ROBUST STATISTICS:
FROM AN INTELLECTUAL GAME
TO A CONSUMER PRODUCT

By

Werner A. Stahel

IMA Preprint Series # 580
October 1989
ROBUST STATISTICS:
FROM AN INTELLECTUAL GAME TO A CONSUMER PRODUCT

WERNER A. STAHEL*

Abstract. The Summer Program on "Robustness, Diagnostics, Computing and Graphics in Statistics" stimulated some discussions about the future of robust statistics with a special focus on the measures to be taken in order to foster wider application. This report is intended to document the exchange of opinions.

Key words. robust statistics, statistical software, data analysis

1. INTRODUCTION AND SUMMARY

"Robustness has been only an intellectual game longer than it should. We should approach it as a consumer product."

This is a key contention which John Tukey wrote down in a letter when he was asked what he would like to be a focus of a workshop on robust statistics. This Technical Report is intended to give some impressions of informal discussions and a series of written comments which were stirred up by Tukey’s letter. The exchange of opinions took place before and during the first four weeks of the 1989 summer program on Robustness, Diagnostics, Computing and Graphics in Statistics organized by the Institute of Mathematics and its Applications, Minneapolis. The basic letter and all written contributions are reproduced in the later sections. The co-authors of this report are thus:


They are not listed formally as co-authors since they were not asked to take any responsibility for the whole document and – please keep this in mind – did not get a chance to polish their partly very spontaneous contributions and rewrite them for publication. Let us begin with a summary.

The baseline of the discussions was the impression that robust statistics has been used much less in practice than it should be. In his letter, Tukey suggested that the most efficient way to move ahead is providing and spreading easy-to-use software. The participants of the workshop should therefore try and identify some robust procedures, notably for the one- and two-sample and the linear regression models, which would have a chance to be applied profitably by everyday users of statistics. These users should not be asked to make choices. On the other hand, the methods need not be perfect or optimal in any sense. They should just be clearly better than the ones which are currently in common use.

Before any method can be “released” in this sense, we may need or want some further knowledge about its properties. A first step would be to identify the items of interest and classify them according to their importance. (Cf. Section 2; a more extensive version should appear in the Proceedings of this IMA workshop.)

No consensus was reached during the workshop, but a proposal for linear regression was nevertheless formulated. It roughly consists of calculating a high breakdown point estimator and using it as a starting point for iterating towards the nearest local minimum for a redescending M-estimator. (A precise formulation is planned for the Proceedings volume.) A list of necessary studies was compiled (comment by E. Ronchetti,

*Seminar f. Statistik, Swiss Federal Institute of Technology (ETH), CH 8092 Zurich, Switzerland. This work results from organizing and attending a Summer Program at the Institute of Mathematics and Its Applications (IMA) at the University of Minnesota.
July 26). It states that the major problem is to find adequate critical values, p-values, or estimated standard errors for finite and unbalanced samples. It seems even difficult to discuss what “adequate” should mean here. How far should p-values allowed to be off – for which error distributions, designs, heteroskedasticity patterns, and sample sizes? The answers depend on the importance one assigns to inference questions (see letters by Tukey, June 9 (11), and June 24, and “clean-up”, Aug. 14).

Why was there no consensus? Leaving apart psychological reasons, there have been a number of arguments why general purpose robust procedures are not urgent, desirable, or even possible. It was felt that good data analysis would require different methods for different circumstances. If the general user should get help, it would rather be by supplying specific methods for specific purposes, or maybe – in a more distant future – by statistical expert systems.

S. Portnoy raised the question how large a fraction of data analyses was carried out by programs written specifically for the analysis of a particular type of data with a specific purpose of analysis, rather than by general statistical packages. This fraction might be considerable and even increasing with the spread of workstations.

F. Hampel stressed the importance of these applications. He suggested that the statisticians-by-training should use the general methodology now available to design such specific procedures. A prototype example in which he is involved consists of the analysis of interlaboratory studies for chemical labs for which a robust analysis will be prescribed by regulations in Switzerland and possibly in other European countries soon.

On the other hand, P. Velleman (in a letter to J. Tukey) expects that robust methods will never get wide explicit use, but might be used implicitly, like black boxes, by expert-type software to produce reliable diagnostics.

More generally, many discussants felt the need to address the strategies of data analysis before they could specify the needs for robust methods. This topic would certainly deserve more attention, but fruitful discussions would require more time and preparation. In very general terms, some robust methods may apply for large, others for small datasets, some may suit exploratory purposes better, while others may be preferable for inference problems. On the other hand, even if these distinctions arise quite naturally, it is not clear that they are necessary.

Most participants seemed to concur with the contention that robust statistics should get wider use. While good software was generally seen to be of primary importance, other factors which are required or desirable in order to spread its use were listed (see comments by W. Stahel, July 11, and L. Mili, July 13). Most prominently, there is a need for applied texts directed towards specific areas of applications of statistics – books and papers in the respective technical journals. Examples should treat real problems with actual datasets of the area, not our old statistical textbook examples (the stackloss data, re-analyzed). Close cooperation with “subject matter scientists” would be essential. Then, books and courses on statistics, from introductory to advanced, will gradually contain more and more about robustness, and we may, of course, do our best to speed this slow process up – but we should not expect too much too soon.

On the other hand, we should notice and value the group of practicing statisticians who are willing to learn about robust methods since they know from their daily practice that they need it. Some will eagerly grab what they get if we bother to write it down in a language that they understand, and if we say something about our insight when to apply which method (see comment by J. Wu, July 11).

This leads back to the question whether the “general user” exists, or how common statistical analyses by the large packages are and will be. A next issue then is how much effort would be needed to improve the quality of these analyses by providing general purpose robust procedures for common models such as linear regression. These questions remained open, and everybody will select a direction based on his or her judgement. John Tukey expressed his view like this (imprecise citation from memory):

2
"If we are convinced that robust statistics is a "medicine" that has a potential to do some good to a large number of "patients" (like Aspirin), we should feel a moral obligation to get it out there."

2. **TUKEY'S BASIC LETTER**

[Titles and bracketed comments by W. Stahel]

July 8, 1988

**INTRODUCTION**

You have asked me about the focus of the workshop in Minnesota next summer. I have marinated this question, and come up with two principal points, derived from a belief that robustness HAS BEEN ONLY AN INTELLECTUAL GAME longer than it should, and that we should APPROACH IT AS A CONSUMER PRODUCT, as we must if we are to make it broadly useful.

The first point is that we need to set out a time-ordered (but probably not scheduled) plan for SUCCESSIVE STAGES OF RELEASE – stages to apply in each of the major areas of analysis, but not necessarily to be synchronous. (These steps should be large enough to make retooling economic.)

[Comment: What is a release? It includes a new procedure in a user-friendly computer package or environment along with a manual or book that describes the beneficial use of the method. A robust procedure to be released is a (type of) method for an area of application, e.g., (a particular) M-estimator(s) in linear regression. Tukey continues:]

The second point is that we need to identify the research questions of greatest importance in making us willing to release a specific level in a specific field. These questions should primarily concern the RELATIVE utility of (a) the proposed release and (b) previous releases (including classical techniques). We should NOT insist on perfection, either of technique or of knowledge about performance, before releasing a technique for general use. (We need only to be reasonably sure it is an improvement.)

[In what follows, Tukey classifies the research questions, suggests some release types, and lists areas of application:]**

**KNOWLEDGE**

This second point consists of classifying presently unavailable knowledge in a specific area as one of:

- $\alpha$: knowledge NEEDED BEFORE we would be willing to make the next release,
- $\beta$: knowledge that would encourage us to make the next release,
- $\gamma$: knowledge that would make some of us rather happier with making the next release,
- $\delta$: knowledge that is NOT important for comparing the next release with its predecessors, BUT which is likely to give us a better understanding of the next release’s remaining deficiencies.
- $\epsilon$: Just knowledge.

**RELEASES**

I think that developing a generic schedule of releases, applicable to a wide variety of areas, is far from easy. However, we will not make progress if we do not start! So I offer a very rough suggestion, in the hope of generating comments comments. (The notation RX alludes to a convention used in medicine.)

RX 1.0: Bi-weight-based analysis (one step wherever effective, else iterated),
RX 2.0: Exact-fit-to-chosen-subset analysis (à la Rousseeuw) followed by one step of iterative improvement,
RX 3.0: Near coeffective estimates (for SEVERAL alternative probability structures) probably generated
by George Easton’s asymptotic methods. Undoubtedly, people will want to alter this list in many ways — I hope its present form will stir up discussion.

AREAS OF ANALYSIS

There will clearly be substantial discussion about the major areas of analysis. I offer the following, again as basis for discussion:

A Single batches and one-way classifications
B Regression, linear
C Factorial data [Analysis of Variance]
D Other general linear models
E Factor analysis/multidimensional scaling
F Clustering
G Survival analysis
H Regression, somewhat nonlinear
I Smoothing
J Goodness of fit
K-L ...

WHAT SHOULD BE DONE?

To be specific about Minneapolis, I would hope that we might:

Move the classification of releases forward enough to work tentatively with the result.

For area A,

(i) divide knowledge for RX 1.0 into \( \alpha, \beta, \gamma, \delta, \) or \( \epsilon, \)

(ii) clarify the extent by which RX 2.0 or RX 3.0 seem likely to differ from RX 1.0 in this area.

For area B,

(i) divide knowledge for RX 1.0 into \( \alpha, \beta, \gamma, \delta, \) or \( \epsilon, \)

(ii) same for RX 2.0

(iii) provide one or more frameworks for RX 3.0 as a basis for discussion.

