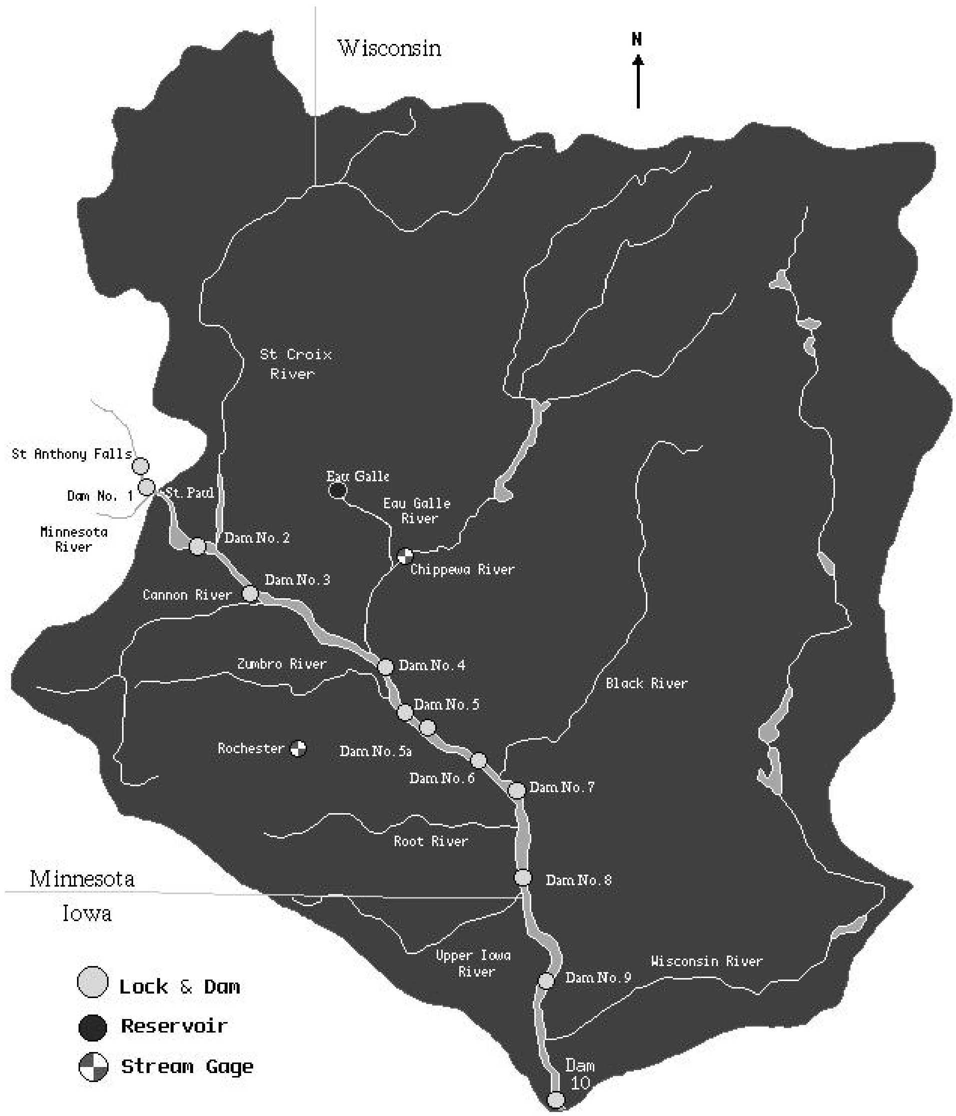

Figure 1: The locations of the 10 dams along the Mississippi whose flow rates are studied.

