Brandt et al.: Assessment of Past, Present and Future... …using the EVA Model System.

Full review, 01 July 2013: Fintan Hurley, IOM, Edinburgh fintan.hurley@iom-world.org

Introductory remarks

First, I apologise to the authors for the long delay between my receiving this paper and preparing this full review.

As requested, I prepared a quick review of the paper in I think January 2013. I don't know if the authors have seen this. I include it as an Annex.

I am unfamiliar with the Journal's interesting and public peer review process and so I didn't understand that occasional requests from the editors, that I contribute to web-based discussions, were in fact requests to complete and post a full review.

I finally understood a few weeks ago that I was to write a full review; here are my detailed comments. They reflect my experience, i.e. within the overall impact pathway approach of this paper, I am not really familiar with the modelling of emissions and concentrations; my expertise is principally in methods for estimating health impacts and how they link with other components, including associated monetary values. In particular, I led the Methodology development of CAFE cost-benefit analysis though I haven't re-visited it in preparing these comments. Nor have I looked at the comments already on the website, including those of the other referee.

SUMMARY OF COMMENTS

This paper has some interesting and quite important results about the extent to which emissions from shipping contribute to air pollution and consequent disease and mortality in Europe. Methods and results are reported on how emissions from shipping (in the Northern Hemisphere, in SECA region) contribute to air pollution concentrations across Europe. Results are given on consequences for health, in terms both of health effects and monetary values. However, corresponding methods (for estimation of health impacts, for monetary valuation) are not given, certainly not systematically. This is a serious limitation which needs to be addressed. Comparison of results with CAFE, without comparison of methods, is no substitute for describing and discussing Health Impact Assessment (HIA) methods properly. The issue is particularly relevant because the EVA Model is applied to Europe as a whole; consequently the paper needs to address whether a common methodology will work between EU and elsewhere Europe in terms of e.g. "exposure-response" functions; background rates of mortality and morbidity; monetary valuation; and if so, what that common methodology might be. There are various other things that need to be addressed.

I think that to have a coherent paper the authors have either of two principal options.

a. Maintain the original vision of the paper but strengthen it very substantially in terms of describing and evaluating methods for HIA and monetary valuation;
b. Focus on emissions from shipping and their effect on air pollution concentrations across Europe, without expressing those implications in terms of health effects and monetary values.

DETAILED COMMENTS

Overview comments – what is new and useful here?

This paper has some interesting and quite important results about the extent to which emissions from shipping contribute to air pollution and consequent disease and mortality in Europe. In particular, it identifies and highlights the benefits in pollution reduction from policies to reduce sulphur emissions from shipping in the Baltic and in the North Sea (SECA Area).

This leads to a proposal, which sounds good and reasonable (though I don't know how difficult it would be to implement) to extend these low-sulphur fuel measures to shipping elsewhere in the Northern Hemisphere, on the grounds that health gains would be significant. The likely benefits to population health in Europe of such a measure could be estimated afresh using the EVA Model. The paper provides a basis for estimating these approximately assuming that the measures internationally would give reductions in resulting population-weighted air pollution *pro rata* those that have come from the SECA measures.

The authors use non-linear atmospheric chemistry modelling of the fate of emissions from particular sources, which they say is an improvement on the linear modelling used in GAINS / RAINS / CAFE – I'm not competent to judge but I'm happy to take the authors' word for this. (Non-linearity might put in question any assumptions of *pro rata* reductions – is that why the authors didn't do this explicitly?)

The authors use the 'impact pathway' approach – I agree. As best I can tell, the paper does not develop any new methodology for this but applies the methodology, as implemented via the EVA Model, to emissions from a particular pollution source, in this case shipping. The main novelty methodologically seems to be "tagging" (p5), though other recent papers from the same team have explained this.

The value of the paper is therefore in the correctness and importance of the application. Because of lack of detail on methods, those unfamiliar with the EVA Model need either to trust that it is 'fit for purpose' for the present applications, or do a wide amount of background reading which I haven't done.

Assuming that the results given here are usable to inform development of policy, the discussion of policy implications seems to me generally sound (though the role of shipping is, I think, somewhat overstated).

On that basis, I think this paper is potentially a useful contribution to discussions on how to reduce outdoor air pollution in Europe and in particular to the role of better controls on international shipping.

Structure of the paper

In principle the structure is good; in practice far too many of the detailed methods are either unreported or appear as asides in the commentary on results, and that is much too late.