For area C,

(i) decide if there is adequate use for an RX 1.0 here,

(ii) divide knowledge for RX 2.0 into \( \alpha, \beta, \gamma, \delta, \) or \( \epsilon. \)

Make some progress on 2 or 3 other areas.

To some, this might seem a modest program, but I, for one, would be well satisfied to see even a large fraction of this actually happen.

Regards, John W. Tukey
3. EXCHANGE OF LETTERS BEFORE THE IMA MEETING

From J. Tukey, March 20, 1989

Dear Werner [Stahel], Elvezio [Ronchetti], Stephan [Morgenthaler], and George [Easton]:

I seem to have had no detailed feedback about my (postal) letter of 8 July 1988 to Werner and Elvezio concerning how to move robustness from an intellectual game to a consumer product.

1) Do you think this should be done?

2) Do you see any alternative to a phased set of releases (software terminology) or of Rx’s (medical terminology)?

3) Could you live with my three early-draft levels of release?

4) Can you suggest better choices of standard kinds or release?

5) Do you favor classification of knowledge?

6) Would you be willing to work on this at IMA?

(Regrettably I shall not be able to get to Minnesota till the 4th week.)

Regards, John.

From S. Morgenthaler, April 25

Hi John,

this letter is in response to your letter dated July 8, 1988 about the Minnesota workshop. ...

In my opinion, Release 1.0 and Release 2.0 have different goals and can well exist parallel to each other. In Release 2.0 one doesn’t really aim at estimating with high efficiency, but rather at identifying outlying points both in the design and among the observations. Other than that comment, I find nothing wrong with this ordering, but at the same time I’m sure that other schools of robustness would substitute their own products for the three levels of releases. I therefore propose a technical description of the three levels.

Release 1.0 includes methods of data analysis that are relatively simple to compute, to interpret and to analyze. The tools for analysis include influence functions, asymptotic variances and neighborhood models. These methods typically suffer from a loss of efficiency at the models that interest us most, and they do not adapt to the sample size.

Release 2.0 comprises methods that have a positive breakdown point.

Release 3.0 consists of those methods that try to minimize the loss of efficiency due to the demands of robustness. Such methods must adapt to the sample size. One example of a Release 3.0 method is a mixture of a few different M-estimators with mixing weights that depend on the sample. Asymptotically, it may be that such a method is fully efficient. This is a reason why an asymptotic analysis is not useful in evaluating the robustness properties of a Release 3.0 method. This means that one must add the requirement that the method be an improvement over the Release 1.0 methods in a finite sample sense. (Fully adaptive, efficient non-parametric methods don’t belong into Release 3.0!) Another demand is that such methods must be flexible. The user can decide in which sense he wants to protect himself. (For example, against short tails rather than heavy tails.) Another example of a Release 3.0 method is provided by the configural approach. The important point about the use of configural methods is that they can be very well approximated in a completely conditional manner, i.e., based entirely on the sample one has at hand. Whether the integrals are best approximated by Taylor expansions around the integrand’s maximum or in some other way is not clear to me.
The other 2 or 3 areas that you mention should in my opinion certainly include E and H. On the point of classifying knowledge I would think that it is sufficient to agree on the type ‘alpha’.

Best regards, Stephan.

P.S. Copy to Elvezio and Werner.

From J. Tukey, May 3

Dear Stephan and Elvezio:

1) Copies to Werner Stahel, George Easton, Chris Field*, Jim Rosenberger*, David Hoaglin* (*= with copy of basic letter)

2) Thank you for your comments on mine of 8 July last. [E. Ronchetti’s comment missing!]

3) Stephan is missing the boat! If we are going to get robustness in use there will have to be a specific method (or, perhaps, 2 or 3 believed-to-be-nearly-equivalent alternative methods) for each release for each problem. A class of methods will do no good.

And there will have to be some agreement on just which method. Negotiations will not be easy, and may produce some camels (the camel is rumored to have been designed by a committee!), but without focusing down we will not get to the consumers.

The only feasible ways to the consumers are

a) through a few makers of statistical software, or

b) less effectively, through many statistical textbook writers. Even though I have been told, apparently reliably, that the only way to a statistical software maker’s heart is by writing a book, I think route (a) is more promising. I hope that a multiply-cosigned release may have the impact of a “book”. But it has no prayer of doing this unless it is completely specific. (2 or 3 alternatives might still persuade, but with much greater difficulty.)

We do not – and dare not – ask for agreement that a specific technique is best, or invaluable. We will be lucky to get agreement that it is

- better than the previous release,
- not easily improved very much within the class of techniques of corresponding complexity. Here “improvement” combines all three of robustness, stringency, and ease of computation. But we won’t have a release unless we focus on a technique (or maybe two).

Besides proposing groupings of methods under “Release 1.0M”, “Release 2.0M”, and “Release 3.0M” (I have added the “M” for “Morgenthaler”) which do not, for reasons sketched above, seem to fly, Stephan made some remarks about “Release 1.0L” and “Release 2.0L” (again I have added “L” for “letter of 8 July ’88”). Let me quote and reply:

“In my opinion, Release 1.0L and Release 2.0L have different goals and can well exist parallel to each other. In Release 2.0L one doesn’t really aim at estimating with high efficiency, but rather at identifying outlying points both in the design and among the observations.” (L’s added).

I do not see Release 2.0L this way at all! My guess is that we will rarely see data in which Release 1.0L fails at all badly (we do see data occasionally on which arithmetic-mean like methods do fail badly!), so that it is hard to see by example that any Release 2.0? method is better than Release 1.0L. What Release 2.0L, with its Rousseau [implementation] followed by one iteration (say a biweight), can provide is

(a) a known high breakdown point and,
(b) hopefully, a limited number of iterations (very hopefully, one) in situations where a Release 1.0L method needs to be iterated.

What other ways are there to seriously improve on Release 1.0L?

Stephan also suggests that it would be sufficient to agree on what knowledge was classified “alpha”, which the basic letter used for “knowledge needed before we would be willing to make the next release” in contrast to “beta” – “knowledge that would encourage us” – and “gamma” – “knowledge that would make some of us rather happier”. I fear that “alpha” is being interpreted as covering what I had in mind for “alpha”, “beta” and “gamma” combined. If we are going to hit the street, we are going to do it on a “preponderance of evidence” standard, not a “beyond any reasonable doubt” standard. A reasonable definition of “alpha” might be that required to change a key potential cosigner from “not over my dead body” to highly reluctant “I guess I can go along”.

Earlier Elvezio [Ronchetti] had e-written “you certainly want to include confidence intervals (or at least standard errors). Since in my opinion this is alpha-knowledge, we seem to be stuck already with release 1.0A (L)”. Those who are quite happy to live even somewhat dangerously, like myself, would point to

(a) asymptotic standard errors and

(b) finite sample tie-down by Kafadar [in Commun. Statist., Theory Meth. 11, No. 17, and JASA 77, both 1982] as providing good alpha-knowledge for Release 1.0A (L). Stephan’s work on various kinds of confidence intervals may then provide considerable beta-knowledge.

I look forward to comments from any and all. Let’s keep the pot boiling!

Regards to all, John.

P.S. I now feel my use of “Release” in this correspondance was injudicious. There is merit in stressing the relationship to software releases, but rather more merit in being clear about the differences. I propose to change to “Replace” unless there is an outpouring of objection!

John.

From W. Stahel, June 2

Dear John,

I feel very sorry that I have not sent you any reaction to your view of directions for our work in robustness. I tend to keep myself busy with day-to-day obligations. As the summer comes very near, I would very much like to have some discussion about your sketch which I might send out to all the robustness participants before the meeting, or at least bring it into the general discussion from the beginning of the program. It might also be used as a kind of prototype, provoking others to write down their view of priorities in research. I shall try and encourage such short informal contributions which might be labeled “directions and problems”. We are currently discussing topics in the area of “estimating random quantities” and might end up with another informal sketch.

But let me come back to your programmatic sketch:

(*) In the meantime, I received your rejoinder, and re-reactions in this letter are marked by (*)

I like your push for making robust methods ready for general use. Your classification of knowledge is to the point.
RELEASES

As you expected, the releases will be defined differently by different statisticians. If this should encourage all of us to contribute to the process, I feel that releases should be labeled much more broadly, and Stephan has given a good working description in his letter. Coming back to being more specific, I arrive at the following classes for the time being:

RX 1.0  M- and R- estimators with monotone Influence Function
RX 1.5  M-estimators with redescending Influence Function
RX 2.0  Estimators based on exact fit to a subset, followed by a few steps of iterative improvement towards a RX 1.5 type estimator
RX 2.5  Other high-breakdown-point methods
RX 3.0  More ambitious methods, probably specific to the problem class.

Remarks:

(*) z. I can see that I am about to miss your boat as Stephan did. I agree that we should RECOMMEND .ie. 2 methods per release. But this amounts to giving defaults. We should RELEASE more than one method. Let us therefore speak about release 1.0 and default 1.0D which roughly corresponds to your Replace 1.0M and 1.0L (should be 1.0T, for Tukey, of course – I do not perceive you as a letter).

a. I distinguish RX1 from RX1.5 because we probably need much more knowledge for 1.5 than for 1 before we can release it.

b. RX 2.5 is just a hope. Maybe somebody will dream up a method which is cheaper than the subsets idea.

c. Stephan Morgenthaler will try to convince us that coeffective estimators are the most superior class for RX 3. I wish him good luck.

d. I included R-estimators because they may turn out to be slightly more easily sold than M-estimators because they do not require a choice of a scale estimator and because the inference problem is solved so nicely for single and multiple batches. They have been used for a while now in MINITAB.

