Keller, Klaus; Miltich, Louise I.; Robinson, Alexander; Tol, Richard S.J.

Working Paper
How Overconfident are Current Projections of Anthropogenic Carbon Dioxide Emissions?

Nota di Lavoro, Fondazione Eni Enrico Mattei, No. 39.2007

Provided in Cooperation with:
Fondazione Eni Enrico Mattei (FEEM)

Suggested Citation: Keller, Klaus; Miltich, Louise I.; Robinson, Alexander; Tol, Richard S.J. (2007) : How Overconfident are Current Projections of Anthropogenic Carbon Dioxide Emissions?, Nota di Lavoro, Fondazione Eni Enrico Mattei, No. 39.2007

This Version is available at:
http://hdl.handle.net/10419/74122

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
How Overconfident are Current Projections of Anthropogenic Carbon Dioxide Emissions?
Klaus Keller, Louise I. Miltich, Alexander Robinson and Richard S.J. Tol

NOTA DI LAVORO 39.2007

APRIL 2007

CCMP – Climate Change Modelling and Policy

Klaus Keller, Louise I. Miltich and Alexander Robinson,
Department of Geosciences, Penn State University Park
Richard S.J. Tol, Economic and Social Research Institute, Whitaker Square,
Institute of Environmental Studies, Vrije Universiteit, Amsterdam, the Netherlands,
Department of Engineering and Public Policy, Carnegie Mellon University

This paper can be downloaded without charge at:
The Fondazione Eni Enrico Mattei Note di Lavoro Series Index:
http://www.feem.it/Feem/Pub/Publications/WPapers/default.htm
Social Science Research Network Electronic Paper Collection:
http://ssrn.com/abstract=981135

The opinions expressed in this paper do not necessarily reflect the position of
Fondazione Eni Enrico Mattei
Corso Magenta, 63, 20123 Milano (I), web site: www.feem.it, e-mail: working.papers@feem.it
How Overconfident are Current Projections of Anthropogenic Carbon Dioxide Emissions?

Summary
Analyzing the risks of anthropogenic climate change requires sound probabilistic projections of CO2 emissions. Previous projections have broken important new ground, but many rely on out-of-range projections, are limited to the 21st century, or provide only implicit probabilistic information. Here we take a step towards resolving these problems by assimilating globally aggregated observations of population size, economic output, and CO2 emissions over the last three centuries into a simple economic model. We use this model to derive probabilistic projections of business-as-usual CO2 emissions to the year 2150. We demonstrate how the common practice to limit the calibration timescale to decades can result in biased and overconfident projections. The range of several CO2 emission scenarios (e.g., from the Special Report on Emission Scenarios) misses potentially important tails of our projected probability density function. Studies that have interpreted the range of CO2 emission scenarios as an approximation for the full forcing uncertainty may well be biased towards overconfident climate change projections.

Keywords: Carbon Dioxide, Emissions, Scenarios, Data Assimilation, Markov Chain Monte Carlo

JEL Classification: Q540

We thank Arnulf Grübler, Jonathan Koomey, Robert Lempert, Michael Obersteiner, Brian O’Neill, Hugh Pitcher, Steve Smith, Rob Socolow, Dan Ricciuto, and Mort Webster for helpful discussions. Mary Haight provided editorial assistance. This paper was conceived at an EMF Snowmass workshop, thanks to the wonderful environment created by Susan Sweeney and John Weyant. The Hamburg University Innovation Fund, the Princeton Environmental Institute, the Princeton Carbon Mitigation Initiative, the National Science Foundation, and the ESRI Energy Policy Research Centre provided welcome financial support. None but the authors are responsible for opinions and errors in this paper.

Address for correspondence:
Klaus Keller
Department of Geosciences
Penn State University Park
PA 16802
USA
Phone: +1 (814) 865 6718
E-mail: kkeller@geosc.psu.edu
Introduction

Anthropogenic carbon dioxide (CO$_2$) emissions cause climatic changes over many centuries (Archer and Ganopolski 2005). Analyzing the risks associated with this anthropogenic perturbation requires sound probabilistic projections over the relevant timescale (Nordhaus and Yohe 1983; Allen 2003). Projections of anthropogenic climate forcing have broken important new ground (e.g., Reilly et al. 1987; Leggett et al. 1992; Nordhaus and Popp 1997; Schmalensee et al. 1998; Nakicenovic 2000; Webster et al. 2002). Yet, many projections still face potential problems due to (i) a lack of explicit probabilistic information, (ii) a mismatch between the timescale of the climate system and the projection timescales, and (iii) potential overconfidence.

One prominent approach to explore possible future CO$_2$ emissions is to derive scenarios that “cover a wide spectrum of alternative futures to reflect relevant uncertainties” without explicitly assigning probabilities to different scenarios (Nakicenovic et al. 2000). Such scenarios can be useful for exploratory modeling (Bankes 1993). However, the lack of probabilistic information can leave scenario users guessing (Wigley and Raper 2001). Furthermore, reporting a range of “plausible” scenarios (Nakicenovic et al. 2000) involves an implicit (but rather opaque) probabilistic statement. Probabilistic information is central to analyzing the risk trade-offs posed by climate change (Nordhaus and Yohe 1983; Kann and Weyant 2000; Pittock et al. 2001; Schneider 2002; Allen 2003; Webster 2003).

