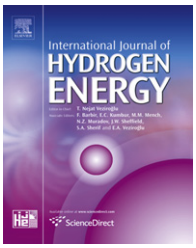
INTERNATIONAL JOURNAL OF HYDROGEN ENERGY 36 (2011) 9382–9395

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/he



Using non-parametric statistics to identify the best pathway for supplying hydrogen as a road transport fuel

Justin D.K. Bishop [a,*], *Colin J. Axon* [b], *David Banister* [c], *David Bonilla* [c], *Martino Tran* [a], *Malcolm D. McCulloch* [d]

[a] *Oxford Martin School Institute for Carbon and Energy Reduction in Transport, c/o Department of Engineering Science, 17 Parks Road, Oxford OX1 3PJ, United Kingdom*[1]

[b] *University of Oxford, Department of Engineering Science, 17 Parks Road, Oxford OX1 3PJ, United Kingdom*

[c] *Transport Studies Unit, School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, United Kingdom*

[d] *Energy and Power Group, University of Oxford Department of Engineering Science, 17 Parks Road, Oxford OX1 3PJ, United Kingdom*

ARTICLE INFO

Article history:
Received 7 December 2010
Received in revised form
22 April 2011
Accepted 22 April 2011
Available online 31 May 2011

Keywords:
Well-to-tank
Non-parametric statistics
Adaptive kernel density estimator
Hydrogen

ABSTRACT

The wealth of estimates quantifying the well-to-tank (WTT) impacts of hydrogen vary significantly. This variation is due to both methodology and the chosen production pathway (gasification, electrolysis, or steam reforming). The statistical distribution of the WTT estimates is non-Gaussian and this work demonstrates the adaptive kernel density estimator as a robust, non-parametric statistical method for determining the underlying probability density function. The approach is flexible, expandable and can be used to investigate the development of hydrogen supply pathways through time. The adaptive kernel density estimator outperforms the first generation (oversmoothed and least squares cross-validation), second generation Sheather and Jones Plug-In and the median. In particular, it represents the multimodal features of the data set better than both the first and second generation methods with less variability than the least squares cross-validation approach. The peak of the distribution represents the most likely pathway (best estimate) for supplying hydrogen. This work suggests that the overall best estimate for supplying hydrogen is by natural gas from Europe via central reforming, subject to a trade-off between the energy impacts and the resultant emissions. Through time, the overall hydrogen production process has become more energy efficient at the expense of greater emissions per MJ delivered to the tank. The best-in-class pathway is that with the lowest greenhouse gas emissions per MJ hydrogen delivered and represents the state-of-the-art. Overall, the best-in-class pathway combination for providing hydrogen is by electricity from renewables via electrolysis.

* *Corresponding author*. Tel.: +44 1865 273 032.
E-mail address: justin.bishop@eng.ox.ac.uk (J.D.K. Bishop).

[1] URL: http://www.eng.ox.ac.uk.

0360-3199/$ – see front matter Copyright © 2011, Hydrogen Energy Publications, LLC. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijhydene.2011.04.173

1. Introduction

A robust, non-parametric statistical method is proposed to answer the key question of what are the best well-to-tank (WTT) pathways to supply hydrogen, based on the numerous estimates available in the literature. The WTT pathway comprises: transformation of the primary resource; transportation of fuel intermediates; and distribution of final fuel [1]. Each stage has associated energy and emissions penalties. To reduce well-to-wheel (WTW) energy use and greenhouse gas (GHG) emissions from road transport, the use of non-conventional fuels and novel vehicle topologies have been proposed and developed. One such fuel is hydrogen. It is difficult to draw defensible comparisons across WTT estimates, within and between studies, on account of the differences in assumptions and system boundaries used. Consequently, estimates vary considerably. Moreover, global trade means that a final fuel which originates from natural resources in one country may be processed in another and consumed in a third. Therefore, all WTT supply pathways and associated impacts are relevant to a global fuel trade.

Non-parametric statistical approaches are used to estimate the probability density of a set of data with an unknown underlying distribution. A list of distribution tests could be performed on the set of WTT estimates to determine its probability distribution. Rather than assume a discoverable, closed-form distribution and attempt to fit the data to it, a more defensible approach is to let the distribution emerge from the data. Therefore, non-parametric density estimation based on Gaussian sample point kernel density estimators is proposed as a robust method to determine the best (highest probability) representative. The emissions associated with the WTT supply of hydrogen dominate those which arise due to its in-vehicle use [2]. Consequently, for hydrogen to be a competitive alternative to the conventional transport fuels, its low emitting pathways must be identified. In the limit, the 'best-in-class' pathway is that with the lowest GHG emissions per MJ delivered and thus represents the state-of-the-art.

