Quality Reviews Comments from SAB Members on the SAB Draft Reports:
(1) Review of the All Ages Lead Model External Review Draft 2.0
(2) Technical review of EPA's 14 Computable General Equilibrium Model, SAGE

List of comments received
June 20, 2020

COMMENTS ON THE DRAFT SAB REPORT: REVIEW OF THE ALL AGES LEAD MODEL EXTERNAL REVIEW DRAFT 2.0 ..... **2**

Comments from Lead Reviewers **2**
Comments from Dr. Janice Chambers 2
Comments from Dr. John Guckenheimer 3

Comments from other Chartered SAB Members **8**
Comments from Dr. Rodney Andrews 8
Comments from Dr. Tony Cox 8
Comments from Dr. Otto Doering 9
Comments from Dr. Kimberly White 9

COMMENTS ON THE DRAFT SAB REPORT: TECHNICAL REVIEW OF EPA'S COMPUTABLE GENERAL EQUILIBRIUM MODEL, SAGE **10**

Comments from Lead Reviewers **10**
Comments from Dr. Otto Doering 10

Comments from other Chartered SAB Members **11**
Comments from Dr. Rodney Andrews 11
Comments from Dr. Hugh Barton 12
Comments from Dr. Tony Cox 12
Comments from Dr. John Guckenheimer 12
Comments from Dr. Kimberly White 13

OTHER COMMENTS ON THE DRAFT REPORTS PROVIDED BY SAB CHEMICAL ASSESSMENT ADVISORY COMMITTEE MEMBERS **15**
Comments from Dr. Dennis Paustenbach 15
Comments from Dr. Ted Simon 16
Comments from Dr. Eric Smith 16

Comments on the Draft SAB Report: *Review of the All Ages Lead Model External Review Draft 2.0*

Comments from Lead Reviewers

Comments from Dr. Janice Chambers

1. Were the charge questions to the Panel adequately addressed?

The charge questions seem to have been adequately addressed. The Panel had detailed answers to all of the questions. The Panel provided an analysis with literature citations to support their points as well as specific information on some of the points being made. The Panel provided evidence of trying to use the model, along with some of the difficulties encountered in these attempts. The Panel identified some deficiencies in the draft report as well as some data needs and some terms which should be clarified in the text. In addition, a few errors were pointed out for correction by EPA staff. Points of confusion in the current draft were identified with suggestions for clarification.

2. Are there any technical errors or omissions or issues that are not adequately dealt with in the draft report?

No error or omissions noted.

3. Is the draft report clear and logical?

Generally the report is clear and logical, but there are a few items that should be addressed for greater clarity and utility to the EPA staff who will use the report to improve the draft:

   a. Some of the recommendations are short and to the point, while others are relatively long and contain additional explanation of the reasoning behind the recommendation (examples include recommendations for Charge Questions 3c, 5, and 7); the latter explanations would be better placed in the text describing the Panel's deliberations preceding the recommendations. It would be more useful to EPA to have all the recommendations short so that EPA will have the Panel's guidance focused on single points within each recommendation. The explanations should be moved out of the recommendations and into the preceding text if they are not already in those sections, or deleted if the explanations are redundant with the preceding text.
   b. Two formats were used for the recommendations, some being declarative and others being imperative. One format should be selected and used throughout the report for consistency.

c. In a few cases the explanatory text used imperative statements that should be changed to be declarative statements for consistency; examples include pg 52, lines 2, 23 and 44.
d. Many EPA reports contain an Executive Summary but this one does not. An Executive Summary would be useful to EPA staff to provide the major points and most important recommendations in a single place. A summary of the major points was not included in the letter to the Administrator, and the major points might be useful in this letter as well.
e. The report should be carefully proofread and edited to make the entire report more consistent in format and more consistent with formal writing; examples include: elimination of contractions (e.g., don't); elimination of the use of the first person (e.g., we); providing consistency in line spacing and eliminating the large space on page 40 for better aesthetics; identification of any definitions of abbreviations that follow instead of precede the first use (e.g., definition of AF, absorption fraction, occurred on pg 29 after it was used on the preceding page).
f. While the report was well written and essentially free of typographical errors, the use of "verse" on pg 50 line 21 probably should have been "versus".