Second, many emission scenarios are limited to a century or less (e.g., Holtz-Eakin and Selden 1995; Schmalensee et al. 1998; Nakicenovic et al. 2000; Webster et al. 2002; Manne and Richels 2004). Analyzing the 21$^{st}$ century is a useful start but implicitly discounts effects beyond 2100. Yet, many climate responses may be triggered in the 21$^{st}$ century but would unfold after 2100 (Stocker and Schmittner 1997; Keller et al. 2004).
Analyzing the risks of long-run climate responses requires long-run emissions projections. The practice of extending the range of the emission scenarios of the Special Report on Emission Scenarios (SRES, Nakicenovic et al. 2000) by assuming a constant atmospheric CO2 concentration beyond the year 2100 (e.g., Gregory et al. 2005) is inconsistent with a business-as-usual (BAU) assumption as it assumes a peculiar trend break at 2100.

Third, projections of CO2 emissions used in scientific assessment may suffer from overconfidence. Overconfidence (i.e., estimating overly tight confidence bounds) is common in laboratory studies and scientific assessments (Alpert and Raiffa 1982; Henrion and Fischhoff 1986). One indication of overconfidence in emission scenarios is the widening of the range of CO2 emissions projected in 2100 as time goes on (Leggett et al. 1992; Nakicenovic et al. 2000). This widening is an indicator of overconfidence because an unbiased error estimate shrinks with information. An additional indication of potential overconfidence in CO2 emission scenarios is that they miss the tails of probabilistic projections (Leggett et al. 1992). Overconfidence in emission scenarios can bias risk analyses as they may underestimate the odds of low-probability high impact events (Gregory et al. 2005; Challenor et al. 2006).

Here we take a step towards data-driven probabilistic projections of population size, economic output, and fossil fuel CO2 emissions that extend beyond the 21st century. We use a Bayesian data assimilation framework to combine prior information with observational constraints. This hindcasting has proven to be a useful (but certainly not sufficient) test to identify structural model weaknesses and to estimate model parameters (Nordhaus and Yohe 1983; Hargreaves and Annan 2002). We then use this model to derive an internally consistent BAU emissions projection to the year 2150. The Bayesian framework provides a rigorous and transparent way to derive probabilistic
projections from a set of observations, structural assumptions, and clearly specified prior information. We use this framework to address three main questions: (i) What is the skill of a simple, mechanistic model in hindcasting the coupled system over several centuries? (ii) How realistic is the assumption that the SRES emission scenarios are a reasonable approximation for the range of relevant outcomes (Wigley and Raper 2001)? (iii) What CO₂ emission projections beyond 2100 are consistent with a simple model and with key observational constraints?


Our analysis captures the dynamics of population, economic output, and CO₂ emissions over the past three centuries reasonably well. The range of the SRES emissions scenarios misses approximately 40% of the tail-areas of our projected CO₂ emissions in 2100. This is consistent with the hypothesis that the SRES scenarios are overconfident. We demonstrate that the current practice of using only decadal-scale observational constraints can result in biased and overconfident predictions by missing important long-term dynamics and underestimating unresolved variability.
Our conclusions are subject to many caveats (discussed in more detail in the section “caveats and research needs” below). First, we choose a very simple model structure. Choosing a model complexity involves a trade-off between approximating the considered process and approximating the uncertainty analysis. Our choice of a simple model enables a relatively careful uncertainty analysis. Second, the analysis of past CO₂ emissions implies that we are deriving only approximately BAU projections. A better characterization of the projections might be a “continuation of past policies” projection (Schmalensee et al. 1998). The projections still seem a reasonable approximation to a BAU case as efforts to reduce CO₂ emissions to date have had only small effects relative to century-scale dynamics. Third, our projections are subject to deep uncertainties associated with the necessary long-range assumptions (Lempert and Schlesinger 2001). It is crucial to stress that our projections are nothing but extrapolations of past observations with a transition to deeply uncertain long-term constraints. These projections can, of course, not predict the timing of unresolved shocks or surprises (Craig et al. 2002).

**Data**

We assimilate globally aggregated observations spanning the years 1700 to 2002. We use estimates of population and economic output made by Maddison (2001). CO₂ emissions due to fossil fuel burning, gas flaring, and cement manufacturing are taken from Marland et al. (2004). We do not consider CO₂ emissions by land use changes. Annual observations of population and economic output span from 1950 to 2002. Prior to 1950, these observations have a much more sparse resolution. CO₂ emissions estimates are annually resolved for the years 1750-2002.
The model

We adopt a simple globally aggregate modeling approach typical of many top-down integrated assessment models (e.g., Nordhaus and Boyer, 2000). The model is composed of three coupled modules describing (i) population size, (ii) economic output, and (iii) anthropogenic CO₂ emissions.

Population

Following previous work, we model population growth by a logistic growth (Verhulst 1838; Cohen 1995), modified by an income sensitive net growth rate. Population \( P \) at time \( t \) is determined by

\[
P_t = P_{t-1} \left[ 1 - \left( \frac{\varphi_1 y_{t-1}}{10^{\varphi_2} + y_{t-1}} \right) \frac{\varphi_3 - P_{t-1}}{\varphi_3} \right] \Delta t,
\]

where \( y \) is per-capita income, \( \varphi_1 \) is the population growth rate, \( \varphi_2 \) is a half-saturation parameter with respect to per-capita consumption, \( \varphi_3 \) is a logistic carrying capacity, and \( \Delta t \) is the time step size (1 year). Table 1 provides definitions of symbols, units, and prior probability density functions (pdfs). This model structure reflects possible interactions between per capita income and population growth. As a technical aside, taking \( \varphi_2 \) as an exponent improved the numerical properties of the assimilation method. Note that Equation (2) implies that population grows and then stabilizes. Other projections assume that the world population peaks and then decreases (Lutz et al. 2001).