The objective of using non-parametric methods should be to estimate a distribution that closely approximates the true, underlying distribution [3]. In practice, the methods attempt to find the optimum bandwidth(s), h, of the data set to create the probability density estimate which balances bias and variability [4]. Specifically, as h increases, the number of bins, m, shrinks resulting in large bias, small variance and an oversmoothed distribution. Conversely, small h leads to a large m, small bias, large variance: the result is an undersmoothed distribution [5,6].

Non-parametric methods can be classified as first generation, second generation and adaptive. First generation methods to determine optimum h assume an underlying unimodal Gaussian distribution [5–7]. The bandwidth, h, is generally chosen to minimize the mean integrated square error (MISE). MISE is sub-optimal for multimodal densities [8] on account of the global h being inappropriate for distributions with many features [6]. Moreover, the error minimization techniques of MISE and asymptotic MISE require prior knowledge of both the distribution function which is being sought and its derivative. Cross-validation (CV) methods are independent of the function and derivative, where the optimal bin count is the minimum CV value over the range of bin count possibilities, from 1 to m [8]. While CV methods are the most studied [5] and used [9,10], they can suffer from too much sample variability [11,12] on account of large bias [7].

More sophisticated second generation bandwidth estimation methods, such as the plug-in and bootstrap methods, are superior alternatives which offer a more sensible trade-off between bias and variability, complete with rates of convergence of $n^{-5/14}$, asymptotic to $n^{-1/2}$, where n is the number of estimates. Thus, they surpass the first generation rules-of-thumb and CV approaches which have rates of $n^{-1/10}$ [6,12]. The Sheather and Jones plug-in method (SJPI) [13] achieves the asymptotic best convergence rate and is computationally more efficient than the bootstrap alternatives [14,15]. However, plug-in methods use pilot estimates for h which are based on a number of prior assumptions about the true h and density. Consequently, if the initial assumptions fail, the estimate is expected to be poor [3,15]. Fixed (first and second generation) bandwidth approaches fail to account for the changes in data density [16]. Given that the detail of the underlying distribution emerges as a result of h (and associated data points within a bin), variable bandwidth or adaptive kernel estimators have been proposed.

Adaptive density estimation uses local smoothing to yield an improved, global estimate by varying the bandwidth chosen as the density of data changes. The bandwidth is reduced in areas with higher density data in order to capture all the features; and the bandwidth is increased at the tails of the distribution, where the data is more sparse [17]. Kernel density estimators can be biased near to the boundary or end point rather than in the body of the data set proper. This is known as the boundary bias or edge effect [7,9] which fixed, symmetric kernels with boundary supports suffer from. Thus, kernels which do not assign weight outside of the extremes of the data set are proposed. In particular, the Gamma kernel is free of boundary bias, always non-negative, obtains optimal convergence rate of the MISE and accommodates sparse areas in the distribution well [9]. Another approach to reducing boundary effects is by data sharpening, where data is moved from areas where it is sparse to those where the density is higher [7]. In the context of the best WTT estimate, the MJ/MJ value must be positive while the g GHG/MJ may be positive or negative. Therefore, acknowledgement of the boundary supports is necessary to ensure a sensible (particularly MJ/MJ) estimate.

Hereafter, the adaptive kernel density estimator method is introduced in Section 2, including the approach to weighting and the metric for choosing the real-world pathway. It is recognized that every primary resource used for producing hydrogen may not have the required potential to displace conventional fuels on a large-scale. Therefore, a weighting factor is added to the WTT estimates based on their resource potential to ensure realistic results. The best estimate result of the statistical method represents the peak of the probability distribution which may not correspond to a real-world pathway. Therefore, the real-world pathway chosen has the smallest, absolute Euclidean distance to the distribution

peak. There are four identified advantages of this method which are discussed. The method is flexible, expandable, can incorporate timing and be used for WTT estimates on other transport fuels. In particular, the inclusion of timing via a quasi-static analysis illustrates how the hydrogen supply pathways have changed through time. These four advantages and n analysis is included to demonstrate the sensitivity of the results to the weighting factor and to compare the performance of the adaptive kernel density estimator to the first and second generation oversmoothed, least squares CV, SJPI methods and the simple median are presented in Section 3.