4. Are the conclusions drawn or recommendations provided supported by the body of the draft report?

The conclusions and recommendations seem to be reasonable and supported by the body of the report.

Comments from Dr. John Guckenheimer

Quality Review Questions (answered at the end of this assessment)

1. Were the charge questions to the Panel adequately addressed?
2. Are there any technical errors or omissions or issues that are not adequately dealt with in the draft report?
3. Is the draft report clear and logical?
4. Are the conclusions drawn or recommendations provided supported by the body of the draft report?

The All Ages Lead Model v2 (AALM) estimates lead concentrations in blood, bone and other body organs resulting from changing lead concentrations in the environment. The model evolved from three models created in the early 1990's: the IEUBK model for lead concentrations in children ages 0-7, a model Leggett developed at Oak Ridge National Laboratory in connection with studies of biological accumulation of radionuclides , and a physiologically based model developed by O'Flaherty, initially fit to experimental data on rats. The original AALM fortran

model has been modified over time, reimplemented in the proprietary acslX language and given an excel interface for users. Both this review panel and previous review panels found it difficult, at best, to run the model and verify that its output matches the Technical Support Document (TSD). Despite these difficulties, this review panel and previous ones judged the AALM to be a good tool that should be supported and further improved. I find this conclusion questionable and present here a more critical analysis of the AALM which addresses the quality review questions listed above.

Charge question 8 for the Review Panel asks whether the AALM is consistent with EPA Regulatory Environmental Model Guidance (cred_guidance_0309.pdf). This EPA guidance document reminds us that the scientific context and objectives of models should be stated clearly in plain language. As noted by the Review Panel (pp.62-63) the AALM does not make such a statement. I looked for existing or potential EPA applications of the software that might highlight its usefulness. I found no documented applications, but here are four potential ones:

1. Estimate how much lead accumulates in varied body tissues from lead in drinking water and how quickly it is depleted when lead concentrations in water are reduced.
2. Estimate how much lead accumulates in varied body tissues from lead dust in houses containing lead paint and how quickly it is depleted when the lead paint is removed.
3. Estimate how quickly lead is depleted in body tissues of individuals who accumulated high lead levels through long time exposure in their work environments.
4. Analyze the significance of lead radio nucleotides as carcinogens in the aftermath of nuclear accidents, Chernobyl being a particular important case that motivated Leggett's work.

Most lead in the body accumulates in blood and bone. The health effects of lead concentrations in these compartments and other organs are outside the scope of the model. Measurements are difficult, so there is relatively little data available to parameterize the AALM. The biokinetics of lead does not appear to be a very active research area: the most recent of the twelve references cited by the Review Panel (Section 2.4.3) as suited for model evaluation is from 2005 and the next most recent is from 2001. I found no data from "high throughput" measurements of lead in the body. Thus, time series of lead concentrations have large intervals between measurements. Nonetheless, the scope of the AALM is broad and has the objective of fitting lead concentrations in several different body. There are no quantitative metrics to assess model fits or even a discussion of the accuracy authors and reviewers expect. Attempts by reviewers to implement simulations of the AALM for comparison with data and with other models have encountered missing information as well as technical problems in running the models successfully. When the AALM has been used to simulate different data sets, fits required ad hoc changes in parameters. Still, prior reviews have not questioned the conceptual structure of the model and have expressed

only modest frustration about how hard it was to run the software. If the model is to be used as a tool for regulatory purposes, we need to be clear about its accuracy and reliability.

Dynamic models like the AALM create time series of interacting variables. The simplest models are either systems of ordinary differential equations that express the rates at which state variables change, or discrete time iterations that give rules for how the state variables change in a single time step. Much more complicated models are also possible, for example hybrid models that include both discrete and continuous time phenomena, and stochastic models that track evolving probability distributions of variables rather than deterministic values. The core of the AALM is a system of differential equations. Solutions of the equations are visualized as trajectories evolving in a multidimensional state space. Numerical integration algorithms are used to compute approximate trajectories step by step. An inherent feature of this process is that errors can accumulate so that the accuracy of the computed trajectory diminishes in time, typically at an exponential rate. The development of numerical integration algorithms and their error analysis is a long standing, mature research area in applied mathematics. A weak point of the AALM is that it uses poor, ad-hoc methods for numerical integration (section 2.3.1 of the TSD). This deficiency has been noted repeatedly in reviews, but fixing it requires a thorough reworking of the Leggett model. The acslX version of the model was one attempt to do that, but the language is proprietary and obscure.