Economic output

To model gross world product, we use a Cobb-Douglas production function accompanied by a Solow-Swan model of economic growth. Total world production \( Q \) at time \( t \) is

\[
Q_t = A_t L_t^\lambda K_t^{1-\lambda},
\]
where $A$ is total factor productivity, $L$ is labor force, $K$ is capital stock and $\lambda$ represents the elasticity of production with respect to labor.

Total production is divided between consumption and investment where all output in a given year is assumed to be either consumed or saved

$$Q_t = Y_t = C_t + I_t = (1 - s)Y_t + sY_t,$$  \hspace{1cm} (3)

where $Y$ is total income, $C$ is consumption, $I$ is investment, and $s$ is the savings. From this relationship, per capita income $y$ follows as

$$y = \frac{Y}{P}.$$  \hspace{1cm} (4)

Growth in total factor productivity occurs exogenously according to the Solow-Swan type growth model. The dynamics of long-term technological change are deeply uncertain (Starr and Rudman 1973; Ausubel 1995). Following Nordhaus (1994), we allow for a saturation of total factor productivity

$$A_t = A_{t-1} + \alpha A_{sat} \left[ 1 - \left( \frac{A_{t-1}}{A_{sat}} \right) \right] \Delta t,$$  \hspace{1cm} (5)

where $\alpha$ is a growth rate parameter for total factor productivity and $A_{sat}$ represents a saturation level of total factor productivity. Capital stock dynamics are governed by the balance between investment levels in the previous year and capital loss due to depreciation

$$K_t = [(1 - \delta)K_{t-1} + I_{t-1}] \Delta t,$$  \hspace{1cm} (6)

with $\delta$ representing the capital depreciation rate.

Labor at time $t$ is determined by
\[ L_t = \pi P_t, \]  
where \( \pi \) is the labor participation rate.

Labor is initialized via an initial population parameter \( P_0 \). Capital is initialized using the steady state relationship

\[ K = \left( \frac{sA}{\delta} \right)^{\frac{\lambda}{\delta}} L \]  
and \( A \) is initialized as a parameter \( A_0 \). This results in six economic model parameters to be determined by calibration: \( \lambda, s, A_0, A_{\text{sat}}, \alpha, \delta \). All other variables in the economic model are determined by identities (\( Y, y, C, c \)) or by parameters (\( Q \) by \( \lambda, I \) by \( s, A \) by \( A_{\text{sat}}, \alpha, K \) by \( \delta, L \) by \( \pi \), and \( P \) by the population model).

**Anthropogenic CO₂ emissions**

We approximate the link between economic output and anthropogenic CO₂ emissions through fossil fuel burning and cement production by a time dependent carbon intensity of production. The carbon emissions \( C_t \) at time \( t \) are

\[ C_t = Q_t \phi_t, \]  
where \( \phi_t \) represents the carbon intensity at time \( t \). The carbon intensity is modeled as the weighted average of four broadly defined technologies

\[ \phi_t = \sum_{i=1}^{4} \gamma_{i,t} \rho_i, \]  
where \( \gamma_{i,t} \) is the fraction of the economy invested in technology \( i \) with carbon intensity \( \rho_i \). The time dynamics of \( \gamma_{i,t} \) are approximated as logistic penetration curves

\[ \gamma_{i,t} = \frac{1}{1 + e^{-\kappa t - \tau_i}}, \]
\begin{align}
\gamma_{2,j} &= 1 - \frac{1}{1 + e^{-\kappa T_{2,j}}} - \frac{1}{1 + e^{-\kappa T_{3,j}}} , \\
\gamma_{3,j} &= 1 - \frac{1}{1 + e^{-\kappa T_{2,j}}} - \frac{1}{1 + e^{-\kappa T_{3,j}}} , \text{ and} \\
\gamma_{4,j} &= 1 - \frac{1}{1 + e^{-\kappa T_{2,j}}},
\end{align}

with \( \kappa \) representing the rate at which technologies penetrate and \( \tau_j (j \in [1..4]) \) representing the times when the technology has penetrated half of the market. Such simple penetration patterns can reasonably approximate observed dynamics (Marchetti 1977; Grübler et al. 1999). We set the carbon intensity of the first technology (\( \rho_1 \)) to zero to mimic a situation with negligible fossil fuel emissions (e.g., a subsistence agriculture with predominant use of biomass energy). Humans did cause fairly small CO\(_2\) emissions prior to 1700 (Ruddiman 2003), but these emissions were largely driven by land-use changes, an emission source we do not consider in our analysis. We estimate the carbon intensities of the second and third technologies from the observations with the constraint \( \rho_2 \geq \rho_3 \). This constraint represents the transition from a high carbon intensity technology (i.e., predominant use of high carbon intensity fuels, such as coal) to a lower carbon intensity technology (i.e., shifting from coal to oil and natural gas as well as a shift towards a more service intense economy). We set the carbon intensity of the fourth technology to zero to allow for a long-term transition to a zero carbon emissions economy. We represent the limited resource base of conventional fossil fuels by constraining the total available fossil fuels to 6000 Gt C (Rogner 1997). Obviously, the constraint due to limited fossil fuel resources is subject to considerable uncertainty. One key uncertain factor is the future role of nonconventional fossil fuels derived from methane hydrates (Milkov 2004; Klauda and Sandler 2005). We quantify the effects of this key assumption with a sensitivity study (discussed below).
Model coupling

The population model affects the economic model by providing the labor force. The economic model affects the population model through the sensitivity of the net growth rate on per-capita consumption. Finally, the CO$_2$ emissions model is driven from the economic model through the economic output. The coupled model yields 17 model parameters (Table 1) that are determined by data assimilation.

