Toward handling uncertainty in prognostic scenarios: Advanced learning from the past

Funded by: Austrian Academy of Sciences (ÖAW), Earth System Sciences (ESS) Research Program
Project duration: 01 September 2015 – 31 August 2016

M. Jonas[1], P. Żebrowski[1], E. Rovenskaya[1], M. Halushchak[1,2], A. Puchkova[3], R. Bun[2]

[1] International Institute for Applied Systems Analysis, Austria, [2] Lviv Polytechnic National University, Ukraine, [3] University of Cambridge, United Kingdom

This poster describes work in progress including contributions by Andriana (2016), Halushchak *et al.* (2016) and Puchkova *et al.* (2016).

Introduction:

The focus of this OAW project is on **uncertainty** and **learning, NOT** on error and 'perfect' predictions into the future (forecasts). We are interested in a metric, an indicator that informs non-experts about the time in the future at which a prognostic scenario ceases to be (for whatever reasons) in accordance with the system's past. The indicator is based on available, observation-based data or estimates and seeks to grasp that part of the future which can be explained by the memory (inertia) contained in the data—or, the explainable outreach of any prognostic scenario in dependence of a given data history. (However, note that we do **NOT** generate prognostic scenarios!) It is also mentioned that (i) we understand 'explainable' more widely than in classical statistics (see below); and (ii) that the indicator is one that is **easily applied** (and not necessarily the most sophisticated from a theoretical perspective).

To better understand both **uncertainty** and **learning**, as perceived in this OAW project, we seek to apprehend the two terms in a wider, more holistic, context. To these ends, we distinguish between **learning in a diagnostic context** and **learning in a controlled prognostic context**, with the OAW project falling under the latter. Both cases start off from available, observation-based data or estimates. The overall purpose of this poster is to **(1)** to provide the wider uncertainty-learning context; while ensuring **(2)** that this context is of global relevance which is why we focus on global greenhouse gas (**GHG**) emissions and concentrations, and mean global temperature data.

In addition, the poster attempts **(3)** to elucidate the scientists' motivations behind both learning in a diagnostic context and learning in a controlled prognostic context. These motivations are **different** which is why these two modes of learning are pursued independently. Finally, the poster envisions (**4)** how learning in a controlled prognostic context is impacted by learning in a diagnostic context. This impact can be considerable.

The poster consists of five parts: **Part I** (below) lists our assumptions and notes of caution. **Part II** (center left) motivates and illustrates learning in a diagnostic context. **Part III** (center right) motivates and illustrates learning in a controlled prognostic context. **Part IV** (center) argues why it is important to combine both learning in a diagnostic and controlled prognostic context. **Part V** (bottom) summarizes our insights from the experiments conducted hitherto.

Part II: Learning in a diagnostic context

Diagnostic uncertainty refers to our ability to estimate current emissions or, **also**, emissions in an agreed future target year. (How a target year is reached is irrelevant; only our real diagnostic capabilities of estimating emissions in the target year matter!) Diagnostic emissions accounting can involve **constant**, **increasing** or **decreasing** uncertainty (compared with a reference year), depending on whether or not our knowledge of emission-generating activities and emission factors becomes more **accurate** and **precise**. This behavior is different from that of prognostic uncertainty (which always increases; see Part III). It is this difference, which suggests that **diagnostic and prognostic uncertainty are independent!**

Motivation and research niche:

Understanding the basic science of GHG sources and sinks requires an understanding of the uncertainty in their estimates **(partial versus full systems view)** (Jonas *et al.* 2015a).

Uncertainty analysis helps designing robust emission compliance regimes (**setting targets**; Fig. 1) and trading schemes (**pricing uncertainty**) and, thus, informing policy (Jonas *et al.* 2015a).

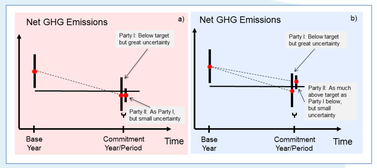

Fig. 1. With compliance and emissions trading in mind—the importance of uncertainty: **Which country (Party) is more credible? a)** Both Parties undershoot the target, but Party I exhibits a greater uncertainty interval than Party II. **b)** Party I exhibits a greater uncertainty interval the mean of which undershoots the target, while Party II exhibits a smaller uncertainty interval, the mean of which, however, does not comply with the target. Source: Jonas & Nilsson (2007: Fig. 10), modified.

Currently, we distinguish between changes in uncertainty due to **(i)** learning (one - sided / bottom - up) and **(ii)** structural changes in emitters, and speculate that the two processes are exponential and can be discriminated (Fig. 2) (Jarnicka & Nahorski 2015; Żerbrowski *et al.* 2015).