PROBLEM AREAS

As to the areas, I guess that we should mention

- "classical" multivariate methods, including discriminant analysis;
- ARIMA-type time series models;
- spectral time series analysis;
- spatial statistics.

Clearly, A and B are long overdue to see some releases for really public use. They have been in use in some places (the rreg function of S has certainly been called for hundreds to thousands of data analyses). We should try and collect this experience more systematically and learn from it.

The priorities for the rest of the categories is less clear. Logistic regression formally counts for H (regression, somewhat nonlinear), but some methods are not far from RX 1.0.

But this is the place for adding some more general comments:
1. There should be a distinction between the various uses of the methods. I feel happy to use RX1 and RX2 for regression as an exploratory tool. I would even be prepared to suggest default choices of psi- or other functions, scale estimates, and tuning constants for this use. But I cannot recommend these methods for general use for inference.

2. Inference is no problem for relatively large, well-balanced samples (use asymptotics). Small sample distributions are important, and there are some recent results which I would like to hear summarized by one or two of us. An important question is: When can we rely on asymptotics? - at least until more refined methods like small sample asymptotics have been decently implemented. (Bootstrap may be easier.)

3. Maybe there would be enough knowledge to make me happy in this area. But others want nonparametric rather than robust inference. This is not the right place to expand about this distinction once more. Inference, for me, needs to be adequate only in a neighborhood of a nominal model. This already means that the results will be more valid than those of classical methods, even if approximations are involved. Nonparametric inference is more ambitious but not necessarily more adequate.

4. Another open but unrewarding area of lacking knowledge concerns the many choices, see 12., when methods should be used outside strictly exploratory uses.

5. Engineers have often been open towards new methods. Median filters instead of average filters are applied in picture analysis. We should talk more intensely to applied scientists who may be prepared to learn from general insight we have – not only for classical statistical applications, but also for other ways of dealing with data.

6. As a summary: I have some difficulties talking about a release without talking about purposes such a release should be used for. I would like to emphasize that we need to think and talk about strategies and do good research in this area. This should help us to focus research on robustness, too. For Minneapolis, strategies are only a side issue, however.

(*) Some more comments on the rejoinder:

7. I do object to the term “replace”. First, I think – along the lines Stephan wrote down – that they will coexist, see also my point on purposes above. Maybe we should replace the default initial estimator by a high breakdown point method – but only where it is feasible. It is unfeasible in high dimensions. Some of us may feel that high dimensional problems do not make much sense anyway. I agree in most instances, but not always. Second, it is difficult to replace methods if they are used in practice unless you can show clear and convincing superiority for the new method – and you argue, John, that the user is not likely to see a difference.

8. Let me come back to the distinction between Release and Default. Releases need:
   - Enough knowledge to make sufficiently many colleagues happy to release the method.
   - A good reference, type American Statistician or Peter Rousseeuw’s book on regression.
   - A program with good documentation. All of that can usually be done as easily for a whole class of techniques as for a single default choice. There are those users who like and are willing to learn how to use non-default choices. Why should we treat those unfriendly when it takes so little additional effort? Do not misunderstand this point: We MUST give default choices.

KNOWLEDGE

9. If “alpha” means “not over my dead body”, then the distinction between alpha and beta is more of a matter of love for life. Anyway: For Huber type and bounded influence regression with monotone psi, I think
that asymptotics may work if all hii are leq .1 – it may work better for BIF [Bounded Influence methods], actually. Before releasing RX 1, I would need some simulation result that make this more precise. Is this alpha or beta? The answer to the same question for RX 1.5 will depend on the tuning constants, I guess. This will therefore require more work.

10. A committee? Doug Martin knows that this is hopeless and is prepared to take responsibility for S-plus. It should not be difficult to go on into packages once a release is there.

Is this hot enough for keeping it boiling? It is certainly long ...

Regards, Werner Stahel.

From John Tukey, June 9

Dear Werner:

1) Would you like “revises” instead of “replaces”? I do not feel it to be strong enough, but might live with it.

2) Your classification of RX’s misses a very important point. It classifies revises by how professionals think about them. It should classify by how they seem to the user.

3) I do not know any estimator with monotone influence function that I would use in practice (with the possible exception of the abhorrent detestable, arithmetic mean!). The only advantages to monotonicity that are known are questions of uniqueness and convergence – neither of which seems to be of real practical importance. In every real problem, y_i can be so extreme that we want to forget it.

If you want to iterate a monotone to convergence and then do one or two cycles of redescending, fine – but I would classify the result as redescending!

4) Read the attachment for why we must focus on single methods! I do not mind also recommending a range of flexibility, to be used by those who can manage it, but the central version must be unique. (Hence no one of us will like it!)

5) Coeffective estimators only come in, if we are prepared to work with at least 6 to 10 alternatives. Otherwise the whole configural approach makes the estimators too special, to be responsive to our choice of alternatives.

6) Why worry about ARIMA-type time series analysis? Who really needs a robust version?

On time-series spectrum analyses, I will try to get you a copy of the paper being drafted by David Thomson and Allen Chave, perhaps for Minnesota.

Which specific problems of “spatial statistics?”

7) Your “releases again 1” seems to say that you will “accept badly inferior least-square regression results because I think I have a better assessment of the variances of the coefficients.” Do you really uphold such a heresy?

Be careful, because, if the variance depends on the x-vector – and who ever is sure that it does not – the variance of the coefficients will not be properly estimated by the conventional least-square calculations!

8) Your “releases again 2” about inference indicates a vain hope that asymptotics (and, tacitly, central models) will save “small-sample” problems. We know more, from work by Gross and by Kafadar, about inference in finite samples using biweights for location, than we can possibly ever learn from “central models.”

9) Your “releases again 3” leaves me rather aghast. How can “adequate only in the neighborhood
of a nominal model" meet the real requirements of practice? Robust inference need not be as grouping as "nonparametric inference" – but it dare not be as over-utopian as "inference in the neighborhood of a nominal model."

10) Much of what you say seems to be focused on regression. Does that mean that location is ready to deliver at some high level?

11) My discussions with Paul Velleman leave both of us with the feeling that deeper uncertainties are important in most regression problems. Even if we have the best inference about these particular coefficients, are we fitting the proper variables, wisely re-expressed? In view of the deeper uncertainties, is it wise to be doctrinaire about the minutiae of inference in the narrow sense?

12) I attach:
   A) A talk on "Consumer dataware, with appendix added, after a first discussion with Paul Velleman.
   B) Most of a letter from Paul.
   C) A letter from me to Paul.

   Regards, John.

P.S. e-copies to Stephan, Elvezio, Frank, Peter Rousseuw, Doug Martin, Ray Carroll, Paul Velleman, Paul Tukey.

From W. Stahel, June 12

Dear John,

a. I am about to get a clearer understanding of the differences between your view of the world and mine, and I am learning. It still seems difficult to me to talk about methods without talking about "purpose" or "general aspects of statistical practice (GASP!)". There is this distinction between "exploratory" and "confirmatory" – who invented these useful terms? I have a perception that New Jersey is attracting explorers.

b. You seem to have a hope to devise methods that work well as exploratory tools and allow at the same time for adequate inference in models that are considerably wrong – but hopefully adequate in the aspects of interest (like considerably wrong distribution of errors, heteroskedasticity, but correct structure of regression function, a coefficient being of interest). This is a very high goal, and I am not sure if such a view might bring us farther away rather than nearer to a "release".

c. Specific answers to your letter: (There is too much defense in these points. You may as well skip them) [Numbers refer to the foregoing letter.]

1) Do you think that LMS regression should replace M-estimation? Should it just replace the starting point for some iterative steps with a redescender? In the latter case, I would agree that it is a revise. Or nothing at all: it should be included in the first release. (I guess that you do not want to call RX 1.0 a revise of least squares.)

2) I can see your point. It may well be that with some discussion in Minneapolis, we will converge, see next point:

3) Conservative redescenders are quite ok. Others do not just show uniqueness and convergence problems, but also small sample problems. This may be connected to what you mention in 8) (Gross and Kafadar – I do not know their work). Frank has some intuition and some experience, and this is why we continue to
talk about redescenders rather than biweight.

4) Agree. One method for one purpose, but there may be several (two?) purposes, essentially exploratory and inference.

5) See b. and 9)

6) Let Doug Martin and Hansruedi Künsch answer ARIMA and spatial statistics. In the latter, geographic interpolation (geostatistics) will certainly benefit from robust methods, to give an example.

7) Did I ever say that I believe least squares better than the most simple asymptotic inference for M-estimates? But we need to convince others ... As for the specific question of heteroscedasticity, we teach and aim to facilitate residual analysis, don't we?

8) I do not quite understand your point. For large samples and well-behaved designs, asymptotics will certainly apply. The question is what this means, and this may well be termed "alpha"-knowledge.

9) See b. In smaller or less balanced samples good inference needs a good model. If you want safe inference for untrained users, this again is a very high goal. "Inference in a neighborhood" may be adequate in a rather large neighborhood even if theory is local. (How far is it adequate? is "beta" or "gamma" knowledge.)

10) For me, Wilcoxon (-Mann-Whitney) / Hodges-Lehmann inference is good enough for location and two samples. I have no evidence that we can do much better than these well-established (except H-L) methods.

11) We certainly should not be doctrinaire. But we should be able to sell the methods and not provoke discussions among experts if we can do without.

d. Let me try and summarize: The ideal method would give a good answer in day-to-day problems. The latter are characterized by rather inadequate models. If the specific question does not depend too much on the inadequacies of the model, the ideal method can work in principle. Example: regression with the "right" explanatory variables in an approximately right equational form, with considerably wrong specifications of the error distribution and variances. (Not yet with dependent errors.) The significance level of a test for a coefficient should be adequate in this situation.