2. Method

The adaptive kernel density estimator is adopted as a robust approach to estimating distributions which may exhibit many features. The method proceeds by:

1. testing for statistical independence between the estimates using the χ^2 test with a threshold of $\rho > 0.05$ (Table 1), as a prerequisite for probability density estimation;
2. formulating a bandwidth estimate using the adaptive kernel density estimator;
3. applying a bivariate Gaussian to each $n \times 2$ WTT estimate of fuel, j, delivered across all pathways, i using the various h estimates:

The Gaussian kernel is employed instead of the Gamma as it is more commonly used in the literature and the kernel has been shown to have less of an impact on the final non-parametric distribution than the choice of h [5–7]. The MJ/MJ component of the non-parametric distribution is truncated at zero to ensure a positive WTT energy impact;

4. weighting each WTT estimate by the resource potential:

A weighting factor is included to account for input resources and primary pathways which may yield low WTT impact hydrogen, but are not sufficiently abundant to form the basis for large-scale hydrogen supply. Therefore, the WTT estimate was scaled by a weighting factor, which is based on the production capacity or yield of its primary resource (Table 2) and serves as a proxy for the long term supply potential. The weighting factor is the quotient of each primary resource capacity to the maximum capacity for the particular pathway (Eq. (1)). Therefore, the primary resources with maximum potential are assigned one and those with smaller potentials are given a positive weight less than one; and

5. repeating the above steps for each pathway, i.

The bandwidth, h, was determined individually for each dimension in the $n \times 2$ data vector of MJ/MJ and g GHG/MJ estimates and later combined to yield the two dimensional density estimate. The algorithm was implemented in Matlab 7.6.0 R 2008a.

1. The adaptive kernel estimator followed the procedure given in [16] to:
   (a) Split the data into a large number of evenly spaced bins, $m\_bins = 100$, noting the frequency and bin position;
   (b) Determine the number of bins with nonzero frequencies. That is, disregard empty bins to yield a new upper bin count, $m\_unique < m\_bins$;
   (c) Perform leave-one-out least squares CV by iterating through the number of bins from 1 to $m\_unique = m$. h_{binned} is chosen corresponding to the minimum value of the CV. The least squares CV is calculated in three parts, given as f_1, f_2 and f_3 in Eq. (1):

The components of Eq. (2) are given in Eqs. (3)–(5) by:

(d) Choose the optimum bin as that which corresponds to the minimum of the least squares CV and the corresponding bandwidths for bins 1 to optimum_bin;

Table 1 – Test for independence for entire hydrogen WTT data set and by pathway. The number of estimates is given by n and $\rho > 0.05$ is the χ^2 test used to determine independence using the Matlab 7.6.0 R 2008a function, *crosstab*. The test for a Gaussian distribution is included for illustrative purposes using *kstest*.

Primary pathway	n	ρ	Independent	Gaussian
All	124	0.50	Yes	No
Gasification	17	0.24	Yes	No
Electrolysis	49	0.10	Yes	No
Steam reforming	59	0.61	Yes	No

Table 2 – Weighting factors based on annual resource potential (MJ/yr) for hydrogen overall and by its gasification, electrolysis and steam reforming pathways. The overall weight for each pathway and input resource is normalized to the maximum resource potential across all pathways and input resources. The weighting factor per pathway is normalized to the maximum resource potential for that particular pathway, across all input resources.

Primary pathway	Input resource	Resource potential (10^{14} MJ/yr)	Weight	
			Overall	By pathway
Gasification	Wood [35]	0.2522	1.0000	1.0000
	Wood waste [36]	1.5000	0.3000	0.3000
Electrolysis	Country mix [37]	0.7321	0.1464	0.3849
	Natural gas [38]	1.1603	0.2321	0.6100
	Nuclear [37]	0.0987	0.0197	0.0519
	Non-combustible renewables and waste [37]	0.0047	0.0009	0.0025
	Coal [38]	1.9022	0.3804	1.0000
	Wood [35]	0.2522	0.0504	0.1326
Steam reforming	Natural gas [38]	1.1603	0.2321	1.0000