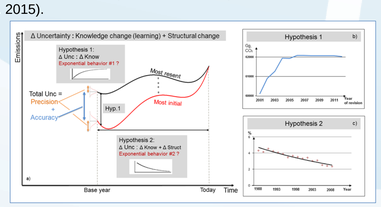

Fig. 2. Illustrating learning in a diagnostic context for uncertain (inaccurate and imprecise) emissions: **a)** Emissions are updated by adding another year of data to the time series, while estimates for earlier years are revised. Uncertainty changes due to **(i)** learning (one - sided / bottom - up) and **(ii)** structural changes in emitters. We hypothesize that the two processes are exponential and can be discriminated (**i** as opposed to **i+ii**). **b)** The figure reflects learning **(i)**. It shows the difference **[most recent – earlier estimates]** for Austria's CO_2 emissions (excluding emissions from land use) for the year 1990. **c)** The figure reflects learning and structural change in emitters **(i+ii)**. It shows the difference **[most recent – most initial estimates]** for Europe's (EU-15) CO_2 emissions (excluding emissions from land use) for 1990–2005. Source: Hamal (2010: Fig. 12), modified.

Part Va: Insights so far from learning in a diagnostic context and outlook

We are able to estimate and distinguish between changes in uncertainty due to **(i)** learning and **(ii)** structural changes in emitters, but so far only for a few countries with good emissions statistics and inventories; and where the first effect outpaces the second. We operate at the limits of skillful resolution. However, we anticipate that, with the advancement of emission inventories, this uncertainty information can be derived for more countries in the future; and that changes in uncertainty due to learning, in particular, will be accounted for in both designing robust emission compliance regimes and trading schemes.

Part I: Assumptions and notes of caution

- Our focus is on observed or (gu-)estimated data which exhibit some sort of memory (inertia)—as it is the case with GHG emissions. (However, we work also with GHG concentrations and global temperature data.) GHG emission data come with diagnostic uncertainty typically comprising **inaccuracy** and **imprecision**.
- Learning in a diagnostic and controlled prognostic context is influenced by our systems view and the data that we consider relevant to understand the system. For example, by assembling GHG emission data globally or nationally only once per year, we implicitly presuppose that the dynamics of emissions at shorter time scales is not worthwhile considering (for whatever reason). However, this can be questioned. For instance, the energy crises in the 1970s led to sharp emission reductions within weeks to months. Thus, by averaging emissions annually, we curtail the system by excluding its most extreme behaviors.
- Our systems view is conservative, meaning that our outreach measure can**NOT** account for 'surprises' which the system has not experienced during its 'one-reality' past.

Part IV: Learning in a diagnostic context to support learning in a controlled prognostic context

Although pursued independently (so far), it is important to realize that both discussed modes of learning should be applied in combination, as learning in a diagnostic context may considerably impact learning in a controlled prognostic context (see Fig. 5 below)

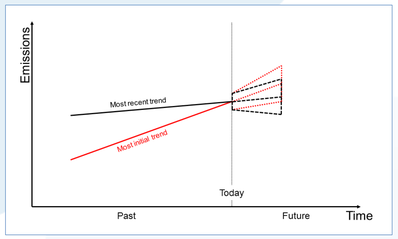

Fig. 5. Illustration of how learning in a diagnostic context may impact learning in a cotrolled prognostic context. **Red:** Most initial trend and uncertainty wedge resulting from learning in a controlled prognostic context based on **unrevised** (most initial) observations (cf. Fig. 2). **Black:** Most recent trend and uncertainty wedge resulting from learning in a controlled prognostic context based on **revised** (most recent) observations (cf. Fig. 2). Learning in a diagnostic context may substantially change our knowledge about the system's past which, in turn, results in grasping both the system's dynamics and the explainable outreach more accurately.

References:

Andriana, D., 2016: Neural Networks to Analyze Uncertainty in Prognostic Model Runs. Report, International Institute for Applied Systems Analysis, Young Scientists Summer Program 2015, Laxenburg, Austria (Ms under scientific review).

Halushchak, M. *et al.*, 2016: Assessment of uncertainty of GHG emission inventories: Learning in a diagnostic context (working title). Manuscript, International Institute for Applied Systems Analysis, Laxenburg, Austria (Ms under preparation for submission to a scientific journal).

Hamal, K., 2010: Reporting GHG Emissions: Change in Uncertainty and its Relevance for the Detection of Emission Changes. Interim Report IR-10-003, International Institute for Applied Systems Analysis, Laxenburg, Austria, pp. 34. http://webarchive.iiasa.ac.at/Publications/Documents/IR-10-003.pdf.

Jarnicka, J. and Z. Nahorski, 2015: A method for estimating time evolution of precision and accuracy of greenhouse gases inventories from revised reports. In: *Proceedings*. 4th International Workshop on Uncertainty in Atmospheric Emissions, 7–9 October, Krakow, Poland [pp. 211, ISBN 83-894-7557-X], 97–102, http://www.ibspan.waw.pl/unws2015/index.php?go=presentations.

Jonas, M. and S. Nilsson, 2007: Prior to economic treatment of emissions and their uncertainties under the Kyoto Protocol: Scientific uncertainties that must be kept in mind. Water Air Soil Pollut.: Focus, 7(4–5), 495–511, doi: 10.1007/s11267-006-9113-7.

Jonas, M., J.P. Ometto, R. Bun and Z. Nahorski, 2015a: Preface. In: *Greenhouse Gas Inventories: Expanding Our Uncertainty Perspective* [J.P. Ometto, R. Bun, M. Jonas and Z. Nahorski (eds.)]. Springer, Dordrecht. Netherlands [pp. 240, ISBN: 978-3-319-15900-3], v–xii.