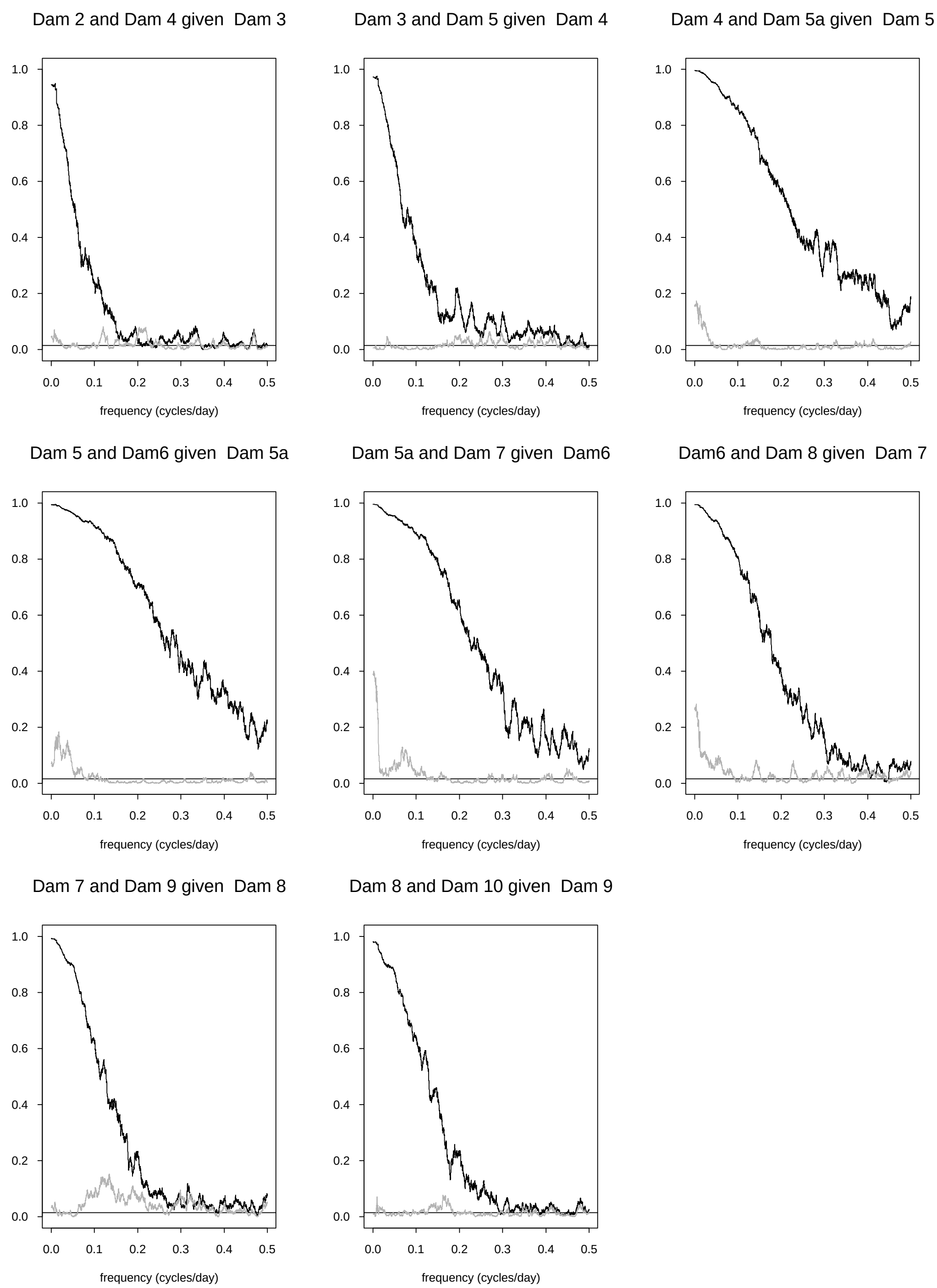

Figure 2: The partial coherences of the successive triples of dams down the river

In each figure the curves that are higher, starting near 1, are the coherence estimates. Unsurprisingly, the series are strongly related at the lower frequencies, the water molecules at Dam i generally flow down to Dam $i+2$. The slowly changing phenomena, e.g. changing seasons, also pass along the river. One explanation for the lower estimated values at the higher frequencies is that rapidly changing phenomena diffuse and are more short-lived. The estimates also appear to be higher for dams that are closer together.

The lower curve, in each figure, near the horizontal line, represents an estimate of the coherence of the series $\epsilon()$ and $\eta()$. For most frequencies the value is consistent with independence, but there are notable exceptions. An explanation could be that sometimes storms will put water into several stretches of the river simultaneously. Another is that there are climatic conditions affecting various of the series simultaneously.