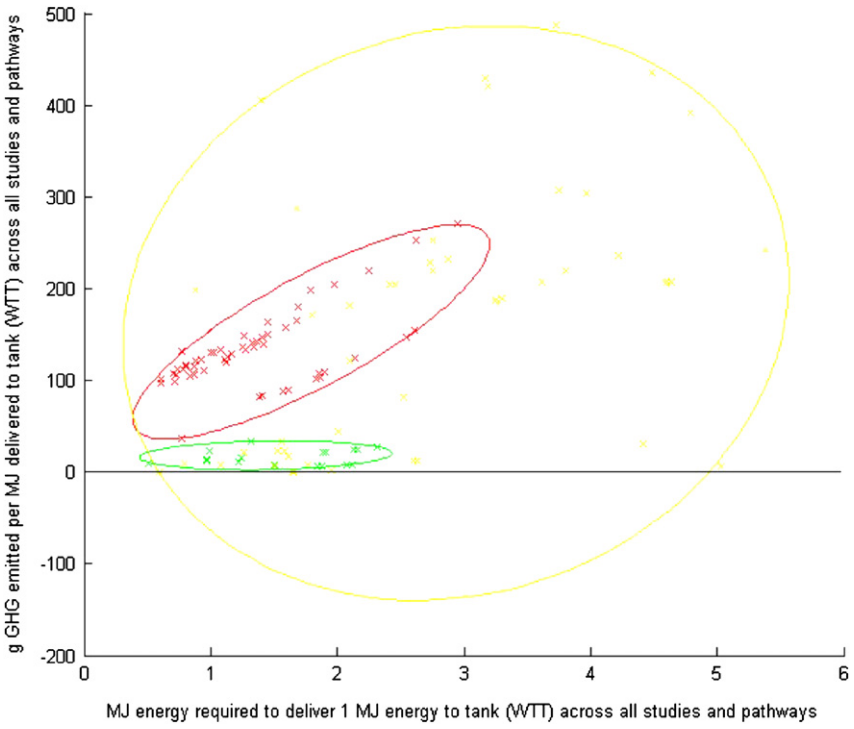


Fig. 1 – Scatter of MJ and emitted g GHG when delivering 1 MJ hydrogen, across the steam reforming (red), electrolysis (yellow) and gasification (green) pathways ([2,22–32]). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(e) Apply this vector of h_{adap} of length = 1:optimum_bin to the data by matching smallest h_{adap} (i) to bin with highest frequency, in decreasing order and apply h_{adap} (i) to all estimates in the same bin, i

(f) Calculate density estimate in Eq. (6) :

2. Use the bandwidth in a bivariate Gaussian[2] function to determine the probability distribution;
3. Identify the values in each dimension of the $n \times 2$ data vector of g GHG/MJ and MJ/MJ estimates corresponding to the indices of the highest probability as the best estimate values;
4. Select the real pathway closest to the peak in the probability distribution by minimum Euclidean distance; and
5. Choose the best-in-class g GHG/MJ pathway as the data vector entry with the smallest g GHG/MJ impact and corresponding MJ/MJ value. The best-in-class estimate is independent of the weighting as it represents the state-of-the-art without regard for large-scale potential.

The primary results of this work are based on a static (time-independent) picture of the state of hydrogen WTT pathways. The method may be repeated at discrete time steps in a quasi-static manner to illustrate how the best estimate moves with time.

The results presented do not consider the cost of supplying hydrogen, however the method and the insights into learning may be extended in that direction. Costs are excluded as they are relatively few studies which provide the WTT MJ/MJ and g GHG/MJ metrics with associated costs of supply (€/MJ). The non-parametric method proposed is most robustly demonstrated using the maximum number of estimates of MJ/MJ and g GHG/MJ. Including best estimate and best-in-class costs would be achieved by a repeat of the non-parametric method using the reduced data set of two studies to maintain consistency across the three metrics. Thus, it is believed that more would be lost in using a reduced data set to justify the non-parametric method than would be gained by including costs.

Using the above adaptive kernel density estimator method, the static analysis proceeds by presenting the best estimate and best-in-class pathway for supplying hydrogen overall and by – gasification, electrolysis and steam reforming – pathway, as presented in the literature. The gasification pathway begins with farmed wood or wood waste which is chipped for transport by road or sea. The biomass may be directed to a centralized gasification plant, from which the hydrogen product is compressed or liquefied and distributed by pipeline or road to the filling station. Alternatively, the biomass may be transported directly to the filling station for decentralized, onsite gasification.