This is of course not the way you would say it. But does it relate to your view? Or am I just stubborn?

e. My point: If we dare to work on this ideal method, we (I) would need more information about the inadequacies of models that occur in practice. I do not have much hope for an omnibus method. At the present stage, I prefer to talk about strategies that lead to good results (not just adequate significance levels) and which will probably need to be rather specific to the purpose of the whole analysis: Is the regression model used for prediction, for testing a theory, for finding causal relations, for getting a useful summary for further analyses, ... ?

f. I am collecting contributions to this discussion. They will be available for the participants of the meeting. I plan to have a first workshop on it in the afternoon of our first or second day, July 10 or 11. Too bad you cannot be there. I hope that some of us can pass the main points on. It would help if you could prepare a summary of your view, including your views about our discussion, before you leave. I am preparing an annotated version of your first letter in order to warm up as many as possible before the meeting. I hope to send a draft very soon.

Yours, Werner.
From John Tukey, about June 24
[Transcription of a hand-written letter]

Dear Werner,

1. I have some difficulty with your paragraph a. I can't feel sure of your distinction between "exploratory" and "confirmatory". Let's go a step further (cf. Mallows & Tukey [1980. "A survey of data analysis emphasizing its exploratory aspects. In: de Oliveira, ed. Recent Advances in Statistics. Academic Press, N.Y., pp. 112-172]) and distinguish

EDA = exploratory DA
OCDA = overlapping confirmatory DA, where exploration is followed by confirmation on the same data.
SCDA = separate confirmatory DA, where exploration was on other data
CCDA = careful confirmatory DA, like SCDA, with lots of randomization, etc.

There is no hope of something close to exact 5%, except in the last two cases. Most of the world is in OCDA. It is ordinarily worthwhile to do a good OCDA, rather than a poor SCDA (of course, a good SDCA is better still).

If we have a good experimental situation, with enough randomization, we can have a 5% that is 5%. If we do a broadly robust SCDA, then, as you said on the phone, we can have a 5% that is no more than 7% (and, hopefully, no less than 3%) if we use "the" asymptotic variance (I return, below, to "which" asymptotic variance). Doing an OCDA is probably a little looser than doing the corresponding SCDA - but probably not much.

Whatever we encourage for packages will, we hope, be used for OCDA. ("Hope" because the alternative is too often a poor SCDA.)

2. We have a three-ball juggling act:

(a) Relationship of functional model to fact;
(b) Stringency of estimate;
(c) Appropriateness of estimated variance. We can't have all three perfect - or, even only two perfect. It's not clear that we can even have one perfect. Since our variance estimate will be off, at least somewhat, we ought not accept [?] unduly for stringency just to make the variance estimate unbiased under ideal conditions, since it will be biased under real conditions.

With that accepted, and getting the variance exact not a driving force, whatever we come up with for SCDA use is almost sure to be good for OCDA use - and even for EDA use. So I don't see any serious tension between these aims. (We may want to explore, instead, with some short cut technique, but this is convenience, not principle.)

2.* Which asymptotic variance? If $A/n$, where $A$ is a reasonable function of the sample, so too is $A/(n-3)$, $A/(n+2)$, $A/(n-1)$, etc. We need to choose wisely among these and others.

The best bases we have - and probably the best bases we could have - for this choice is finite-sample results. These can be simulations of various situation-statistic pairs, or average values for linear statistics and sums of squares (rather Gauss-Markoff flavored). We need finite sample results if we are to use asymptotics properly!

3. I also have trouble with your b, especially where it says "correct structure of regression function, a coefficient being of interest". What the coefficient estimates is something that depends both on how that $x$ enters the regression and on what other $x$'s are involved and how they enter. Should we feel free to introduce, leave out, and re-express $z$'s in such a way as to get the apparently best fit? Certainly, if we want a descriptive fit - and promise to avoid interpreting its coefficients. Certainly not, when introducing/leaving-out makes a large difference in the value of the coefficient. Re-expression of $z$'s, and even of $y$, to improve the quality
of the fit is probably both allowed and recommended, most of the time. Indeed, if re-expressing makes the coefficient different, it is likely to be wrong to not re-express.

I would like to see a discussion of, say, three actual situations where not re-expressing makes sense though re-expressing *x’s* does not.

4. (Your c(1)). I am happy with LMS followed by a few redescenders. I would *not* be happy with LMS alone.

5. (Your c(3)). I always thought the biweight is a redescender. Does c(4) ask for something more than a different shape for the *ψ*-function? For Kafadar see JASA some years back (Princeton is not always exploratory.)

6. (Your c(4)). Why are we sure to want two methods? We must discuss re-expression, which carriers for what purpose, and looking for easily described heteroscedasticity *within* all methods. It will be seven times easier to sell one specific than two!


8. (Your c(7 and ...)). You seem to worry about convincing others, but (elsewhere) say, that a packaged “reg” has already been used many, many times. Who persuaded those users?

9. (Your c(8)). I don’t understand “large samples ... asymptotic will certainly apply”. Do you mean Gaussian asymptotics will apply when the model is not Gaussian? Do you mean unsuspected superadaptive asymptotics? Or do you mean only that asymptotic-variance formulas will work well enough over a range of distributions, so long as the estimate’s behavior is not too terrible?

10. (the same). I think “largeness of sample” should be defined relative to the design! A high nose count is not enough! (Of course, this will make many more practical problems into small-sample problems!)

11. (Your c(9)). What do you mean by “safe inference for untrained users”? My basic standard is that it should be safer than what they will do now with least squares. I think a reasonably iterated biweight meets that standard. Surely some will get harmed on high leverage points in both cases, but I suspect the burning will, on the average, be less for the biweight. All we can ask is improvement, not perfection! (And we need to judge this in realistic situations.)

12. (continuing). Sometimes people will get burned with a high-leverage point far off the regression through the other points because the high-leverage point is a good, reliable, solid point – and something like LMS puts it aside. Nothing will give a perfect fix. The case for LMS is that it helps us *on balance* curing more trouble than it makes. (I believe this, do you?)

13. (Your c(10)). As I said on the phone, if you want to work out as far as slash, Wilcoxon will not have a good efficiency. It also seems not very easy to generalize freely.

It seems to me the virtue of robustness is getting back to stringency (e.g. efficacy). Since we will usually be doing OCDA, our validity can never be perfect! Throwing away efficiency to make part of the process valid is then very very hard to justify! (I can’t do it!)

14. (your c(11)). I don’t believe we can sell methods without expert agreement that a specific is better than what is in use! This can be accompanied by suggestions of how to do still better, but the default should be unique!

15. (your d). I don’t find enough in your paragraph about how, since changing the list of variables that
are included changes the definition (meaning!) of what a particular fitted coefficient estimates, we are to choose the list of variables included. I don’t think its enough to say “take the coefficient for the list of variables that seems to fit most carefully!” Indeed, I know this isn’t so, because sometimes there are x’s that would improve the fit, although it is clearly wrong (for the desired interpretation of a specific coefficient) to include them.

16. (first part of part e). I am heartily in agreement about distinguishing
   (i) prediction,
   (ii) testing a theory,
   (iii) finding causal relations,
   (iv) getting a useful summary for further analysis,
       but insist on analyzing each of these carefully. So let us begin!

17. If the purpose is prediction it is hard to see why I want to do inference about coefficients (plural). There would seem to be no reason for concern about any one coefficient. And I feel somewhat uncomfortable about using the “asymptotic variance” formula for all variances and covariances together. The simplest thing would seem to be to jackknife “the works”, and come out with 10 to 20 pseudoprediction formulas, which could be used for any future x vector, with Student’s t applied to the 10 or 20 predictions – this would give a confidence statement for something like an optimum prediction, but would do nothing about the anticipated variability of future values about the “best prediction”, so I wonder who wants it, anyhow?

   Another approach would be to fit as well as we can – by IRLS [Iteratively Reweighted Least Squares] – coming out with a fitting weight for each data point. Then we could re-coordinatize the z-vectors (separately by future point) into \( z_1^*, z_2^*, \ldots, z_k^* \), where \( z_2^* \) up to \( z_k^* \) vanish at the future point concerned, and \( z_1^* \) is orthogonal to \( z_2^* \) up to \( z_k^* \) (using the fitting weights). This means the \( z_1^* \) is made up of the residuals after fitting \( z_2^* \) up to \( z_1^* \) to some one – or some linear combinations of the original x’s (using the final fitting methods of course). Then we can make a picture of \( y - \hat{y} + b_1^* z_1^* \) or even \( y - \hat{y} \) against \( z_1^* \) and do whatever we think is best for a univariate problem! This way, we might get a good start on including future individual variation.

   Diagnosis: Not sure why inference is wanted. If one wants careful inference, one would still not seek the answer from a simple multiple regression. (As always, we may also need to think hard about re-expression.)

18. If the purpose is testing a theory we must be careful about what the theory says. (I regard this case as uncomnon.) If the theory is about one coefficient, usually the theory does not specify which other x’s should go into the fit, although this usually makes a difference. This makes it hard to formulate a sensible inference question.

   Theories that certain variables do not matter are even more of a problem. It is now much more important to settle which x’s are to be in the regression, and even more likely that we won’t know (from the reported theory).

   Diagnosis: If 1, or 3, or even 10 instances of “testing a theory” could be identified and be carefully thought through, it would be a very fine thing! Absent such care, is this alternative more than a stalking horse? (Presumably, if we gather enough data, the answer for any theory is “no!” – so why do we need to be careful about real cases?)