Estimation methods

We use a Bayesian data assimilation approach to calibrate the model and to estimate the effects of parametric uncertainty on future predictions. The Bayesian approach is an efficient and statistically consistent method to assimilate observations into a nonlinear model with prior information and to derive probabilistic predictions (Brand and Small 1995). The use of prior information is a transparent and parsimonious way to represent information exogenous to the analyzed observations. Consider, for example, the task of projecting population levels over several centuries. Extrapolating the current trends results in projections that are at odds with independent assessments of the overall carrying capacity (Cohen 1995; Lutz et al. 2001). One approach to this problem would be to explicitly represent the calculations leading to the estimates of carrying capacities in the model. We choose a second, arguably more transparent approach by representing these estimates as a Bayesian prior constraint on a model parameter.

Bayes theorem then allows us to combine prior estimates with new information to yield a posterior estimate

\[
p(\tilde{\theta} | \tilde{y}) = \frac{L(\tilde{y} | \tilde{\theta})p(\tilde{\theta})}{\int_{\theta} L(\tilde{y} | \theta)p(\theta)d\theta},
\]  

(15)
where \( p(\hat{\theta} | \bar{y}) \) is the posterior probability of the parameter vector \( \hat{\theta} \) given the observations \( \bar{y} \), \( L(\bar{y} | \hat{\theta}) \) is the likelihood of the observations given the parameters, and \( p(\hat{\theta}) \) is the prior probability of the parameters. We use prior values from previous studies and our own assessments (cf. Table 1). The denominator is a normalization constant to normalize the integral of the pdf to unity.

The Bayesian approach requires evaluating the likelihood of the observations given the model parameters (and, implicitly, the adopted model structure). For identically independently distributed (iid) residuals from a Gaussian distribution, the likelihood function is

\[
L(y | \theta_k) = \sum_{i=1}^{n} \frac{1}{\sqrt{2\pi \sigma_i^2}} \exp \left[ -\frac{1}{2} \left( \frac{f(\theta_k, t_i) - y_i}{\sigma_i} \right)^2 \right],
\]

where the joint likelihood of a set of independent observations is the product of the likelihoods of the individual observations. The iid assumption is violated for the model residuals for the annual population, economic output and CO2 emissions as they show statistically significant autocorrelation. We follow Zellner and Tian (1964) and modify the likelihood function for these observations to account for the autocorrelation effects by whitening the residuals. Following Pizer (1999), we neglect potential correlations between the model residuals for the different observational constraints (e.g., between population size and CO2 emissions).

We approximate \( \sigma_k \) for each observational dataset as the standard deviation of whitened residuals, calculated using the root mean square error (RMSE) best model fit to the data. We derive appropriate \( \sigma \) for each observation type using a genetic algorithm (Storn and Price 1995) to minimize the percent RMSE of the model fit, given the observational data.

For the segment of each time series where annual data are available, the autoregressive
properties of these residuals are assessed and the residuals are whitened using the first order autocorrelation coefficient, in order to satisfy the assumption of normal $iid$ residuals. For non-annual data, we neglect possible effects of autocorrelation. This process results in five $\sigma$ terms. One $\sigma$ term is assigned to the entire emissions time series and two $\sigma$ terms are assigned to each population and gross world product (one for sparse data observed prior to 1950, and the other for annual data observed after 1950).

We use the Metropolis-Hastings algorithm (Metropolis et al. 1953; Hastings 1970) to derive the posterior pdf and the probabilistic hindcasts and projections. This Monte Carlo method allows us to account for the full nonlinear model behavior and to characterize the posterior pdf without making parametric distribution assumptions. This algorithm constructs a Markov chain of parameter vectors that represent direct draws from the posterior pdf. The algorithm samples efficiently from the posterior pdf by probabilistically accepting a parameter vector into the chain based on the gradient of the likelihood function. Using the gradient information allows for preferential sampling in regions with high posterior likelihood. The probabilistic formulation of the acceptance rule minimizes the risk of getting trapped on a local maximum of the posterior pdf.

The Markov chain starts from an initial parameter guess $\hat{\theta}_i$. The likelihood of this parameter vector is computed and compared to the likelihood of a candidate parameter vector $\hat{\theta}_c$ chosen by stepping a specified distance from the previous parameter vector in a direction chosen from a uniform proposal distribution. The candidate vector is accepted into the chain if

$$\frac{L(y|\hat{\theta}_c) p(\hat{\theta}_c)}{L(y|\hat{\theta}_i) p(\hat{\theta}_i)} > U,$$  \hspace{1cm} (17)
where $U$ is a uniform random variable between 0 and 1. This rule allows a higher probability of moving up gradient towards regions of high likelihood mass but also allows the possibility of moving down gradient. Allowing the algorithm to move down gradient as well as up reduces the risk of misconvergence to a local maximum of the posterior pdf. If a candidate point is accepted, the process is repeated with $\tilde{\theta}_i$ replacing $\tilde{\theta}_i$. The chain eventually retains no memory of the starting guess. After a sufficient “burn-in” period, which is removed from the analysis, the chain converges to a stationary sample from the posterior pdf. We use simple heuristic rules to identify the necessary burn in and chain lengths (Raftery and Lewis 1995). This procedure results in a final chain length of roughly $10^6$ samples. Using two independent chains with different initial conditions as well as Bayes Monte Carlo — an alternative assimilation method without the use of gradient information (Dilks et al. 1992) — resulted in consistent estimates of the posterior pdf. We sample from this posterior pdf to derive the hindcasts and projections. The model code, the analyzed data, and the assimilation algorithm are available from the authors by request.