The electrolysis pathway has the most diverse primary resource option set, ranging from natural gas, coal and other fossil-fuels, both combustible and non-combustible renewables and waste to nuclear power plants. In all cases, the electricity is an intermediate product. It may be transmitted to a centralized electrolysis facility, from which the hydrogen is

[2] The m-file available at http://www.mathworks.com/matlabcentral/fileexchange/22999 was modified for this purpose.

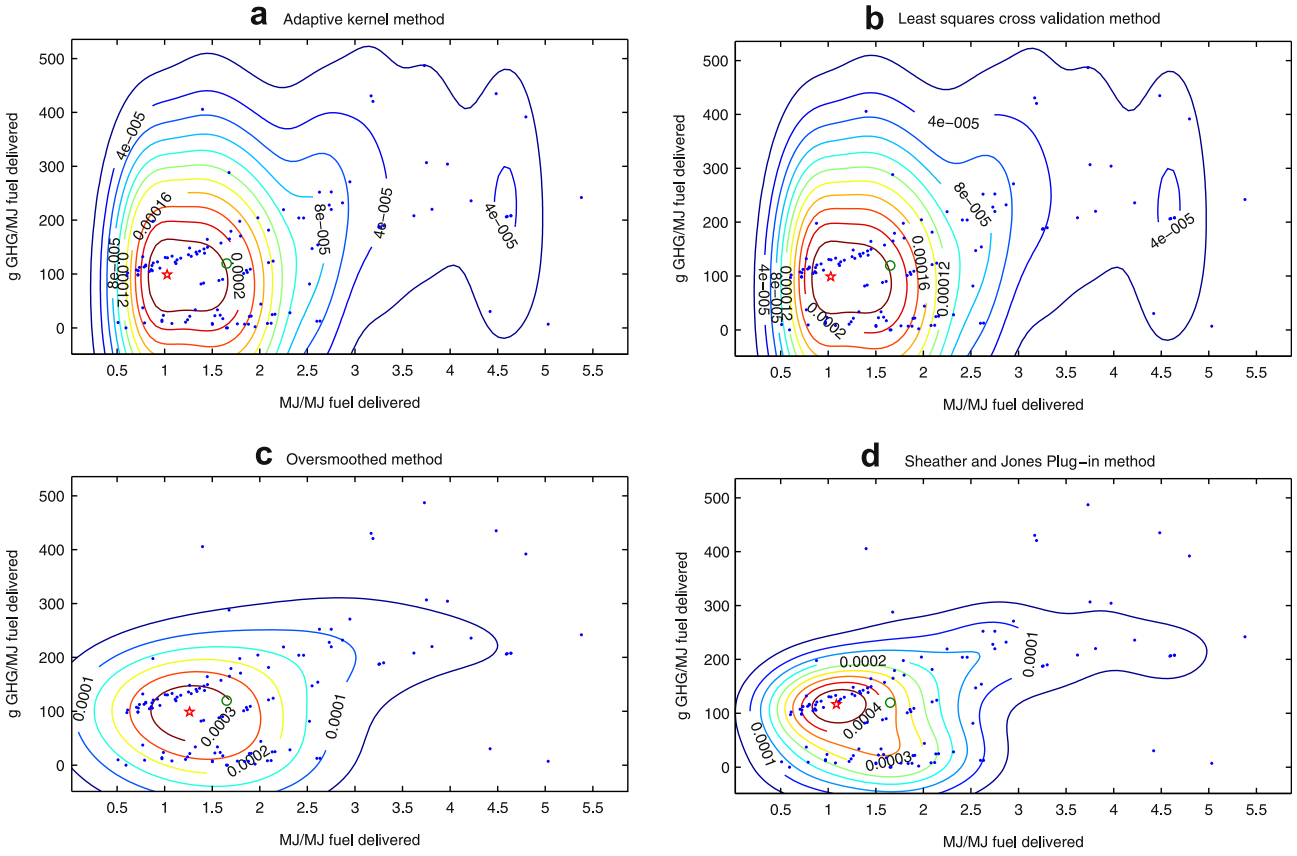


Fig. 2 – WTT estimates (blue dots) of supplying hydrogen across all pathways with contour plots (lines) of probability distribution of WTT MJ/MJ and g GHG/MJ impacts across the a) adaptive kernel, b) least squares CV, c) oversmoothed and d) SJPI methods, respectively. The distribution peak, indicating the best estimate, is denoted with a green circle and the median of the distribution given by a red star. Iso-contours are labelled with the corresponding distribution probability. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

transported by pipeline or road to the filling station. Alternatively, the electricity may be transmitted to the filling station for decentralized, onsite electrolysis.