19. If the purpose is establishing a causal relationship then no one set of data has a chance to suffice. On the stochastic side alone, a reasonable number of data sets is the minimum! And our attention to internal errors for any one data set ought to be almost negligible! Student was right when he emphasized multicenter, multiyear agricultural trials as the minimum standard, one where the proper error term was not what went on at individual centers. Fisher was only half-right when he said (in J. Min.Agr.) that hard knowledge came
not from one study with $P = 10^{-6}$ but with an ability to reliably repeat studies each of which is safely significant at 5%. Half-right because a number of studies all leaning the same way (at 50%) is almost as good. (We do not have to get the individual P-values down, though its more nice when we can.)

Diagnosis: There is probably no important place for local inference in any problem of studying causation.

20. If the question is getting a useful summary for further analysis, then I presume you mean that our regression takes place in each study separately and we want to combine the results. This may be a simple comparison of two studies, a multiple comparison of several studies, or the sensible case of 19 above.

In the first two of these subcases, I expect to be much more worried about matching lists of x’s, about matching ranges of x-variation, and about finding good re-expressions that work across the various studies, than I do about exactness of confidence statements, (what do you think?). It would be nice to have exact confidence statements, but its unlikely to be worth 12 cents on the dollar, or even 7 cents or 4 cents! But with sizes of data sets as restricted as in the real world, giving up 10% or 20% in efficiency causes (just) appreciable pain! Yet this is almost certainly the most important use for “inference” in a multiple regression situation.

In the sensible case of 19 above, local errors – and local confidence – are usually not important.

Diagnosis: The comparison of (always somewhat disparate) studies, though difficult because of many more reasons than local confidence can command, is probably the most important use for inference about regressions. It doesn’t seem to me to urge closer to exactness of validity at the price of poorer efficiency.

The main excuses for confidence intervals on multiple regression coefficients are:

1) keeping the consumer from going wild and overbelieving the coefficients to one unit in the last place reported.

2) preparing for comparison with some future study (for this propose we must be careful to document the location, scale and shape of the x-vector distribution as seen with the final IRLS [Iteratively Rewighted Least Squares] weights).

A reasonable approximation to P-values can be important— the precision P-values we could never get would help us only a little.

21. I await your counter-blast with interest!

Anything you can do to promote discussions will be very worthwhile.

Regards, (John W. Tukey)

From C. Field, June 18

Comments on the proposals by John for the Minnesota meetings:

I generally agree with the procedure as laid down. In order to get the users to pay attention and respond, I think we need to have several substantive examples for which a robust analysis is clearly more effective and informative. For instance in a oneway analysis, we could have an example in which the classical analysis picks up a difference which a robust analysis suggests is spurious as well as an example in which a robust analysis picks up important differences which are missed by the classical analysis. These examples should involve current data and not simply be re-worked old data where our distance from the experimenter is large.
4. WRITTEN COMMENTS EXCHANGED DURING THE MEETING

From Luciano Molinari, July 11

The need to step from an "intellectual game" to a "consumer end product" in robustness suggested to me the following consideration which may not correspond to Tukey's intention.

Sometimes statisticians develop quite fancy techniques, possibly or even most often, very useful potentially. They carry them from infancy into a prepubertal or pubertal, i.e., into a decently, but not very, user friendly stage; now the pleasure of the intellectual game is over and they leave their children by themselves. I think they should assume the responsibility for them until they have reached the adult age. This means that they should leave their techniques, actually in most cases programs, in a form which can easily be integrated into some existing, widely used, package, and also that they at least spend some time and energy to supervise and steer this integration process, not just leave it to a variety of commercial software companies.

[Luciano is working at the Children's Clinic in Zurich.]

From Werner Stahel, July 11

A main feature of our discussion on Monday, July 10, might be summarized as follows:

Several things are needed before a method can be released and has a chance to be widely used:
1. Technical paper(s) covering theory, including finite sample results (probably Monte Carlo)
2. Tutorial paper which can be used as a basic reference, including "real world" examples,
3. Computer program implemented in a (or a few) experimental and, later on, standard statistical package(s),
4. A manual indicating correct use and interpretation (along with the tutorial, see 2.).

Of course, we knew this before. Who volunteers for the next step? We have some who work on programs. Who will write the tutorial? (It exists for some methods.) And most importantly: What technical research needs to be done before somebody can write the tutorial that we would be willing to use as a basic reference? This is Tukey's assignment for us. In his terminology, we should list items of knowledge needed, and classify them into "existing", or "alpha", "beta", ...

Who goes first to open the game?

From Jeremy Wu, July 11

If robust method were a consumer product, I would be one of the first potential customers in the "real" world. At the U.S. Department of Agriculture where I work as a mathematical statistician, there is a wide range of challenging estimation and testing problems for inspecting and grading agricultural products ranging from apples to zucchinis. Robustness has long been an attractive concept. However, its status may be described by an old Chinese saying: "There is a lot of loud thunder but very few raindrops".

Agriculture is changing rapidly. We no longer know our butchers or bakers personally. Business is getting so big that inaccurate and inconsistent estimates may mean an enormous financial loss that no one can afford. We no longer can simply look and touch and smell and say: these are "good" apples and those are "bad" grapes. Our organoleptic capabilities cannot detect the presence of traces of pesticides, poisons, and other chemical residues in our foods. We no longer dominate the world trade market. The United States competes with foreign countries in selling grain and other agricultural products with measurable quality factors and nutrient values.
In this environment, breeders have made positive contributions by improving genetics and therefore yield and disease resitance. Engineers are building machines to measure wheat hardness, which is currently determined by visual and tactile methods, and to detect gases that cause odor. Microbiologists and chemists are developing quick methods to detect salmonella in poultry products, aflatoxin in corn and peanuts, and free fatty acid in rice. They proceed with or without a statistician's help; they accomplish their tasks in an efficient or an inefficient manner. Statisticians, on the other hand, continue the time-honored tradition with regression and least squares method to convert near infrared spectral values into estimates of protein content. Compared to the past, they are faster, less expensive and more flexible, but not necessarily more innovative.

Breeders, engineers, chemists, and statisticians alike are interested in “better” statistical methods, but they need readily available tools to make use of these methods. From the perspective of these people in the “real” world, “better” is understandably not defined in terms of whether the estimator is L-type, M-type, or R-type, or whether the breakdown point is 0.4 or 0.5. Not that these concepts are not recognized as important and necessary, but it is certainly more intuitive for a broad audience to consider a method “better” if the predicted protein values are actually “closer” to the chemically determined true values over time. Such, I submit, is also the intent of a robust method.

While lack of theory will not satisfy statisticians, lack of performance will satisfy neither statisticians nor non-statisticians. If robust method were to become a consumer product, it must not only be technically correct according to the manufacturer but also meet the satisfaction of the consumer. That is, it must also be easy to understand, easy to use, and perform well, the meanings of these terms are subject to definition by the consumer, but not the manufacturer.

In this regard, there appears to be a reluctance, or perhaps a fear, that a perfectly good robust method may be abused and ultimately destroyed by the naive consumer, just as regression and least squares method have been abused (but not destroyed!). While some consumers will look for a push-button solution, many will examine the conditions of application and make thorough investigations to satisfy their criteria of performance. In reality, a majority of scientists, statisticians, and other users are able and willing to learn and use new statistical methods.

In summary, the following may be useful in making robust method a consumer product:

1) Stop bashing non-robustniks.
2) Have some finished products available in the market.
3) Educate and increase base of potential users.
4) Let others use and judge.
5) Acquire new knowledge (even from users).
6) Continue to improve (by developing new releases).
7) Go to (2).

Thanks for the opportunity to express these views.

From Lamine Mili, July 13

I suggest the achievement of the following tasks in order to make robustness theory available to a wider public:

1) Write text-books on robust statistics for non-statistitians.
2) Publish papers in non-statistical journals.
3) Organize summer courses on robust statistics for non-statistitians.
4) Write user-friendly packages which implement the LMS estimator followed by a one-step redescending
M-estimator. Due to the possible presence of leverage points, I do not recommend the use of any
M-estimator in the regression case without the availability of a robust starting point.

5) Try to incorporate these software programs in the SAS package. Robust methods must be presented
as techniques which complement LS methods.

6) Write user-friendly programs which implement ARIMA time series methods. Here, some research has
to be done in order to find robust starting points.

From Werner Stahel, July 13

“Robustness is for small to moderate samples.”

This statement needs some explanation. In large samples, parametric models will (almost) never fit
– to the extent that they can be improved on the basis of the data. Even if a model would fit precisely, no
one would need fine tuned robust inference since an effect would be either very significant as judged by the
most conservative method (as well as by naked eye) or so tiny that it is irrelevant anyway. (There may be
rare exceptions, of course.) If inference is needed, at least the distributions of errors may be estimated non-
parametrically, and other deviations may also be assessed and adjusted for. This is the realm of adaptivity
or non-parametrics.

Robust statistics, in my opinion (which in this case reflects the “Zurich dogma!”) is a topic in
parametric estimation and inference. Parametric models are a useful device to formalize questions and to
obtain the redundancy needed for making efficient inference in small to medium sized samples. Fine tuned
robust inference is needed to exploit the redundancy, but not overly so.

There is a peculiar contradiction between this statement and the fact that most robustness criteria
are based on asymptotics. This should allow us to take asymptotic optimality with a grain of salt – yet any
optimality makes us feel better. Clearly, finite sample results are needed to back up the “intuition” that we
gain from (first order) asymptotic approximations. (Finite sample results may include better small sample
approximations.)