The numerical integration issue is a fundamental weakness of the AALM. Abstractly, differential equation models have the form $x'=f(x)+i(t)$ where x is a vector encoding the state variables (e.g., amounts of lead in each body compartment), f is the "right hand side" which expresses the rates of change in the state variables and i(t) is a vector of inputs (lead from the environment). Approximate trajectories are computed in discrete time steps. Numerical integration algorithms require the user to code the right hand side which evaluates f(x) and specify initial values of x at the start time and the input functiont i(t). The AALM is not organized in this way. It incorporates an explicit low order formula (Eq. 2.3-2 in the TSD) for its time stepping that is intermingled with the formulas that evaluate the right hand side. This yields unacceptable run times for simulations having sufficient accuracy, prompting further ad-hoc modifications to the software. Furthermore, the current version of the AALM buries the integration algorithm beneath an excel spreadsheet that calls fortran libraries with a run-time library that is specific to the visual basic language used by excel The review panel attempted to extract f(x) for use in a matlab reimplementation with standard numerical integration algorithms, but they encountered difficulties in making this work. Almost all modern computational science software includes fully implemented and documented examples that can be used to test the software on different computers. The lack of such examples and the arcane computing environment required by the AALM are serious deficiencies. I do not think further efforts to remedy these deficiencies are warranted because the conceptual foundations of the Leggett model are also weak.

The Leggett model begins with the assumption that the right hand side of the model is linear. This limits the number of parameters required to specify the model and yields solutions given by sums of exponential functions. Engineering systems are often linear. When data gives poor fits to the engineering models, a common strategy is to add more state variables that yield more exponential time constants. For the AALM, this results in splitting single compartments into multiple compartments solely for the purposes of fitting the data. However, biokinetic processes are seldom linear, leaving the Leggett model with weak scientific foundations. The models created by O'Flaherty are much better in this regard, but the AALM still follows the Leggett approach at its core. When time series of lead concentrations are not fit by single exponentials, new compartments are created to yield sums of exponentials with different time constants. No pretense is made to establish biological principles for the extended models. Thus, the Leggett model can be viewed as "data driven" with a tenuous relationship between model variables and identified physical quantities.

Would the resources of the EPA to be better used to create a new model rather than to improve and extend the AALM as recommended by the Review Panel? The deficiencies described above and additional limitations noted by the Review Panel are strong arguments for reimplementation. Little would be lost scientifically by abandoning the AALM because there are few, if any, applications of the AALM to studies of lead accumulation and retention in the scientific literature. I note that the AALM played no role in the recent SAB report on lead in water systems.

The accuracy of the AALM as tool to estimate the uptake, retention and excretion of lead by the body is hardly clear, so it does not seem ready for use in the context of EPA rule making. For example, the EPA is reconsidering the "action level" for replacing lead pipes and fittings in public water systems to reduce the toxic effects of lead concentrations in children. Since most of the lead in the body is retained in blood and bones, the EPA might use the AALM to estimate the correlation of lead concentration in people with the lead concentrations in an individual's environment. However, there is still too much uncertainty in the model to trust its results. Reference data sets should have been chosen and used for quantitative evaluation of the model with regard to its regulatory objectives. Moreover, as described by the Review Panel, comparison of model output and data ideally would take account of individual variation, stochastic fluctuations and measurement errors. I see no evidence that the AALM can be modified with reasonable effort to meet these objectives. If the SAB agrees with this conclusion, then we should make alternative recommendations to the. Here are a few suggestions.

General modeling principles recommend using models of minimal complexity that capture the essential aspects of the system being studied. Ideally, a lead model should incorporate scientific understanding of the key processes in its biokinetics. The O'Flaherty model could provide a starting point that focuses upon the interchange of lead between blood, plasma, trabecular bone

and cortical bone within the body,. Systematic analysis of models that are not too large or complicated can quantify the sensitivity of model output to changes in model inputs and parameters. Such analysis may also identify aspects of the system that a small model is unable to reproduce and point toward improvements. Even with small models, the number of quantities to be measured or estimated is large enough that ad hoc "tuning" of model simulations is unlikely to find optimal fits of model and data. Systematic sensitivity analysis is usually a more effective strategy. This is likely the case for modeling lead concentrations in the body.