The steam reforming pathway begins with natural gas extraction, with three possible downstream paths. In the first instance, the hydrogen is produced by steam reforming of the natural gas at the point of extraction (wellhead). The second option is to compress or liquefy the natural gas for long distance transport by pipeline or ship to a centralized steam reforming plant. The hydrogen thus produced is liquefied or compressed for transport by pipeline or road to the filling station. The remaining option is to pipe natural gas to filling stations for decentralized, onsite reforming.

The non-parametric method is repeated on a quasi-static basis. An error analysis is conducted between the adaptive kernel and other non-parametric density estimation methods – oversmoothed and CV approach (first generation), the SJPI (second generation) – and the median. There is a further investigation into the effects of including the weighting factor. Statistical robustness refers to the insensitivity of an estimate to small deviations from the assumptions, resulting in a large breakdown point. The breakdown point is the smallest fraction of errant observations which skews the estimate to an arbitrarily large (magnitude) value [18]. The median is used instead of the mean as the former has the maximum breakdown point of 0.5, whereas the mean has the smallest breakdown point of $1/n$ [19]. The oversmoothed estimator is given by $h_{os} = 1.144 * \text{data}_{iqr} * n^{-1/5}$ [7], where data_iqr is the interquartile range of the data and n is the number of data points. The SJPI method was given in [20] [3]to yield h_{sjpi}. Both h_{os} and h_{sjpi} were applied to the bivariate Gaussian function the method already described.

3. Analysis and discussion

A scatter plot of primary WTT estimates across the hydrogen pathways illustrates the range values, demarked by ellipses [21] (Fig. 1). The 124 estimates used in this study derive from 13 publicly available, published sources [2,22–33]. Two of the most popularly cited sources are the Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation

[3] An m-file implementation of this method is available at http://www.mathworks.com/matlabcentral/fileexchange/22999.