I appreciate that it is not easy to judge well how much detail on general methods to include in a specific paper from a longer series but, while I have focused on the practical implications of the results, the authors also have a methodological purpose (c/f extensive comparisons

with CAFE; and the final sentence p16, lines 18-20). The comparisons with CAFE really need a lot more information about methods. In fact, there is too little detail on methods even for a substantive paper. Methods on HIA and on monetary valuation need to be reported much more clearly and more fully, early in the paper, where Methods are usually to be found; and their strengths and limitations discussed later. Then the reader can understand what's behind the results that otherwise we are asked to take on trust; and the methodological comparisons with CAFE can be much more incisive.

NB – I do not take CAFE as a gold standard. I think its methodology was good, but it can be improved. However, without description of methods, we don't know the similarities and differences.

Limitations in description of methods – and perhaps in the methods themselves

My main concern is that description and discussion of methods are largely missing. I support the authors' ambition of using the "best available and most accurate" methodology (p4, line 30) throughout the impact pathway chain, even if that is computationally demanding. I do not know how well that has been achieved but if the authors have sufficient grasp of the issues to develop a leading edge methodology, it is strange that several important aspects of this have been neither described or discussed.

As I see it, estimation of health impacts involves integrating four kinds of information:

a. The population exposed / at risk – in this paper "Europe" (undefined, but we get an idea from most of the Figures, and some text in finally 14, line 13 – this needs to be defined clearly, early) and "Denmark".
b. The relevant pollution levels experienced by that population. These vary by scenario (c/f Table 1); the non-linear (with "tagging") methodology is described in at least some detail; and results are given (c/ Figures) for the effects of several scenarios – enough for the reader to get a feel for what is or isn't going on.
c. Concentration-response functions (CRFs – more accurate than the authors' name of exposure-response functions) – these are not given, not even the pollutant-outcome combinations. Tables 2-4 show us what health outcomes are used – but which pollutant-outcome pairs are used? And which CRFs?
d. Finally, implementation needs to use background rates of mortality and morbidity, often age-specific (depending on the CRF). These are not even mentioned. (In CAFE we integrated CRFs and background rates into 'impact functions', for ease of implementation; but the component parts need to be described separately.)

Given the comparisons with CAFE, and the similarity of health outcomes, I assumed the CRFs might be the same as CAFE. But I recognise at least two differences: cardiovascular hospital admissions are sub-divided here, though not in CAFE; and there is reference to "exposure-response" functions for SO2, not included in CAFE core.

So what was done, and why? If CAFE is a kind of benchmark, what's similar, what's different here from CAFE?

- What functions in SO2 are used?

- What are the possibilities of double-counting, when aggregated across pollutants or "compounds" (Table 5)? Chances are not much with just PM and O2, as in CAFE core. But with SO2 also?
- And especially: CAFE was designed for applications EU-wide, not Europe-wide, with a population about twice that of Europe, the difference primarily from including countries of Eastern Europe (Ukraine, Belarus, Turkey, parts of Russia are mentioned – were the same CRFs used Europe-wide? How applicable are they?
- How were YOLL estimated from cohort studies – how well does that work for Europe as a whole? What account, if any, of cessation lag?
- And YOLL from time series studies – are these from effects of ozone only? And calculated how?
- Did ozone quantification use the recommendation from WHO TFH, and implemented in CAFE, of quantifying only at daily 8-hr max concentrations higher than 35 ppb?
- And other questions…

And with monetary valuation, there are further issues

e. What monetary values were used? Same as CAFE, or different? Same major question – how applicable to countries in Eastern Europe? The actual valuations used Europe-wide need to be considered; and the issue of whether or not to apply the same values Europe-wide surely needs discussion. (From this paper I have no idea of what was done.)
f. The variation in CAFE results (Watkiss et al.) comes largely from 4 different methods of monetary valuation – based on Value of a Statistical Life (VSL) or Value of a Life Year (VOLY), and using mean or median values. In CAFE, VSL results were higher than those using VOLYs; and results based on mean values were higher than those using means. Yet in this paper, comparisons with CAFE were made using VOLY results from EVA and VSL results from CAFE. Why? In my opinion this methodological difference makes a substantial contribution to the CAFE results quoted being higher than the EVA results (per unit population).
g. And finally – what, if anything, has been done about Discounting? And why?