**Results and Discussion**

The model hindcasts the long-term dynamics of population levels, economic output, and carbon dioxide emissions over the last three centuries reasonably well (Figure 1). The CO$_2$ emissions hindcasts capture the declining growth rate of the last few decades. The median CO$_2$ emission projections show a peak of 15 Gigatons of carbon per year (GtC/a) around 2060, with a subsequent decline to 3 GtC/a in 2150. This projected reduction in emission is driven by the constraint on total available fossil fuel resources and the observed decrease in carbon intensity in the last few decades (Nordhaus 1994; Grübler and Nakicenovic 1996).
The population hindcast captures the recent transition to decreasing growth rates (Cohen 1995) reasonably well and projects a median population level of eight billion in 2100. This median estimate compares to a best guess implemented in the prior of 11 billion (Table 1, Lutz et al. 1997). The projected economic output continues to grow, but with a decreasing growth rate. This decreasing growth rate of economic output is driven by decreasing growth rates for population and for total factor productivity, two key inputs to the production function.

The estimation problem is nonconvex (Figure 2) because several parameter pdfs show multiple modes. Parameter estimation methods relying on unimodal assumptions, such as the Ensemble Kalman Filter (Evensen 1994), may face nontrivial misconvergence problems.

The probabilistic CO2 emission projections extend considerably beyond the SRES range (Figures 3 and 4). Webster et al. (2002) assume unimodal pdfs for their uncertainty analysis and report a unimodal CO2 emission projection. We also start with a unimodal assumption for the prior pdf. However, the data assimilation results in a multimodal posterior pdf (Figure 2) and multimodal projections for population size and CO2 emissions (Figure 3). Our projected 95% confidence interval for CO2 emissions in 2100 is 44 GtC/a. This is a wider range than the 31 GtC/a reported by Leggett et al. (1992) or the 34 GtC/a of Nakicenovic et al. (2000) (Figure 3). In contrast, the 95% confidence interval of the projections for population and economic output are narrower than the SRES (Nakicenovic 2000) range (Figure 3). This indicates that the uncertainty in our projections of the carbon intensity is considerably higher than those in the SRES assumptions.
We use our model to analyze potential biases due to neglecting the effects of autocorrelated residuals. This is illustrated by the projected CO₂ emissions in 2010 (Figure 5). Neglecting the effects of autocorrelation results in biased modes and artificially tight confidence limits. The conclusions of previous studies that neglect this effect (e.g., Dowlatabadi and Oravetz, 2006) may therefore need to be revisited.

We quantify the artifacts introduced by relying on only a few decades of observational constraints (Figure 6). Many century-scale projections of industrial CO₂ emissions are calibrated using decadal-scale observations only (e.g., Holtz-Eakin and Selden 1995; Tschang and Dowlatabadi 1995; Heil and Selden 2001; Tol 2005). The overall effects of limiting the calibration range to a small fraction of the projection time-scale reflect several competing effects. First, withholding observations can widen the prediction confidence limits because less information is available to constrain the parameter estimates. On the other hand, withholding observations can artificially narrow the prediction confidence limits when the considered observation subset covers a period of reduced variability (compared to the full sample). The current praxis of using observations over the last few decades arguably falls into this category, in that this period does not include such large scale shocks as World War II or the global flu pandemic of 1918 (Lindmark 2002). In the considered example, the typical praxis of limiting the observational constraints to the last four decades results in artificially tight confidence limits of the projected CO₂ emissions for 2010 and artificially narrow confidence limits for the projected CO₂ emissions for 2100 (Figure 6).

We use a sensitivity study to quantify the effects of two structural assumptions on the projected CO₂ emissions (Figure 7). The model incorporates several expert assessments about the long-term system dynamics that are poorly constrained by the past observations. Key examples of these assumptions include (i) the population carrying
capacity (parameter $\varphi_3$, equation 1), (ii) the saturation level of technological growth (parameter $A_{sats}$, equation 5), and (iii) the available fossil fuel resources. The effects of these long-term future assumptions are minimal for the hindcasts (results not shown), but they become increasingly important the further the projections extend beyond the observational constraints. This effect is demonstrated for the carbon dioxide emission projections. The base case (Figures 1 to 4) adopts a fossil fuel resource constraint and a prior for the saturation level of technological growth approximating (Nordhaus 1994) (Table 1). Relaxing either of these assumptions has only a small effect on the projected carbon dioxide emissions in 2010 (Figure 7). Relaxing the fossil fuel resource constraint and adopting a more diffuse prior for the saturation level of technological growth shifts the mode of the projections by less than 15% and the width of the 95% confidence limits by less than 30%. As expected, the effects of long-term assumptions on the projected CO$_2$ emissions grow with the projection timescale (Figure 7).

**Caveats and Research Needs**

Our conclusions are subject to several caveats that point towards research needs. First, the model is a simple approximation that neglects likely important structural details. Potentially important structural refinements include dynamical representations of (i) changes in the economic sectors (Kongsamut et al. 2001), (ii) population carrying capacity (Cohen 1995), (iii) the savings rate (Deaton and Paxson 2000), (iv) endogenous technological change (Romer 1996), and (v) capital stock effects (Edmonds and Reilly 1983). Furthermore, we aggregate observations to a global scale.

Second, our projections neglect potentially important structural uncertainties. Structural uncertainty is a key concern in models of socioeconomic systems (Koomey 2002; Webster 2003). In our analysis, we consider only a small subset of structural uncertainties. A more complete consideration of structural uncertainties would widen
the projection range (Draper 1995). As a result, our projections still suffer from overconfidence. A more complete consideration of structural uncertainty would, however, also strengthen our conclusion that the SRES scenarios may be overconfident.