Jonas, M., P. Żebrowski and E. Rovenskaya, 2015b: A metric for the prognostic outreach of scenarios: Learning from the past to establish a standard in applied systems analysis. 4th International Workshop on Uncertainty in Atmospheric Emissions, 7–9 October, Krakow, Poland [pp. 211, ISBN 83-894-7557-X], 78–89, http://www.ibspan.waw.pl/unws2015/index.php?go=presentations.

Puchkova, A., A. Kryazhimskiy, E. Rovenskaya and M. Jonas, 2016: Cells (working title). Manuscript, International Institute for Applied Systems Analysis, Laxenburg, Austria (Ms under preparation for submission to a scientific journal).

Żebrowski, P., M. Jonas and E. Rovenskaya, 2015: Assessing the improvement of greenhouse gases inventories: Can we capture diagnostic learning? In: *Proceedings*. 4th International Workshop on Uncertainty in Atmospheric Emissions, 7–9 October, Krakow, Poland [pp. 211, ISBN 83-894-7557-X], 90–96, http://www.ibspan.waw.pl/unws2015/index.php?go=presentations.

Part III: Learning in a controlled prognostic context

Prognostic uncertainty comes with any prognostic scenario and **increases** the more the further we look into the future. This behavior is different from that of diagnostic uncertainty (which can increase or decrease; see Part II). It is this difference, which suggests that **diagnostic and prognostic uncertainty are independent!**

Note that—although we seek to approximate this increase in uncertainty with the help of learning in a controlled prognostic context for which we assume **accurate but imprecise** data (Fig. 3)—we do **NOT** aim at grasping prognostic uncertainty!

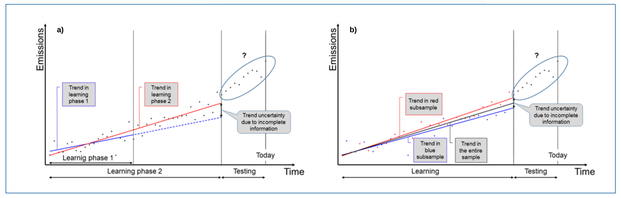

Fig. 3. Learning in a controlled prognostic context builds on **more** than one learning phase to account for our inability to grasp the dynamics of a given (historical) data series perfectly well. Many ways of learning are conceivable. **a)** The figure illustrates learning by means of time steps. **b)** The figure illustrates learning by means of data samples. The preferred mode of learning will influence the way of how we construct the explainable outreach (dealt with in Figure 4). Sources in spirit: Andriana (2016), Puchkova *et al.* (2016).

Motivation and research niche:

An easy-to-apply metric or indicator is long overdue, which informs non-experts about the time in the future at which a prognostic scenario ceases to be (for whatever reasons) in accordance with the system's past. This indicator **(i)** accounts for our inability to grasp the dynamics of a given (historical) data series perfectly well, resulting from the temporal limitedness of the data and, in addition, the uncertainty in the data; and **(ii)** it allows defining the explainable outreach resulting from the memory (inertia) contained in the data. The explainable outreach can be pictured as an uncertainty wedge. We speculate that the opening and extension of the uncertainty wedge can be derived solely from the historical data at hand. For deriving the explainable outreach we foresee (at least) two learning phases and (at least) one testing phase (Fig. 4). It is this process that we refer to when we speak of 'learning in a controlled prognostic context' (Jonas *et al.* 2015b).

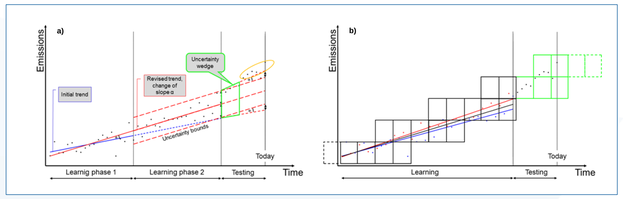

Fig. 4. Expanding Figure 3: Many ways of constructing the explainable outreach—pictured as an uncertainty wedge—are conceivable. **a)** The explainable outreach is constructed with the help of the different trends resulting from learning by means of time steps. When tested, we learn that the outreach must be cropped at the point where data begin to fall outside. **b)** The explainable outreach is constructed with the help of rectangular cells the height of which was chosen so that the rectangles span across the different trends resulting from learning by means of data samples. When tested, we learn that the explainable outreach holds across the entire testing interval. Sources in spirit: Andriana (2016); Puchkova *et al.* (2016).

Part Vb: Insights so far from learning in a controlled prognostic context and outlook

We see increasing evidence that this research issue is legitimate and under-explored. However, we still need to conduct more experiments in order to identify the most promising way forward for balancing three things of a given (historical) data series: the 'right' order of its dynamics (the memory or inertia of the system) and both the 'right' extension and the 'right' opening of the explainable outreach; and, in addition, for making learning in a controlled prognostic context easy to apply throughout model building. We anticipate that accounting for learning will bring models closer to their skillful level of complexity.