4.4 Discussion

What are the limitations of this study? Well correlation is not causation, but it is pretty easy to design an experiment to show that some of the water at Dam 8 say does go on to reach Dam 9 and then Dam 10. The technique of partial coherency analysis does appear to have worked well here, as it should. The estimated partial coherences are much reduced at many frequencies. There are interesting frequency bands where some association apparently remains. In speaking about frequency-side as opposed to time side analysis in [1950k] JWT remarks that by using frequency analysis the experimenter

> *"will have a chance to discover something unexpected, should it occur"*

A related analysis of point processes of the times of extreme flows, based on the same data set, is presented in Brillinger (2001).

5 Extensions

5.1 Robustness

One of the topics on which JWT worked for his whole statistical career is that of robustness. He has remarked

> *"Some years of close contact with the late C. P. Winsor had taught the writer to beware of extreme deviates."*
>
> Tukey [1960f]

In Mosteller and Tukey (1977), p. 211, a "resistant and robust analog" of the correlation coefficient is set down namely

$$\pm\frac{1}{\sqrt{1+\hat{bar}(y-bx)/b^2\hat{bar}(x)}}$$

where bx denotes a robust fitted value for y (e.g. biweight-based) and $\hat{bar}$ denotes a robust variant of $\hat{var}$. Also $\pm$ goes with the sign of b. They give an example where the classical correlation is -.7333 and the robust estimate is .99992

In the same section they also speak of the calculation of correlation coefficients as *"deplored by some"*.

5.2 Mutual Information

The correlation coefficient and the coefficient of determination may be viewed as measures of the strength of dependence of statistical variates. They have the oft-quoted disadvantage that their taking on the value 0 does not imply that the variates are independent, despite the converse holding. The mutual information does not suffer this disadvantage. Time series variants of this quantity are discussed in Granger and Lin (1994) and Brillinger (2001).

Presumably JWT would have been dissatisfied with the summary being the value of the mutual information, even be it a function of time or frequency. He would have wished some estimate of the inherent relationship.

6 A Few Stories

1. The story is that when that compendium of formulas and tables, Abramowitz and Stegun (1964), was being prepared Marvin Zelen and JWT had lunch at the Cosmos Club to discuss methods of generating random numbers and their applications. JWT came up with a number of ideas. Zelen asked JWT why he had not published them. JWT's reply was that he had just thought of them.

2. When this author told JWT of a simple way that he had heard about for converting between (some) temperatures in degrees Centigrade and degrees Fahrenheit, namely 61 F flips to 16 C and 82 F flips to 28 C, after a moment's thought JWT came up with and 40 F goes over to 04 C.

3. One of the things that went along with working with JWT was telephone calls. He could be anywhere in the country and would call with ideas. This was particularly noticeable when I was at Princeton. The call would come very early, eg. 7am. John would always start with, "I hope that I haven't woken you up." Of course we lied.

4. Once there was a very heavy early snowfall in Princeton and the roads were closed. John was so concerned about the birds that he walked out some distance from his house to get a large packet of bird seed.

5. Once we had a dinner party at our apartment in Princeton. The Tukeys and various others were there. Being young, my wife and I were showing slides of our recent trip to europe. John was helping out identifying various places and things like unusual zoo animals. Then his remarks stopped for 15 or 20 minutes. When they started again his wife, Elizabeth, said "Now John Tukey don't you pretend that you haven't been asleep." There was a pause then John said, "Well if Lorie hadn't cooked me such a nice dinner then I wouldn't have been."

7 Summary and Conclusions

The concept of a Society for the Suppression of Correlation Coefficients is just wonderful. It is probably more needed now than it was back in the 1940's. Perhaps someone will start a website.

It may be noted that the quotes I have are mainly from the 1950's. I certainly heard him talk about the issue when I was a graduate student and later. I just haven't found any writings.

My conclusion is that we want to remember John's comments on the matter, particularly striving to use regression coefficients and their extensions whenever possible.

Acknowledgements

I thank Manny Parzen for the Tukey-Zelen story and many friends for stories and comments while I have been preparing various articles concerning John. The work has benefited from several published interviews, particularly Fernholz and Morgenthaller (1997,2000).

References

Abramowitz M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions.* U. S. Government Printing Office, Washington.

Andrews, D. F. and Tukey, J. W. (1973). Algorithm AS 61-Six-line plots. *J. Roy. Statist. Soc.* C 22, 265-269.

Anscombe, F. J. (1988). Frederic Mosteller and John W. Tukey: a conversation. *Statistical Science* 3, 136-144.