Computers and computer software have changed dramatically during the twenty five years since the AALM was first created, but the evolution of the model has taken relatively little advantage of these improvements. Open source software (including packages specifically designed for investigation of dynamical models) and high level programming languages like Python, Julia and Matlab (or its open source alternative Octave) make it far easier to implement models of moderate complexity and enable much more extensive exploration of their properties. The deficiencies of the AALM strongly suggest that the EPA stands to benefit far more from creating a new model to simulate the biokinetics of lead than by investing more resources into the AALM. One possible strategy is for the agency to support an interdisciplinary team that will create and implement a new model while simultaneously conducting empirical studies to calibrate and test their model. Joint programs of the NSF and NIH to support collaborations that span the interface between mathematics and biology have demonstrated the effectiveness of this strategy in quantitative modeling of biological systems. Based on this analysis, here are answers to the questions that I was asked:

1. Were the charge questions to the Panel adequately addressed?

Most of the charge questions were adequately addressed. However, the charge questions and the Panel responses hardly address fundamental issues about the utility and quality of the AALM.

2. Are there any technical errors or omissions or issues that are not adequately dealt with in the draft report?

The weakness of the scientific foundations and the numerical methods in the AALM should have been given more emphasis in the report of the Review Panel. The report does a good job in pointing to extensions of the software needed to deal with random effects and systematic comparisons of the software with data .

3. Is the draft report clear and logical?

Yes.

4. Are the conclusions drawn or recommendations provided supported by the body of the draft report.

No. Individual items in the report identify many shortcomings in the AALM software. The Panel was charged to review the draft technical support document rather than the software itself. Nonetheless, they made recommendations for improving the software. The poor numerical and scientific foundations of the AALM support the conclusion that the EPA should create new software tools for analyzing lead in the body rather than investing further effort in the AALM.

Comments from other Chartered SAB Members

Comments from Dr. Rodney Andrews

1. Were the charge questions to the Panel adequately addressed?

Yes.

2. Are there any technical errors or omissions or issues that are not adequately dealt with in the draft report?

None that I could discern as a non-SME. The statements made in the report are clearly supported by either literature citation or by inclusion of data or example.

3. Is the draft report clear and logical?

Yes. The report was clear and logical, often including specific examples or references.

4. Are the conclusions drawn or recommendations provided supported by the body of the draft report?

Yes.

Comments from Dr. Tony Cox

Preliminary Comments on *SAB Review of the All Ages Lead Model External Review Draft 2.0*

1) Were the charge questions to the Panel adequately addressed? Yes. This is a clear, technically cogent review.

2) Are there any technical errors or omissions or issues that are not adequately dealt with in the draft report? Not that I found.
3) Is the draft report clear and logical? Yes.
4) Are the conclusions drawn or recommendations provided supported by the body of the draft report? Yes, they are well supported. This is a good piece of work.

Comments from Dr. Otto Doering

1. I believe that the charge questions to the Panel were adequately addressed.
2. I did not find technical errors or omissions or issues that were not adequately dealt with in the draft report.
3. I believe that the draft report is clear and logical. Good organization of responses after charge questions and very helpful prioritization of recommendations.
4. I believe that the conclusions drawn and/or recommendations provided were supported by the body of the report.

Comments from Dr. Kimberly White

1. Question: Were the charge questions to the Panel adequately addressed?
Answer: The All Ages Lead Model Review Panel was tasked with addressing nine charge questions. The draft report is organized to address each charge question and recommendations are provided based on priority for completion (i.e. tier I, II or III). All questions appear to be adequately addressed in the draft report.

2. Question: Are there any technical errors or omissions or issues that are not adequately dealt with in the draft report?
Answer: I did not identify technical errors, omissions or issues that were not adequately addressed.

3. Question: Is the draft report clear and logical?
Answer: The draft report appears clear and logical. The draft report could benefit from shortening of some of the text associated with each charge question. For example, in several areas throughout the report there are paragraphs proceeding the recommendations which appear to recap or introduce several of the recommendations that follow those sections.