On the other hand, robustness DOES apply to the problem of finding nice simple DESCRIPTIONS
of large as well as small datasets – without the need of any probabilistic inference. For description, tools
are needed which are just practical and give “good” simplifications. There isn’t much around to judge
what “good” means. If it means that we want to describe the majority of the data well and separate out the
deviating data, then robust tools, with high breakdown points, are clearly adequate. In this sense, robustness
also applies to large samples.

From S. Morgenthaler, July 13

[Reply to the foregoing comment]

I agree with your distinction between robust inference and resistant exploratory analysis. I may even
agree that resistant exploratory methods are useful tools in analysing small samples. But that does not
mean that they should be based on asymptotic justifications.
From W. Stahel, July 14

A. A further point which came up in the discussions

There was some guessing about how difficult it was to convince SAS to incorporate robust methods. It would certainly be easy to contribute a SAS procedure to the "SUGI" library (distributed, but not supported by SAS). It was felt that robust methods should be sold as a complement, not a replacement of least squares. Diagnostics have always been considered that way, which certainly is a reason why they were implemented so quickly.

Two points should be discussed in the present week:

B. Which procedures should be applied in which situations?

The first distinction that comes to mind is between exploratory and (various flavors of) confirmatory or inference stages of analysis. The latter need a more reliable assessment of the variability of robust estimators and test statistics. LMS, for example, can easily be recommended for exploratory analysis already (see Peter Rousseeuw's talk and book).

On the other hand, there may be procedures which are suitable for "all" tasks. A good candidate is "Yohai's IMA proposal", which is roughly as follows:

1. Calculate a high breakdown point estimator (as good as possible at present);

2. Improve by full iteration (or maybe only a few steps of a reweighting algorithm) for a redescending M-estimator. The tuning constant should be chosen in a mildly adaptive way.

Victor will give us a detailed proposal for multiple regression.

The issue if and when other procedures are necessary or preferable will be discussed in an open discussion on Wednesday or Thursday.

C. List and classify knowledge which is needed or desired before a specific procedure can be "released".

Why don't we try this for the foregoing procedure? The points to be checked will include good choices of various details and finite sample properties. We want to start this on e-mail and discuss the list sometime this Friday.

Everybody is very welcome to join these discussions.

Werner Stahel.

From L. Molinari, July 17

The Friday discussion left me a little enervated, because it seemed to me it would not lead far, and this in a situation which seems to me quite clear (I may well be wrong, as I began to realize this morning, Monday). I will write down in a synthetic form what seemed clear to me, and let others expand or correct it.

1) The main obstacle to a more extensive use of robust statistics is the lack of a user friendly interface to some widely used package. Optimally the robust procedures should be callable from within the package.

2) The main areas where robust procedures should be made available as soon as possible:

   a) location and scale;

   b) regression;

   c) robust covariances, as a first step to many multivariate procedures;

   d) time series.
3) It makes sense to me to distinguish between the use of robust techniques for exploration and for inference.

4) I believe that the knowledge available for topics a), b), c) can be compiled in a clear state of the art form and that this knowledge warrants making the corresponding techniques available to the public (am I wrong for c) above?).

From 1) to 4) it follows that the discussion should by now focus on how getting some of the available techniques into widely used packages: we should try to work out some kind of guidelines for this process (but I am not thinking of the commercial aspects, and for me these guidelines should also be applicable to newly discovered diagnostic or exploratory tools).

Some points to be considered for such guidelines.

0) For a) and b) above it seems that software available in ROBETH reflects the state of the art and should be made available to packages. What about c) ? For d) some robust smoother should also be available!

1) A researcher, who honestly thinks he has developed a method of practical usefulness, should also provide high quality software for it, or at least supervise this process. (This is crucial, is it also utopian?)

2) This software should remain public domain and its inclusion within a few widely used packages still be under the supervision of its developers.

3) It would not bother me if a package's version *+1 includes a different estimator than version *, following new developments in robust theory.

4) It would not bother me either, if the included/recommended estimator fails from time to time. But it would be useful if the package contains at least two essentially different estimators for the same task.

P.S. As a result of some discussion a provocative question: if there are problems with c) above why not give up affine equivariance?

From W. Stahel, July 19

This is a reply to L. Molinari. I guess he addresses a different issue. We should clearly distinguish between “alpha test releases” for a selected range of users and releases for a general audience. Some of us are very happy to try several methods in order to gain a deeper understanding of the data or of the methods (or both). Certainly, this will take place only if we do not have to write all the programs ourselves. This is a call for S-functions.

The general audience needs a method which lets them do inference. Let me emphasize once more Tukey’s call for just one method (per problem area). This means that we should avoid to make a distinction between exploratory and various degrees of confirmatory as well as between good and bad designs in regression, for example, if possible. (This point will be debated this afternoon.) It does not mean that we should wait for the perfect method, however. The users will most probably be confronted with a next release – Tukey called them “replace” or when in a milder mood, “revise” to make clear that the new one should make the old one obsolete rather than complement it. Successive releases should have clearly improved properties to make the replacement worthwhile. The first robust release need only be a substantial improvement over classical methods.

(I start to be too repetitive. Let me close.)
From E. Ronchetti, July 26

SOME KNOWLEDGE NEEDED IN ROBUST REGRESSION (WITHOUT CLASSIFICATION)

It seems difficult to me to come up with a complete list of knowledge needed in regression without having first a reasonable set of proposals. For instance, the inference questions for LMS are obsolete if this estimator is used only as a starting point for a fully iterated procedure. On the other hand, these questions become important in the case of a one-step procedure. Anyhow here are some of the questions we might want to address.

For any proposal:
- In which situations will a given proposal perform best and in which cases do we expect problems?
- Finite (small) sample behavior, in particular
  * small sample approximation of the distribution;
  * bias, variance;
  * when can we trust the asymptotic distribution?
- Choice of the tuning constants;
- Model selection;

For LMS and high breakdown point methods:
- How far is the exact solution from that obtained by the resampling algorithm?

For proposal *:
- How large is the (asymptotic) bias?

Proposal *: Start with LMS and fully iterate with a redescending M-estimator.

From R. Staudte, July 21

Robust estimation has spread throughout the world, but it is still largely undecipherable to most practicing statisticians.

A cynical outsider’s view of the workshop may be: “Before the workshop, statistics was a blackbox for producing a p-value. After the workshop, statistics is a shinier black box which produces two p-values. The user may choose the p-value which confirms his views, rejecting the robust one on the basis that the new methods are not to be trusted or rejecting the classical one because he wants to use the latest methods available.”

The statisticians’ view is that statistics is an iterative, interactive process. Their view will be increasingly replaced by the black box view as expert systems are developed, especially if such systems are not written by statisticians. Can statisticians come to any agreement what an expert system would do?

Perhaps at a future workshop we would work for a few days to determine whether statisticians under pressure would reach similar answers to specific questions, using the same data and other evidence of experts who provide it. (A few good solutions would also provide examples of analyses for students.) If there is consensus among statisticians, expert systems may be feasible.
"Clean-up", August 14

Some additional points which came up in the discussions:

a. Several participants felt the need for a kind of decision tree which would help to choose an appropriate robust method in a given situation. Peter Rousseeuw drafted such a tree as follows (in a graphical form, of course):

   (1) Fixed design (1) or observational $x$'s (2)?
   (11) Two-sample (11) or general (but balanced) $z$'s (12)?
   (12) $M_-, L_-, R_-$, or $L_1$-estimation (all are o.k.).
   (2) Generate plot of resistant residuals on resistant leverages (Mahalanobis type distance in $z$-space). If there are "no problems", continue with a (redescending) M-estimation.

   Rousseeuw was very positive about "forcing" the user to look at the diagnostic plot first.

b. There were different thoughts about how the robust methods should be presented (sold!). A prominent view said that they should generally be given along with the classical results, if possible side by side on the display. This means that the robust results must come in the same form. (What about the ANOVA-table?) They should be advertised as an add-on rather than an exclusive alternative to classical methods. This is how diagnostics found their way into packages very fast.

c. Inference is often neglected by robustniks. Model building is not always necessary. There may be only one or two explanatory variables, and the functional form may be known from previous studies. In two-sample problems (which are frequent in medical statistics) there is not much to explore beyond the appropriate re-expression (which is even irrelevant for rank methods).

d. Users have quite different views about models. Engineers usually take a very pragmatic point of view, and this seems to contrast often with economists. It may be easy to convince them that inference methods should be robust.

e. In large samples, efficiency is much less important than validity (that is, p-values which are correct or conservative are welcome, but they should never be too small).
5. QUESTIONS SUBMITTED TO A PANEL AND AN OPEN DISCUSSION

[Answers are not provided here.]

Questions posed to the panel, July 26, 1989

1. Are 50% breakdown estimates really necessary? Can’t I live with 10% breakdown? In particular, for multivariate covariance estimation, am I better off with a simple M-estimate than an MVE?

2. All robust estimation methods have tuning constants. How should they be chosen in practice (if one were interested in inference, not diagnostics)?

3. What are the relative merits / demerits of LMS plotting and LS diagnostics? Is masking a problem for either one? Should one use one or both approaches for data analysis?

4. With the exception of one talk, all the data sets used have had small n, small p. What should I use in terms of numerical diagnostics, graphical diagnostics and robust estimation / inference if n=2000, p=7?

5. Robust methods were criticized for “too many choices – too many criteria”. Isn’t the same becoming true for diagnostics? If I can have one robust estimation / inference method and one numerical / graphical diagnostic, what should I choose?

6. In a regression problem, if I were to use an LMS initial estimate followed by an M-estimate, what diagnostics are necessary and which diagnostic tools should I use?