Basford, K. E. and Tukey, J. W. (1999). *Graphical Analysis of Multiresponse Data.* Chapman and Hall/CRC, Boca Raton.

Brillinger, D. R. (1975a), *Time Series: Data Analysis and Theory,* Holt Rinehart, New York. Republished by SIAM in 2001.

Brillinger, D. R. (2001a). Does anyone know when the correlation coefficient is useful?: a study of the times of extreme river flows. *Technometrics* 43, 266-273.

Brillinger, D. R. (2001b). Networks of time series and point processes. *Proceedings of the Conference Statistics 2001 Canada, Concordia University.* (Chief Editor Y. P. Chaubey).

Burks, A. W., Goldstine, H. H. and Von Neumann, J. (1946). Preliminary discussion of the logical design of an electronic computing instrument. Pp. 17-82 in *Papers of John Von Neumann on Computing and Computer Theory.* (Eds. W. Aspray and A. Burks). MIT Press, Cambridge.

Dolby, J.L., Resnikoff, H.L. and Tukey, J. W. (1969). A ruly code for serials. *Proc. of the Amer. Soc. Info. Sci.* 6, 113-124. [1969g]

Dolby, J. L. and Tukey, J. W. (1973). *Index to Statistics and Probability: The Statistics Cumindex.* Vol. 1, R and D Press, Los Altos. [1973e]

Fernholz, L. T. and Morgenthaler, S. (1997). A conversation with John W. Tukey. Pp. 26-45 in *The Practice of Data Analysis* (Eds. D. R. Brillinger, L. T. Fernholz and S. Morgenthaler). Princeton University Press, Princeton.

Fernholz, L. T. and Morgenthaler, S. (2000). A conversation with John W. Tukey and Elizabeth Tukey. *Statistical Science* 15, 79-94.

Gnanadesikan, R. and Hoaglin, D. (1993). "A discussion". Videotape. American Statistical Association.

Granger, C. W. J. and Lin, J. L. (1994). Using the mutual information coefficient to identify lags in nonlinear models. *J. Time Series Analysis* 15, 371-384.

Mosteller, F. and Tukey, J. W. (1977). *Data Analysis and Regression.* Addison-Wesley. Reading. [1977b]

Ross, I. C. and Tukey, J. W.(1973). *Index to Statistics and Probability: Locations and Authors*, Vol. 5. R and D Press, Los Altos. [1973g]

Ross, I.C. and Tukey, J. W. (1975). *Index to Statistics and Probability: Permuted Titles, A-Microbiology*, Vol. 3. R and D Press, Los Altos, CA. [1975d]

Ross, I. C. and Tukey, J. W. (1975). *Index to Statistics and Probability: Permuted Titles, Microclimatic-Z*, Vol. 4. R and D Press, Los Altos, CA. [1975e]

Tompkins, M. L. and Tukey, J. W. (1966). Permuted (circularly-shifted) indexes to abbreviations: a mechanically prepared aid to serial identification. Pp. 347-355 *Proceedings of the 29th Annual Meeting of ADI.* [1966I]

Tukey, J. W. (1950). "Discussion". *J. Clinical Psychology* VI, 61-74. [1950a]

Tukey, J. W. (1954). Unsolved problems of experimental statistics. *J. American Statistical Association*, 49, 706-731. [1954a]

Tukey, J. W. (1954). Causation, regression and path analysis. Pp. 35-66 in *Statistics and Mathematics in Biology* (Eds. Kempthorne, O. et al). Iowa State College Press, Ames. [1954b]

Tukey, J. W. (1969). Analyzing data: sanctification or detective work? *American Psychologist*, 24, 83-91. [1969f]

Tukey, J. W. (1962). Keeping research in contact with the literature: citation indices and beyond. *J. Chemical Documentation* 2, 34-7. [1962c]

Tukey, J. W. (1963). A citation index for statistics and probability. *Bull. ISI* 40, 747-755. [1963f]

Tukey, J. W. (1973). *Index to Statistics and Probability: Citation Index.* R and D Press, Los Altos, CA. [1973f]

Tukey, J. W. (1979). Introduction to the dilemmas and difficulties of regression. Unpublished manuscript.

Winsor, C. P. and Clark, G. L. (1940). *J. Marine Research* 3, 1.