**Conclusion**

We assimilate observations of population size, economic output, and CO₂ emissions into a simple model to derive probabilistic hindcasts and projections on a century timescale. The nonconvexity of the posterior pdf poses nontrivial methodological challenges. We demonstrate that the common practice of deriving century-scale economic projections from decadal-scale observations can result in biased and overconfident conclusions. The 95% confidence range of our CO₂ emission projections extends considerably beyond the range covered by past CO₂ emission scenarios (Nakicenovic et al. 2000).

Our results are consistent with the hypothesis that the SRES CO₂ emission scenarios (Nakicenovic et al. 2000) are overconfident and cover only a subset of the potential future outcomes. An alternative hypothesis is that assimilating less aggregated observations into a structurally more complex model would result in a tighter projection range. Testing this alternative hypothesis is possible using our method, but would be beyond the scope of this proof of concept study.

Analyzing the risks of anthropogenic CO₂ emissions requires a careful assessment of the odds of different CO₂ emissions over the relevant multicentury timescale. Deriving these projections is a daunting (and so far open) challenge. We show that a Bayesian data assimilation in conjunction with a simple, transparent, and concise model can provide reasonable probabilistic hindcasts over the last three centuries. A quantitative understanding of the past dynamics is, of course, just a necessary condition for a skillful probabilistic projection of possible futures, but it is a start.
Acknowledgements: We thank Arnulf Grübler, Jonathan Koomey, Robert Lempert, Michael Obersteiner, Brian O’Neill, Hugh Pitcher, Steve Smith, Rob Socolow, Dan Ricciuto, and Mort Webster for helpful discussions. Mary Haight provided editorial assistance. This paper was conceived at an EMF Snowmass workshop, thanks to the wonderful environment created by Susan Sweeney and John Weyant. The Hamburg University Innovation Fund, the Princeton Environmental Institute, the Princeton Carbon Mitigation Initiative, the National Science Foundation, and the ESRI Energy Policy Research Centre provided welcome financial support. None but the authors are responsible for opinions and errors in this paper.

References


response to increasing atmospheric CO₂ concentration." Geophysical Research Letters 32(12).


Pittock, A. B., R. N. Jones and C. D. Mitchell (2001). "Probabilities will help us plan for climate change - Without estimates, engineers and planners will have to delay decisions or take a gamble." Nature 413(6853): 249-249.


Table Header and Footnotes

Table 1: Prior ranges of the model parameters. Lower and upper values are absolute bounds for uniform probability density functions and 95% confidence limits otherwise.

Footnotes:

a The lower bound is the peak population in the 0.025% scenario; the upper bound the 2100 population in the 0.975% scenario.

b The lower bound is the minimum of the four alternative estimates of Maddison (2001, Table B-1) minus the standard deviation; the upper bound the maximum plus the standard deviation.

c The average gross savings rate between 1970 and 2002 according to WRI (2005; http://earthtrends.wri.org) is 22% with a standard deviation of 1%. The range given here is the 95% confidence interval, with the standard deviation arbitrarily doubled.

d We use the ratio of the 2005 level to the long-term saturation level (Nordhaus 1994) with a uniform probability density function of +/- 50%. For the sensitivity study with an uninformative prior, we increase this range by a factor of $10^{90}$.

e $A_0 = Y_0 \left( \frac{\delta}{\pi P_0} \right)^{\frac{1}{2\lambda}} \left( \frac{\pi P_0}{\lambda} \right)^{\frac{1}{\lambda+1}}$. This gives the best guess for $A_0$. 

**Figure Legends**

**Figure 1:** Observations, hindcasts and projections for population (panel a), economic output (panel b), and carbon dioxide (CO$_2$) emissions (panel c). The CO$_2$ hindcasts and projections between 1950 and 2010 are shown in greater detail in panel d. The circles represent the observational constraints. The model hindcast and projections are plotted as the median (solid line) as well as the 5 and 95% confidence limits (dashed lines) and the 1 and 99% confidence limits (dotted lines).

**Figure 2:** Posterior probability density functions (pdf) of the model parameters. Shown are the marginal pdfs of the joint pdf. Note the multimodal properties (e.g., $s$ or $\pi$).

**Figure 3:** Model projections for population (left column), economic output (middle column), and carbon dioxide (CO$_2$) emissions (right column) for the years 2050 (upper row), 2100 (middle row), and 2150 (bottom row). The horizontal lines denote the range spanned by the 40 SRES scenarios (Nakicenovic et al. 2000).

**Figure 4:** Comparison of the range of the SRES emission scenarios with the 1$^{\text{st}}$ and 99$^{\text{th}}$ percentiles of the probabilistic projection. The range of the SRES emission scenarios (Nakicenovic et al. 2000) are indicated by the horizontal lines.

**Figure 5:** Effects of neglecting autocorrelation in the model residuals on the projections. Shown are the projected carbon dioxide emissions in 2010 for the case of neglecting (white bars) and accounting for the effects of autocorrelated residuals (black bars).

**Figure 6:** Effects of introducing the century scale observational constraint on the projected carbon dioxide (CO$_2$) emissions. Shown are projections for the years 2010
(panel a) and 2100 (panel b) with observation constraints spanning back to the year 1965 (dashed lines) and to the year 1700 (solid lines).