4. Question: Are the conclusions drawn or recommendations provided supported by the body of the draft report?
Answer: The conclusions and recommendations appear to be supported by the body of the draft report.

<u>Comments on the Draft SAB Report: Technical Review of EPA's Computable General Equilibrium Model, SAGE</u>

Comments from Lead Reviewers

Comments from Dr. Otto Doering

A Lead Review of the 'SAB Technical Review of EPA's Computable General Equilibrium Model, SAGE'
Otto Doering
Purdue University

Overall I believe this is an excellent and most constructive review that should be most valuable for EPA in the development of this new analytical capacity. I come away from reading it with admiration for the work of the review panel.

1. I believe that the charge questions were adequately addressed by the panel. EPA's staff provided charge questions that were excellent and most relevant to assessing and improving the model which helped the review team to be most effective in its review.

2. I did not identify important technical errors or omissions or issues that were inadequately dealt with in the draft report.

3. I found the draft report both clear and logical. This was driven by the clear charge questions and then the review committee's excellent job of organizing their responses into the relevant sub-topics. The prioritization of recommendations under these sub-topics added to the clarity and logic of the review committee's response and should aid EPA in directing its resources towards improving the model.

4. The review team provided clear concise support for the recommendations in both theory and practice.

Comments:

First, I believe that the recommendations and the priority given to the recommendations are well targeted to key characteristics of CGE models and what should be expected of them.

It is important that EPA recognizes that the review panel's Tier 1 Modeling Recommendations are essential if the model is to do what a GCE model is intended to do - i.e. capturing the aggregate welfare and/or distributional impacts of a policy while taking cross-price and cross-

market effects into account and being driven by a framework of consumer and producer maximization. The three Tier 1 modeling recommendation areas highlighted in the transmittal letter are particularly important for this task. There is one recommendation that I might elevate to tier 1 status. That is recommendation CQ10-9 on page 23, Allow Imperfect Competition. The assumption of such things as perfect competition in CGE models is violated in the real world by government price setting or deal making and monopoly situations. The model should be prepared from the outset to consider this situation as an important sensitivity analysis scenario. I also support a related concern of the review panel reflected by the Tier 1 recommendation on page 13 to Relax the Small Open Economy Assumption of the model.

Last, when one constructs and operates a CGE model, ones bias may be to fervently believe it. This amplifies the importance of getting it right as much as possible at the outset, approaching it's results with healthy skepticism, and creating the flexibility to run different scenarios based on very different basic economic and behavioral assumptions.

Comments from other Chartered SAB Members

Comments from Dr. Rodney Andrews

Comments on the Technical Review of EPA's Computable General Equilibrium Model, SAGE

1. Were the charge questions to the Panel adequately addressed?

Yes. Charge Question 4 was very broad and it would be unrealistic for the panel to address all possible outcomes; however, the response was as adequate as possible within that context. The response to Charge Question 10 may go beyond the request for near term improvements, but is well within the scope of the overall report.

2. Are there any technical errors or omissions or issues that are not adequately dealt with in the draft report?

None that I could determine as someone without specific expertise in this area.

3. Is the draft report clear and logical?

Yes. I found both the text and the organization very clear and logical.

4. Are the conclusions drawn or recommendations provided supported by the body of the draft report?

Yes.

Comments from Dr. Hugh Barton

Comments on the Technical Review of EPA's Computable General Equilibrium Model, SAGE

1) Were the charge questions to the Panel adequately addressed? No answer as this is not my field.

2) Are there any technical errors or omissions or issues that are not adequately dealt with in the draft report? No answer as this is not my field.

3) Is the draft report clear and logical? Yes.

4) Are the conclusions drawn or recommendations provided supported by the body of the draft report? Yes

Comments from Dr. Tony Cox

Preliminary Comments on *SAB Technical Review of EPA's Computable General Equilibrium Model, SAGE*

- Model validation is not discussed much, but is crucial for assessing the value of the information provided by SAGE (or other CGE or AGE models). (https://www.researchgate.net/publication/304452270_Catching_up_with_history_A_methodology_to_validate_global_CGE_models)
- How should SAGE be updated in response to events such as the COVID-19 pandemic?