So, many questions. All are important for understanding what the results mean, especially in applying the methodology / EVA Model to Europe as a whole. And all are essential for understanding similarities / differences with CAFE. I find it remarkable to have a model "validation" vis-à-vis CAFE without explicit consideration of issues such as these – which will determine the similarities and differences in results, especially because there do not seem to be major differences between CAFE and the EVA Model in modelling current or future pollution under the All/all scenarios.

Focus of the paper

As I've said, I think the focus on shipping, and on reducing emissions from shipping, is good. Beyond that, I have difficulty with two aspects of the focus of the paper. The first of these is the focus on modelling the effects of all air pollution from all sources. I read p8 Lines 17-18 as saying this is the main objective of the paper though on re-reading it's not quite as clear as that. (It should not be ambiguous.)

I find the focus on all air pollution from all sources strange – it's been done already, e.g. CAFE, and I can't see what this paper does that is new in that regard, except extend the

domain of application beyond the EU to Europe as a whole – while ignoring completely the methodological issues involved in such an extension.

Within the paper there does seem to be a rationale for Question 3:

- It gives a basis for relative effects, to put the effects of shipping in context – this is done, though it could be developed more incisively.
- Results are used to give a basis for comparison with CAFE – but I think a comparison of results is superficial without at least statement of similarities and differences of methods. [These include that the paper includes emissions from natural sources, whereas CAFE doesn't (because these are not amenable to policy measures) – a difference not noted by the authors here.]

Secondly, I do not understand the focus on results from Denmark specifically – why do this, unless some general conclusions are drawn, in addition to the obvious ones that if emissions are reduced in one part of Europe, i.e. the SECA Region, then there are greater benefits in nearby countries. Half of the text on Conclusions is about Denmark. Why?

Tables

Table 1 is helpful in listing the scenarios examined. Tables 2-4 have a lot of information which has not been used fully. Several aspects deserve comments.

i. In each Table, usually the trend with time is similar for different health outcomes. But there are differences in the sharpness of the slope. I think that lack of proportionality reflects two things – that difference pollutants affect different outcomes, and that the quantification of various outcomes is specific to particular age-groups which differ by outcome. The authors should explain.
ii. As a specific example: Why in Table 3 are acute YOLL in 2020 higher than in 2011, when the converse is true for all other outcomes? (My guess is that it is because acute YOLL reflect an effect of daily ozone only – surely deserves some comment?)
iii. I'm puzzled with results from Table 4. Assuming that we can add across compounds (and nothing is said that we can't), it suggests that <20% of the burden of air pollution across Europe is from primary PM2.5. Is this really so? It may well be true but I didn't expect it to be so small. What is the ratio of primary to secondary PM2.5 in terms of population-weighted average concentrations Europe-wide? Once again, what CRFs were used for SO2?

Figures

These need to be checked and corrected in at least several respects, e.g. (a) Figure 3 is not what the caption says it is; (b) The captions for Figures 1-3 refer to three scenarios when the text refers to two. And (c) Figures 4 and 6 refer to 2020 but the detailed text within the Figures refers to 2000. I assume these errors reflect carelessness – it's surprising that none of the 13 authors noticed them.

To summarise once more: There is something good and worthwhile here but substantial further work is needed. Scaling back to a paper showing effects of measures on concentrations should be considered.

Annex – quick review in Jan-Feb 2013 – now superseded by longer review

This paper highlights that, compared with CAFE in particular, the EVA system has an important advantage – it can and does model the effect of emissions from particular sectors. The present paper uses that attribute of EVA to model the contribution of the shipping sector in the Northern Hemisphere and in the Baltic Sea.

The idea of modelling the effect of emissions from a sector is not new but as far as I know its application to shipping in and around Europe is and the results are important from a policy point of view.

The EVA model has four components – the chemical transport model; population data; 'exposure-response relationships'; and monetary valuation. Of these four, the paper gives detail only on the 1st component. Otherwise, it gives results from the model (in terms of health effects and their monetary value) but without any transparency about the methodology – for this, we are referred to other papers.

The focus on the chemical transport model is I expect of interest to (readers of) this journal; and the 'translation' of effects on PM into effects on health, with corresponding monetary values, is I think of interest also.

I see little that is new (=previously unpublished) methodologically in the paper. The EVA model itself has been described elsewhere. There is a limited 'validation' against CAFE in terms of results for air pollution generally. However, the similarities and differences are not discussed incisively.

I see the paper as being of reasonable interest policy-wise and of limited interest methodologically.

The paper is well written – clearly structured, avoiding jargon, clearly written – though I think to some extent repetitive.