**Figure 7:** Effects of long-term assumptions on the short (2010, panel a), medium (2100, panel b) and long-term (2200, panel c) projections for carbon dioxide emissions. Shown are the cumulative density functions (cdf) for the base case (with a fossil fuel resource constraint and an informative prior on the saturation of technological growth) and the cases where these two assumptions are relaxed. See text and table 1 for details.
### Table 1

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Units</th>
<th>Distribution</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varphi_1 )</td>
<td>population growth rate</td>
<td>1/year</td>
<td>uniform</td>
<td>0.0001</td>
<td>0.15</td>
<td>this study</td>
</tr>
<tr>
<td>( \varphi_2 )</td>
<td>half-saturation constant of growth scaling</td>
<td>$ / (year capita)$</td>
<td>uniform</td>
<td>0</td>
<td>5</td>
<td>this study</td>
</tr>
<tr>
<td>( \varphi_3 )</td>
<td>population carrying capacity</td>
<td>billions</td>
<td>normal</td>
<td>6.9</td>
<td>14.4</td>
<td>(Lutz et al. 1997)(^a)</td>
</tr>
<tr>
<td>( \varphi_4 )</td>
<td>population in 1700</td>
<td>billions</td>
<td>normal</td>
<td>0.3</td>
<td>0.9</td>
<td>(Maddison 2001)(^b)</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>elasticity of production with respect to labor</td>
<td>dimensionless</td>
<td>uniform</td>
<td>0.6</td>
<td>0.8</td>
<td>(Romer 1996)</td>
</tr>
<tr>
<td>( s )</td>
<td>savings rate</td>
<td>dimensionless</td>
<td>normal</td>
<td>0.18</td>
<td>0.26</td>
<td>this study(^c)</td>
</tr>
<tr>
<td>( \delta (&lt;s) )</td>
<td>capital depreciation rate</td>
<td>1/year</td>
<td>uniform</td>
<td>0.01</td>
<td>0.14</td>
<td>(Nordhaus 1994)(^d)</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>rate of technological progress</td>
<td>1/year</td>
<td>uniform</td>
<td>0.0007</td>
<td>0.0212</td>
<td>this study</td>
</tr>
<tr>
<td>( A_{sat} )</td>
<td>saturation level of total factor productivity</td>
<td>dimensionless</td>
<td>uniform</td>
<td>5.3</td>
<td>16.11</td>
<td>(Nordhaus 1994)(^d)</td>
</tr>
<tr>
<td>( \pi )</td>
<td>labor participation rate</td>
<td>dimensionless</td>
<td>normal</td>
<td>0.51</td>
<td>0.67</td>
<td>this study</td>
</tr>
<tr>
<td>( A_0 )</td>
<td>initial total factor productivity</td>
<td>dimensionless</td>
<td>uniform</td>
<td>0</td>
<td>3</td>
<td>this study(^e)</td>
</tr>
<tr>
<td>( \sigma_2 )</td>
<td>carbon intensity technology 2</td>
<td>kg carbon / 1990 $</td>
<td>uniform</td>
<td>0</td>
<td>0.5</td>
<td>this study</td>
</tr>
<tr>
<td>( \sigma_3 )</td>
<td>carbon intensity technology 3</td>
<td>kg carbon / 1990 $</td>
<td>uniform</td>
<td>0</td>
<td>0.5</td>
<td>this study</td>
</tr>
<tr>
<td>( \tau_2 )</td>
<td>half-saturation time technology #2</td>
<td>year</td>
<td>uniform</td>
<td>1700</td>
<td>2100</td>
<td>this study</td>
</tr>
<tr>
<td>( \tau_3 )</td>
<td>half-saturation time technology #3</td>
<td>year</td>
<td>uniform</td>
<td>1700</td>
<td>2100</td>
<td>this study</td>
</tr>
<tr>
<td>( \tau_4 )</td>
<td>half-saturation time technology #4</td>
<td>year</td>
<td>uniform</td>
<td>2010</td>
<td>2500</td>
<td>this study</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>time constant of technological penetration</td>
<td>1/year</td>
<td>uniform</td>
<td>0.005</td>
<td>0.2</td>
<td>(Grübler 1991)</td>
</tr>
</tbody>
</table>
Figure 1

Population

Gross World Product

Carbon Dioxide Emissions

Carbon Dioxide Emissions (Detail)
Figure 2
Figure 3
Emissions (Gt C/a)