7. Several authors have suggested that many of the case-deletion diagnostics are relatively simple functions of residuals and leverage. They suggest a plot of residuals versus leverage possibly augmented by contour lines for the different diagnostic functions. Is such a plot more attractive than numerical tables? Would a robustified version be safer (not failsafe, just safer)?

8. There are diagnostics for all sorts of aspects of model fit checking (outliers, heteroscedasticity, nonlinearity, etc.). Indeed, the number of such diagnostics has become a bit too large. Nevertheless, the common perception is that robust methods have been largely confined to estimation. Is this enough? Are there good robust analogues which will help me with the checking of model fit?

9. What do you think about the following statements:
   a. One-case or two-case deletion is not of much help, because you never know how many outlying observations there are (masking).
   b. Robust residuals, i.e., residuals from a robust fit, standardized using a robust estimate for the variance, do not tell me the whole story: They tell me that there are outliers, but not what is wrong (typing errors vs model misspecification).
   c. Graphical methods could tell me – if I could find the right projection – what went wrong; is there a group of gross outliers, do I need to add terms to the model, etc.

So as a data analyst I would like to have the best of both worlds: High breakdown robust methods which tell me that there is something wrong, and diagnostics (especially graphics) which tell me what is wrong.

10. If one starts at a high breakdown (say LMS) estimate, should I iterate just once or until convergence?

11. What are the properties of various regression estimators under departures from symmetry?

12. Why isn’t there a user-friendly robust package for DOS and/or UNIX machines? Or is there?

13. There have been a number of references to discussions during and after the Princeton robustness study. How can newcomers to the field get access to these discussions, which remain largely unpublished?
Questions for the open discussion of Aug 2

A. General area of questions: who are the consumers?
   (1) What do consumers really want and/or need?
   (2) Are the graphical presentations we develop for statisticians really meaningful to ultimate users?
   (3) Should we spend more effort developing (and/or tuning) tools for very specific purposes?
      If my experience is not atypical, the majority of data analyses are accomplished using rather special programs or packages (often, in fact, with little or no statistical analysis incorporated in the program). Shouldn’t we be spending much more time with our clients, finding out what they really want to do and producing appropriate specific tools for their purposes? Won’t this lead to rather different emphases and, if not different ideas, at least somewhat different means of presentation?

B. It seems to me that $2 \cdot 10^6$ (PC’s or users, which kind?) has been somewhat over-emphasized in John Tukey’s talk. Should we really direct our efforts, in the short run, to help all of them?

C. Some of John Tukey’s remarks seemed to push in the direction of a statistical expert system. Is it time to develop such systems?

D. What are the relations, analogies, syntheses, and basic differences of robust procedures and diagnostics?
   Let me try and summarize what has been said before:
   a) It is useful to examine diagnostics which make use of a robust fitting procedure. This avoids masking. See Peter Rousseeuw’s talk.
   b) Diagnostics (effect of case deletion as well as local influence) for robust estimators may be considered.
   c) A refined version of a) is exemplified by Sheather’s talk: What is the distribution of residuals from a robust fit? More generally, one may look at the effect of single observations (or other perturbations) on a non-robust estimator, starting from a robustly estimated model. (May need oral explanation)
      This is different from b).
   d) How could one make added-variable plots and other useful diagnostics for developing models robust? Does this make sense in principle? We certainly lose the geometrical interpretation for added-variable plots. Some work exists for robust transformation diagnostics, if I am not mistaken.
   e) Deletion diagnostics are versions of empirical influence functions, and these are well approximated by (asymptotic) influence functions. This leads to analogies like splitting Cook’s distance into the leverage and residual components and splitting the influence function into influence of position and influence of residual. Some of this seems to have been examined in Cook and Weisberg’s book.
   f) Are there further analogies (maybe even of a merely formal nature) which might be profitably transferred to “the other camp”? Can local influence diagnostics be approximated by aspects of asymptotic influence functions? Would this lead to new robust procedures?

E. On influence assessment: Cook and Weisberg (1982) discuss the external and internal norm approaches to assessing influence on parameter estimation. Very little has appeared on internal norm diagnostics since then. Roy Welsch and Dennis Cook indicated at Roy’s talk yesterday that they prefer the relative assessment of influence measure values as a batch over the use of cutoffs suggested by sampling theory. The internal norm approach would appear to be a natural way of extending relative assessment back to internally measure the changes in the parameter vector estimate. Has the lack of interest in the internal norm approach been due to computational requirements or are there more fundamental reasons? (Note: E-norm and i-norm calculations are comparable in cost for single cases.)