Comments from Dr. John Guckenheimer

Assessment of the two SAB Panel Report on Computable General Equilibrium (6-4-20)
John Guckenheimer

1. Were the charge questions to the Panel adequately addressed?

The Panel Report gives thorough answers to the charge questions, including many recommendations for changes and additions to the SAGE model.

2. Are there any technical errors or omissions or issues that are not adequately dealt with in the draft report?

I did not identify any. However, further assessment of numerical issues associated with SAGE would have been helpful. The Report implicitly assumes that the software just runs.

3. Is the draft report clear and logical?

Yes.

4. Are the conclusions drawn or recommendations provided supported by the body of the draft report?

Yes.

Comments from Dr. Kimberly White

SAB draft report titled "Technical Review of EPA's Computable General Equilibrium Model, SAGE"

General Comments: The Computable General Equilibrium Model, SAGE, is outside my area of expertise but I have reviewed the draft report and addressed the quality questions below.

1. Question: Were the charge questions to the Panel adequately addressed?
Answer: The CGE Model Review Panel was charged with addressing nine questions related to three topics: (1) technical accuracy and defensibility of the model, (2) approach to model versioning and peer review of future model updates and (3) priorities for improvements of the model framework. All the charge questions appear to be addressed and recommendations for improvement, as applicable, are included.
2. Question: Are there any technical errors or omissions or issues that are not adequately dealt with in the draft report?
Answer: I did not identify technical errors, omissions or issues that were not addressed in the draft report.

3. Question: Is the draft report clear and logical?
Answer: The draft report appears to be clear and logical.

4. Question: Are the conclusions drawn or recommendations provided supported by the body of the draft report?

Answer: The conclusions and recommendations are supported by the body of the draft report and the authors have presented their recommendations in three tiers, thus offering the Agency priorities for implementation of the recommendations.

Other Comments on the Draft Reports Provided by SAB Chemical Assessment Advisory Committee Members

Comments from Dr. Dennis Paustenbach

Review of EPA's "All Ages Lead Model" (External review Draft II)

I have reviewed the Charge Questions and the Responses to Charge Questions that were distributed on June 4, 2020. I have had quite a lot of experience with the various models used to predict blood lead concentrations based on classic input parameters.

I have reviewed many SAB reviews of EPA work and I found this document, signed by Hugh Barton, to be among the top three evaluations that I have read over the past 30 years.

I can only say that it would be terrific if the Agency can accommodate all of the recommendations.

It is extraordinarily well-done.

Review of EPA's Computable General Equilibrium Model (Sage)

I have reviewed the Charge Questions and the Responses to Charge Questions that were distributed on June 4, 2020. Because this body of work is not one in which I have much experience, my comments are limited.

1) Charge question 1: I agree with the recommendations offered on Charge question 1. I agree that it should be more approachable for the readers/users who don't work in this space. The other comments appear reasonable.

2) Charge question 2: I agree with the panel's recommendations regarding Model Structure and Assumptions. The specific comments should be weighed by the Agency

3) Charge question 3: I agree with the recommendations.

4) Charge question 4: No comments

5) Charge question 5: I agree with the recommendations.

6) Charge question 6: No comments

7) Charge question 7: No comments.

Comments from Dr. Ted Simon

Comments on the draft SAB report, *Review of the All Ages Lead Model External Review Draft 2.0*

Page 9, line 12. A useful way to present the context and applications of the model is as a domain of application, as is done for assays or testing platforms
Page 9, line 39. How do these O'Flaherty growth curves compare to recent NHANEs data?

Page 11, line 11. Region 8 used a probabilistic lead model in the 1990s. Perhaps some aspects of that model could be used to ameliorate the proposed AALM. Phil Goodrum, an SAB member, developed that model and will probably have much to add.

Page 20, line 42. Bone remodeling is subject to hormonal influences and these influences account for mobilization of lead from bone stores in post-menopausal women. Since nursing mothers are exposed to a variety of hormonal influences, the underlying science may be quite complex and consideration of such no easy task.

Page 37, line 3. This section partially addresses my previous comment.