this study
99th percentile

observations

1st percentile

Figure 4
Figure 5
Figure 6
Figure 7
### NOTE DI LAVORO PUBLISHED IN 2007

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PRCG</td>
<td>3.2007</td>
<td>Cesare Dosi and Michele Moretto: <em>Concession Bidding Rules and Investment Time Flexibility</em></td>
</tr>
<tr>
<td>IEM</td>
<td>4.2007</td>
<td>Chiara Longo, Matteo Manera, Anil Markandya and Elisa Scarpa: <em>Evaluating the Empirical Performance of Alternative Econometric Models for Oil Price Forecasting</em></td>
</tr>
<tr>
<td>PRCG</td>
<td>5.2007</td>
<td>Bernardo Bortolotti, William Megginson and Scott B. Smart: <em>The Rise of Accelerated Seasoned Equity Underwritings</em></td>
</tr>
<tr>
<td>CCMP</td>
<td>6.2007</td>
<td>Valentina Bosetti and Massimo Tavoni: <em>Uncertain R&amp;D, Backstop Technology and GHGs Stabilization</em></td>
</tr>
<tr>
<td>CCMP</td>
<td>10.2007</td>
<td>Valentina Bosetti, Emanuele Massetti, and Massimo Tavoni: <em>The WITCH Model. Structure, Baseline, Solutions</em></td>
</tr>
<tr>
<td>SIEV</td>
<td>11.2007</td>
<td>Margherita Turvani, Aline Chiabai, Anna Alberini and Stefania Tomin: <em>Public Policies for Contaminated Site Cleanup: The Opinions of the Italian Public</em></td>
</tr>
<tr>
<td>CCMP</td>
<td>12.2007</td>
<td>M. Berrittella, A. Cerra, M. Enea and P. Zito: <em>An Analytic Hierarchy Process for The Evaluation of Transport Policies to Reduce Climate Change Impacts</em></td>
</tr>
<tr>
<td>NRM</td>
<td>13.2007</td>
<td>Francesco Bosello, Barbara Buchner, Jacopo Crimi, Carlo Giupponi and Andrea Povellato: <em>The Kyoto Protocol and the Effect of Existing and Planned Measures in the Agricultural and Forestry Sector in the EU25</em></td>
</tr>
<tr>
<td>NRM</td>
<td>14.2007</td>
<td>Francesco Bosello, Carlo Giupponi and Andrea Povellato: <em>A Review of Recent Studies on Cost Effectiveness of GHG Mitigation Measures in the European Agro-Forestry Sector</em></td>
</tr>
<tr>
<td>CCMP</td>
<td>15.2007</td>
<td>Massimo Tavoni, Brent Sohngen, and Valentina Bosetti: <em>Forestry and the Carbon Market Response to Stabilize Climate</em></td>
</tr>
<tr>
<td>ETA</td>
<td>16.2007</td>
<td>Erik Ansink and Arjan Ruijs: <em>Climate Change and the Stability of Water Allocation Agreements</em></td>
</tr>
<tr>
<td>ETA</td>
<td>17.2007</td>
<td>François Gusdorf and Stéphane Hallegatte: <em>Compact or Spread-Out Cities: Urban Planning, Taxation, and the Vulnerability to Transportation Shocks</em></td>
</tr>
<tr>
<td>NRM</td>
<td>18.2007</td>
<td>Giovanni Bella: <em>A Bug’s Life: Competition Among Species Towards the Environment</em></td>
</tr>
<tr>
<td>IEM</td>
<td>19.2007</td>
<td>Valeria Termini and Laura Cavallo: <em>Spot, Bilateral and Futures Trading in Electricity Markets. Implications for Stability</em></td>
</tr>
<tr>
<td>ETA</td>
<td>20.2007</td>
<td>Stéphane Hallegatte and Michael Ghil: <em>Endogenous Business Cycles and the Economic Response to Exogenous Shocks</em></td>
</tr>
<tr>
<td>CTN</td>
<td>21.2007</td>
<td>Thierry Bréchet, François Gerard and Henry Tulkens: <em>Climate Coalitions: A Theoretical and Computational Appraisal</em></td>
</tr>
<tr>
<td>CCMP</td>
<td>22.2007</td>
<td>Claudia Ketterer, Angela Köpp, Stefan P. Schleicher and Gregor Thenius: <em>Stringency and Distribution in the EU Emissions Trading Scheme –The 2005 Evidence</em></td>
</tr>
<tr>
<td>NRM</td>
<td>23.2007</td>
<td>Hongyu Ding, Arjan Ruijs and Ekko C. van Ierland: <em>Designing a Decision Support System for Marine Reserves Management: An Economic Analysis for the Dutch North Sea</em></td>
</tr>
<tr>
<td>CCMP</td>
<td>24.2007</td>
<td>Massimiliano Mazzanti, Anna Montini and Roberto Zoboli: <em>Economic Dynamics, Emission Trends and the EKC Hypothesis New Evidence Using NAMEA and Provincial Panel Data for Italy</em></td>
</tr>
<tr>
<td>ETA</td>
<td>25.2007</td>
<td>Joan Canton: <em>Redealing the Cards: How the Presence of an Eco-Industry Modifies the Political Economy of Environmental Policies</em></td>
</tr>
<tr>
<td>ETA</td>
<td>26.2007</td>
<td>Joan Canton: <em>Environmental Taxation and International Eco-Industries</em></td>
</tr>
<tr>
<td>CCMP</td>
<td>27.2007</td>
<td>Oscar Cacho and Leslie Lipper (lxxxi): <em>Abatement and Transaction Costs of Carbon-Sink Projects Involving Smallholders</em></td>
</tr>
<tr>
<td>CCMP</td>
<td>28.2007</td>
<td>A. Caparrós, E. Cerdá, P. Ovando and P. Campos (lxxxi): <em>Carbon Sequestration with Reforestations and Biodiversity-Scenic Values</em></td>
</tr>
<tr>
<td>CCMP</td>
<td>29.2007</td>
<td>Georg E. Kindermann, Michael Obersteiner, Ewald Rametsteiner and Ian McCallum (lxxxi): <em>Predicting the Deforestation-Trend Under Different Carbon-Prices</em></td>
</tr>
</tbody>
</table>
(lxxxi) This paper was presented at the EAERE-FEEM-VIU Summer School on "Computable General Equilibrium Modeling in Environmental and Resource Economics", held in Venice from June 25th to July 1st, 2006 and supported by the Marie Curie Series of Conferences "European Summer School in Resource and Environmental Economics".

(lxxxii) This paper was presented at the Workshop on “Climate Mitigation Measures in the Agro-Forestry Sector and Biodiversity Futures”, Trieste, 16-17 October 2006 and jointly organised by The Ecological and Environmental Economics - EEE Programme, The Abdus Salam International Centre for Theoretical Physics - ICTP, UNESCO Man and the Biosphere Programme - MAB, and The International Institute for Applied Systems Analysis - IIASA.