[End of Technical Report.]
<table>
<thead>
<tr>
<th>#</th>
<th>Author/s</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>516</td>
<td>P. Singh, Ph. Caussignac, A. Fortes, D.D. Joseph and T. Lundgren</td>
<td>Stability of Periodic Arrays of Cylinders Across the Stream by Direct Simulation</td>
</tr>
<tr>
<td>517</td>
<td>Daniel D. Joseph</td>
<td>Generalization of the Foscolo-Gibilaro Analysis of Dynamic Waves</td>
</tr>
<tr>
<td>518</td>
<td>A. Narain and D.D. Joseph</td>
<td>Note on the Balance of Energy at a Phase Change Interface</td>
</tr>
<tr>
<td>519</td>
<td>Daniel D. Joseph</td>
<td>Remarks on inertial radii, persistent normal stresses, secondary motions, and non-elastic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>extensional viscosities</td>
</tr>
<tr>
<td>520</td>
<td>D. D. Joseph</td>
<td>Mathematical Problems Associated with the Elasticity of Liquids</td>
</tr>
<tr>
<td>521</td>
<td>Henry C. Simpson and Scott J. Spector</td>
<td>Some Necessary Conditions at an Internal Boundary for Minimizers in Finite Elasticity</td>
</tr>
<tr>
<td>522</td>
<td>Peter Gritzmann and Victor Klee</td>
<td>On the 0-1 Maximization of Positive Definite Quadratic Forms</td>
</tr>
<tr>
<td>523</td>
<td>Fu-Chu Pu and D.H. Sattenger</td>
<td>The Yang-Baxter Equations and Differential Identities</td>
</tr>
<tr>
<td>524</td>
<td>Avner Friedman and Fernando Reitich</td>
<td>A Hyperbolic Inverse Problem Arising in the Evolution of Combustion Aerosol</td>
</tr>
<tr>
<td>525</td>
<td>E.G. Kalnins, Raphael D. Levine and Willard Miller, Jr.</td>
<td>Conformal Symmetries and Generalized Recurrences for Heat and Schrödinger Equations in One</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spatial Dimension</td>
</tr>
<tr>
<td>526</td>
<td>Wang Jinghua and Gerald Warnecke</td>
<td>On Entropy Consistency of Large Time Step Godunov and Glimm Schemes</td>
</tr>
<tr>
<td>527</td>
<td>C. Guillopé and J.C. Saut</td>
<td>Existence Results for the Flow of Viscoelastic Fluids with a Differential Constitutive Law</td>
</tr>
<tr>
<td>528</td>
<td>H.L. Bodlaender, P. Gritzmann, V. Klee and J. Van Leeuwen</td>
<td>Computational Complexity of Norm-Maximization</td>
</tr>
<tr>
<td>529</td>
<td>Li Ta-tsien (Li Da-qian) and Yu Xin</td>
<td>Life-Span of Classical Solutions to Fully Nonlinear Wave Equations</td>
</tr>
<tr>
<td>530</td>
<td>Jong-Shenq Guo</td>
<td>A Variational Inequality Associated with a Lubrication Problem</td>
</tr>
<tr>
<td>531</td>
<td>Jong-Shenq Guo</td>
<td>On the Semilinear Elliptic Equation $\Delta w - \frac{1}{2} y \cdot \nabla w + \lambda w - w^{-\beta} = 0$ in $\mathbb{R}^n$</td>
</tr>
<tr>
<td>532</td>
<td>Andrew E. Yagle</td>
<td>Inversion of the Bloch transform in magnetic resonance imaging using asymmetric two-component</td>
</tr>
<tr>
<td></td>
<td></td>
<td>inverse scattering</td>
</tr>
<tr>
<td>533</td>
<td>Bei Hu</td>
<td>A Fiber Tapering Problem</td>
</tr>
<tr>
<td>534</td>
<td>Peter J. Olver</td>
<td>Canonical Variables for BiHamiltonian Systems</td>
</tr>
<tr>
<td>535</td>
<td>Michael Renardy</td>
<td>A Well-Posed Boundary Value Problem for Supercritical Flow of Viscoelastic Fluids of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maxwell Type</td>
</tr>
<tr>
<td>536</td>
<td>Michael Renardy</td>
<td>Ill-Posedness Resulting from Slip As a Possible Explanation of Melt Fracture</td>
</tr>
<tr>
<td>537</td>
<td>Michael Renardy</td>
<td>Compatibility Conditions at Corners Between Walls and Inflow Boundaries for Fluids of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maxwell Type</td>
</tr>
<tr>
<td>538</td>
<td>Rolf Rees</td>
<td>The Spectrum of Restricted Resolvable Designs with $r = 2$</td>
</tr>
<tr>
<td>539</td>
<td>D. Lewis and J.C. Simo</td>
<td>Nonlinear stability of rotating pseudo-rigid bodies</td>
</tr>
<tr>
<td>540</td>
<td>Robert Hardt and David Kinderlehrer</td>
<td>Variational Principles with Linear Growth</td>
</tr>
<tr>
<td>541</td>
<td>San Yih Lin and Yisong Yang</td>
<td>Computation of Superconductivity in Thin Films</td>
</tr>
<tr>
<td>542</td>
<td>A. Narain</td>
<td>Pressure Driven Flow of Pure Vapor Undergoing Laminar Film Condensation Between Parallel</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Plates</td>
</tr>
<tr>
<td>543</td>
<td>P.J. Vassiliou</td>
<td>On Local Equivalence for Vector Field Systems</td>
</tr>
<tr>
<td>544</td>
<td>Brian A. Coomes</td>
<td>On Conditions Sufficient for Injectivity of Maps</td>
</tr>
<tr>
<td>545</td>
<td>Yanchun Zhao</td>
<td>A Class of Global Smooth Solutions of the One Dimensional Gas Dynamics System</td>
</tr>
<tr>
<td>546</td>
<td>H. Holden, L. Holden and N.H. Risebro</td>
<td>Some Qualitative Properties of $2 \times 2$ Systems of Conservation Laws of Mixed Type</td>
</tr>
<tr>
<td>547</td>
<td>M. Slemrod</td>
<td>Dynamics of Measured Valued Solutions to a Backward-Forward Heat Equation</td>
</tr>
<tr>
<td>548</td>
<td>Avner Friedman and Jürgen Sprekels</td>
<td>Steady States of Austenitic-Martensitic-Domains in the Ginzburg-Landau Theory of Shape</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Memory Alloys</td>
</tr>
<tr>
<td>549</td>
<td>Avner Friedman and Bei Hu</td>
<td>Degenerate Hamilton–Jacobi–Bellman Equations in a Bounded Domain</td>
</tr>
<tr>
<td>550</td>
<td>E.G. Kalnins, Willard Miller, Jr., and M.V. Tratnik</td>
<td>Families of Orthogonal and Biorthogonal Polynomials on the N-Sphere</td>
</tr>
<tr>
<td>551</td>
<td>Heinrich Freistühler</td>
<td>On Compact Linear Degeneracy</td>
</tr>
<tr>
<td>552</td>
<td>Matthew Witten</td>
<td>Quantifying the Concepts of Rate and Acceleration/Deceleration of Aging</td>
</tr>
<tr>
<td>553</td>
<td>J.P. Albert and J.L. Bona</td>
<td>Total Positivity and the Stability of Internal Waves in Stratified Fluids of Finite Depth</td>
</tr>
<tr>
<td>554</td>
<td>Brian Coomes and Victor Zurkowski</td>
<td>Linearization of Polynomial Flows and Spectra of Derivations</td>
</tr>
<tr>
<td>#</td>
<td>Author/s</td>
<td>Title</td>
</tr>
<tr>
<td>----</td>
<td>-------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>555</td>
<td>Yuriko Renardy</td>
<td>A Couette-Poiseuille Flow of Two Fluids in a Channel</td>
</tr>
<tr>
<td>556</td>
<td>Michael Renardy</td>
<td>Short wave instabilities resulting from memory slip</td>
</tr>
<tr>
<td>557</td>
<td>Daniel D. Joseph and Michael Renardy</td>
<td>Stokes’ first problem for linear viscoelastic fluids with finite memory</td>
</tr>
<tr>
<td>558</td>
<td>Xiaxi Ding</td>
<td>Superlinear Conservation Law with Viscosity</td>
</tr>
<tr>
<td>559</td>
<td>J.L. Ericksen</td>
<td>Liquid Crystals with Variable Degree of Orientation</td>
</tr>
<tr>
<td>560</td>
<td>F. Robert Ore, Jr. and Xinfu Chen</td>
<td>Electro-Optic Modulation in an Arbitrary Cross-Section Waveguide</td>
</tr>
<tr>
<td>561</td>
<td>M.V. Tratnik</td>
<td>Multivariable biorthogonal continuous-discrete Wilson and Racah polynomials</td>
</tr>
<tr>
<td>562</td>
<td>Yisong Yang</td>
<td>Existence of Solutions for a Generalized Yang-Mills Theory</td>
</tr>
<tr>
<td>563</td>
<td>Peter Gritzmann, Laurent Habsieger and Victor Klee</td>
<td>Good and Bad Radii of Convex Polygons</td>
</tr>
<tr>
<td>564</td>
<td>Martin Golubitsky, Martin Krupa and Chjan. C. Lim</td>
<td>Time-Reversibility and Particle Sedimentation</td>
</tr>
<tr>
<td>565</td>
<td>G. Yin</td>
<td>Recent Progress in Parallel Stochastic Approximations</td>
</tr>
<tr>
<td>566</td>
<td>G. Yin</td>
<td>On H-Valued SA: Finite Dimensional Approximations</td>
</tr>
<tr>
<td>567</td>
<td>Chien-Cheng Chang</td>
<td>Accurate Evaluation of the Effect of Diffusion and Conductivity in Certain Equations</td>
</tr>
<tr>
<td>568</td>
<td>Chien-Cheng Chang and Ruey-Ling Chern</td>
<td>The Effect of Viscous Diffusion in Discrete Vortex Dynamics for Slightly Viscous Flows</td>
</tr>
<tr>
<td>569</td>
<td>Li Ta-Tsien (Li Da-qian) and Zhao Yan-Chun</td>
<td>Global Existence of Classical Solutions to the Typical Free Boundary Problem for General Quasilinear Hyperbolic Systems and its Applications</td>
</tr>
<tr>
<td>570</td>
<td>Thierry Cazenave and Fred B. Weissler</td>
<td>The Structure of Solutions to the Pseudo-Conformally Invariant Nonlinear Schrödinger Equation</td>
</tr>
<tr>
<td>571</td>
<td>Marshall Slemrod and Athanasios E. Tzavaras</td>
<td>A Limiting Viscosity Approach for the Riemann Problem in Isentropic Gas Dynamics</td>
</tr>
<tr>
<td>573</td>
<td>P.J. Vassiliou</td>
<td>On the Geometry of Semi-Linear Hyperbolic Partial Differential Equations in the Plane Integrable by the Method of Darboux</td>
</tr>
<tr>
<td>574</td>
<td>Jerome V. Moloney and Alan C. Newell</td>
<td>Nonlinear Optics</td>
</tr>
<tr>
<td>575</td>
<td>Keti Tenenblat</td>
<td>A Note on Solutions for the Intrinsic Generalized Wave and Sine-Gordon Equations</td>
</tr>
<tr>
<td>576</td>
<td>P. Szolonyan</td>
<td>Heteroclinic Orbits in Singularly Perturbed Differential Equations</td>
</tr>
<tr>
<td>577</td>
<td>Wenxiong Liu</td>
<td>A Parabolic System Arising In Film Development</td>
</tr>
<tr>
<td>578</td>
<td>Daniel B. Dix</td>
<td>Temporal Asymptotic Behavior of Solutions of the Benjamin-Ono-Burgers Equation</td>
</tr>
<tr>
<td>579</td>
<td>Michael Renardy and Yuriko Renardy</td>
<td>On the nature of boundary conditions for flows with moving free surfaces</td>
</tr>
<tr>
<td>580</td>
<td>Werner A. Stahel</td>
<td>Robust Statistics: From an Intellectual Game to a Consumer Product</td>
</tr>
<tr>
<td>581</td>
<td>Avner Friedman and Fernando Reitich</td>
<td>The Stefan Problem with Small Surface Tension</td>
</tr>
<tr>
<td>582</td>
<td>E.G. Kalnins and W. Miller, Jr.,</td>
<td>Separation of Variables Methods for Systems of Differential Equations in Mathematical Physics</td>
</tr>
<tr>
<td>583</td>
<td>Mitchell Luskin and George R. Sell</td>
<td>The Construction of Inertial Manifolds for Reaction-Diffusion Equations by Elliptic Regularization</td>
</tr>
<tr>
<td>584</td>
<td>Konstantin Mischaikow</td>
<td>Dynamic Phase Transitions: A Connection Matrix Approach</td>
</tr>
<tr>
<td>585</td>
<td>Philippe Le Floch and Li Tatsien</td>
<td>A Global Asymptotic Expansion for the Solution to the Generalized Riemann Problem</td>
</tr>
<tr>
<td>586</td>
<td>Matthew Witten, Ph.D.,</td>
<td>Computational Biology: An Overview</td>
</tr>
<tr>
<td>587</td>
<td>Matthew Witten, Ph.D.,</td>
<td>Peering Inside Living Systems: Physiology in a Supercomputer</td>
</tr>
<tr>
<td>588</td>
<td>Michael Renardy</td>
<td>An existence theorem for model equations resulting from kinetic theories of polymer solutions</td>
</tr>
<tr>
<td>590</td>
<td>Luigi Preziosi</td>
<td>An Invariance Property for the Propagation of Heat and Shear Waves</td>
</tr>
<tr>
<td>591</td>
<td>Gregory M. Constantine and John Bryant</td>
<td>Sequencing of Experiments for Linear and Quadratic Time Effects</td>
</tr>
<tr>
<td>592</td>
<td>Prabir Daripa</td>
<td>On the Computation of the Beltrami Equation in the Complex Plane</td>
</tr>
<tr>
<td>593</td>
<td>Philippe Le Floch</td>
<td>Shock Waves for Nonlinear Hyperbolic Systems in Nonconservative Form</td>
</tr>
<tr>
<td>595</td>
<td>Mark J. Friedman and Eusebius J. Doedel</td>
<td>Numerical computation and continuation of invariant manifolds connecting fixed points</td>
</tr>
<tr>
<td>596</td>
<td>Scott J. Spector</td>
<td>Linear Deformations as Global Minimizers in Nonlinear Elasticity</td>
</tr>
</tbody>
</table>