Page 42, line 42. Having the model available in open source software such as R or Julia would be a good suggestion. ACSLXtreme is no longer available but the package Magnolia is very similar and available without charge at https://www.magnoliasci.com/. Magnolia was developed by Conrad Housand, who also developed ACSL. MATLAB is hugely expensive. Excel, although widely available and widely used, is a poor modeling platform for time-dynamic models involving ODEs. A version in R would address the Apple/PC divide as well. EPA has many staffers highly knowledgeable in R, esp. in the NC CompTox group. They might be able to provide additional insight.

Page 51, line 40. Whilst I've not had a chance to use the software, testing with various numerical methods including those mentioned here. Excel

Comments from Dr. Eric Smith

E.P. Smith comments on **TECHNICAL_SUPPORT_FOR_AALM_2.0-MAY2019**

1. Equation 3.4 has an intercept term of -3.89. This does not seem correct (note figure 3.13 seems to go through the origin). Although there are 406 measurements on

individuals, the number of values in the graphs are not 406, please give the number of points or groups.

2. Equation 4.3 is not f'(x) rather it is f'(x)/f(x) / dx/x. Can f'(x) be dropped?
3. Equation 4.6 defines "residual". This definition is not consistent with standard statistical definitions that define residual as "observed – predicted". Perhaps it was intended to call this a standardized residual (as in equation 3.1) as it is a residual divided by standard deviation. It also seems unusual to divide by the standard deviation of the mean or equivalently the standard errorThe criteria that all residuals be between -2 and +2 seems restrictive given this definition since as sample size becomes large the denominator should get small. The criteria that r^2>= 0.7 seems low (is there a reference for this rule?), especially if there is any tuning of the model. So
    a. Define the denominator with a formula. Perhaps more should be discussed about the data reduction of the 406 measurements (how many groups, etc). The denominator is probably the standard deviation of the means of the observed values. Consider just calling it the standard deviation of observed values.
    b. Rethink the formula - Normally one would divide by the standard deviation of the numerator.
    c. Comment that the average of the residuals will not be zero but will reflect the average bias in the model.
    d. A useful reference is below (Pineiro et al., 2008)
    e. Note that Equation 3.1 is for a standardized residual but still might need a correction to the denominator's name.
4. The figures associated with the residuals: Figure 4-15. Illustrates possible confusion related to the definition of residuals. The top right figure gives a plot predicted versus observed (note that this is not observed vs expected as would be suggested by the formula for the residuals). The bottom two figures are associated with residuals and shows values between -2 and +2. The figure is atypical in that the values are not centered around zero as one might expect. The bottom left plot shows bias for the AALM-LG model. This is indicated by the value of the intercept of regression model: -1.01. Similarly, the other model has a bias of 1.64. The relevant questions seem to me to be (1) is the intercept different from zero (ie no bias)? and (2) is the slope different from 1.0? These can be addressed either statistically through a test or by a pre-determined criteria (I.e. is being off on average by 1 unit relevant?). Also, the pattern in the residuals in the second plots suggest a nonlinear relationship.

    Reference: G. Piñeiro, S. Perelman, J.P. Guerschman, J.M. Paruelo
    How to evaluate models: observed vs. predicted or predicted vs. observed?
    Ecol. Modell., 216 (2008), pp. 316-322, 10.1016/j.ecolmodel.2008.05.006
5. Equation 2.2-1 and others. Shouldn't *n* be defined for completeness?
6. Page 19 line 1. dYj/dt is defined as the change in Pb mass in compartment *j* over time *t*, but is should be over time dt.

7. Page 56 line 6. Why specify $r^2>0.8$ when you can give the exact value?

Miscellaneous
Page 81 line 6 observation is misspelled
Page 61 line 23 (r^2 0.81). should be ($r^2 = 0.81$)
Page 65 Line 23 refers to 409 adults not 406 as in line 26 page 59
Figure 3.14 indicates n=25 and n=7 however only 4 circles and 4 squares are displayed. Are there multiple values that are not displayed? Perhaps consider jittering the points.
Figure 3.15 Again specify number of means that are displayed. Are the bars standard errors? Why in 3.13 are there error bars in both directions but in 3.15 they are only in the y direction?
Figures 3.16, 3.17 – it is often useful to also plot the 1-1 line to get a sense of deviations.
Figure 3.17 indicates the model is biased as the intercept is 2.3 (same for 3.18)