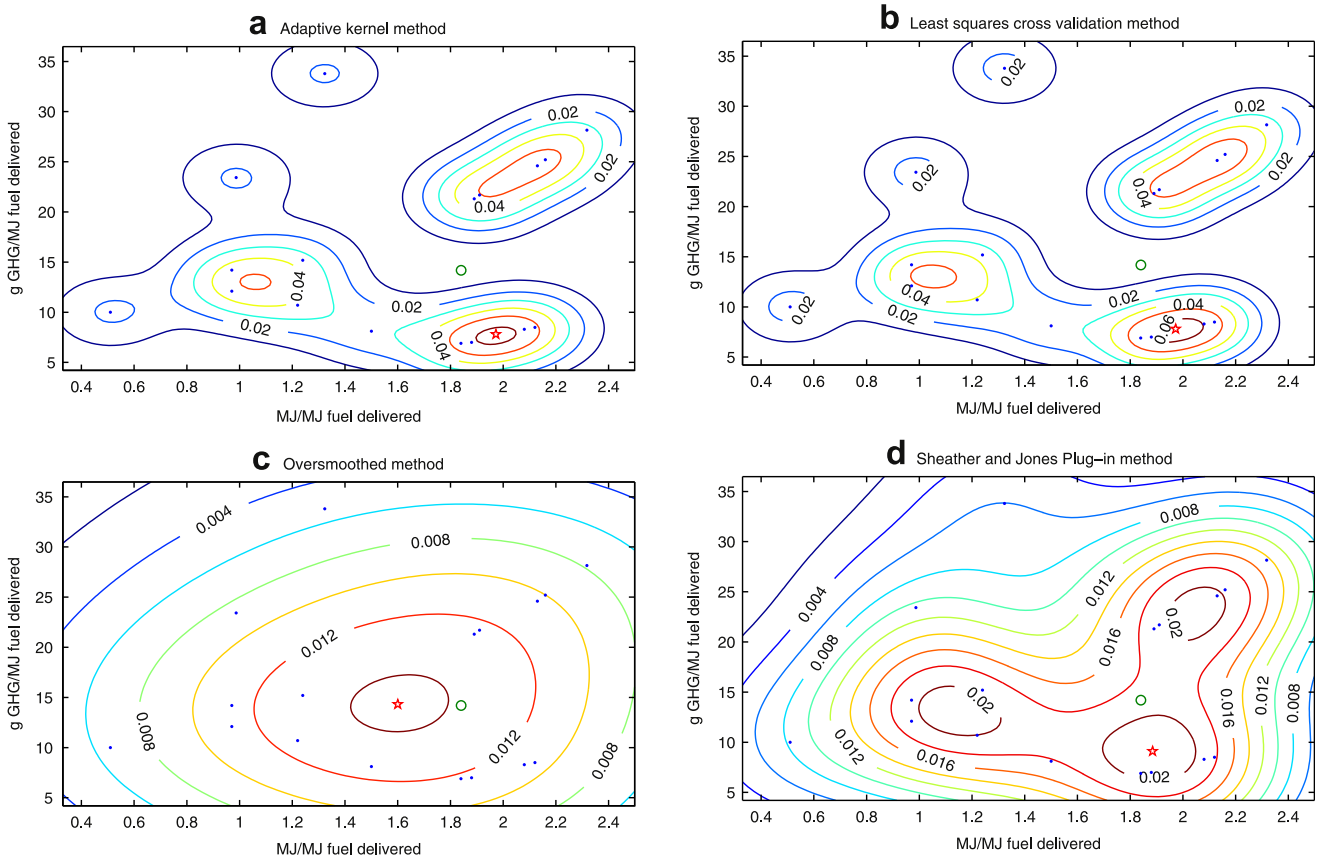


Fig. 3 – WTT estimates (blue dots) of supplying hydrogen by gasification with contour plots (lines) of probability distribution of WTT MJ/MJ and g GHG/MJ impacts across the a) adaptive kernel, b) least squares CV, c) oversmoothed and d) SJPI methods, respectively. The distribution peak, indicating the best estimate, is denoted with a green circle and the median of the distribution given by a red star. Iso-contours are labelled with the corresponding distribution probability. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(GREET) Model [25] and Concawe reports [2] and are included in this analysis. Studies which re-state the findings from any work already included in this research, are omitted to avoid duplication.

A noticeable feature of the probability density functions (Figs. 2–5) is the presence of multiple modes which can be attributed to the different primary resources which are used in the particular production pathway and year of WTT estimates. The modes are most noticeable for gasification (Fig. 3) and electrolysis (Fig. 4).

Across all pathways, there are three *strata* for low, intermediate and high g GHG/MJ impacts across all MJ/MJ energy use. The group of estimates with low GHG impacts consist of gasification (green ellipse in Fig. 1) and electrolysis using electricity from renewable sources. The intermediate GHG impacts arise from steam reforming (red ellipse in Fig. 1) of natural gas and electrolysis using electricity from gas or country mixes with a large proportion of gas. The high GHG impacts *stratum* represents electrolysis from electricity using coal and country mixes with a large proportion of coal. The majority (102, or 82%) of the 124 WTT estimates lie at less than 2.8 MJ/MJ and in the intermediate band of GHG impacts, coinciding with the position of the distribution peak (red star in Fig. 2).

The probability density function for gasification indicates three modes which depend more so on the wood/biomass type than on the gasification process. The three modes reflect high energy and high GHG impacts, low energy with intermediate GHG impacts and low energy use with high GHG impacts. The first case represents the pathways using poplar wood as the primary resource. Low energy use and intermediate GHG impacts derive from both farmed and waste wood, in central and onsite gasification. The high energy, low GHG impact pathways are generally supplied with woody biomass, farmed or waste. The results are more sensitive to the wood/biomass source, as demonstrated with the poplar, implying that the wood type be included in future WTT estimates for clarity.

The probability density function for hydrogen from electrolysis indicates three *strata* of increasing g GHG/MJ impact, corresponding to electricity from renewables, gas and coal, respectively. The three groups of WTT estimates are: those with g GHG/MJ less than 100; between 100 and 300 g GHG/MJ; and greater than 300 g GHG/MJ. The estimates in the first

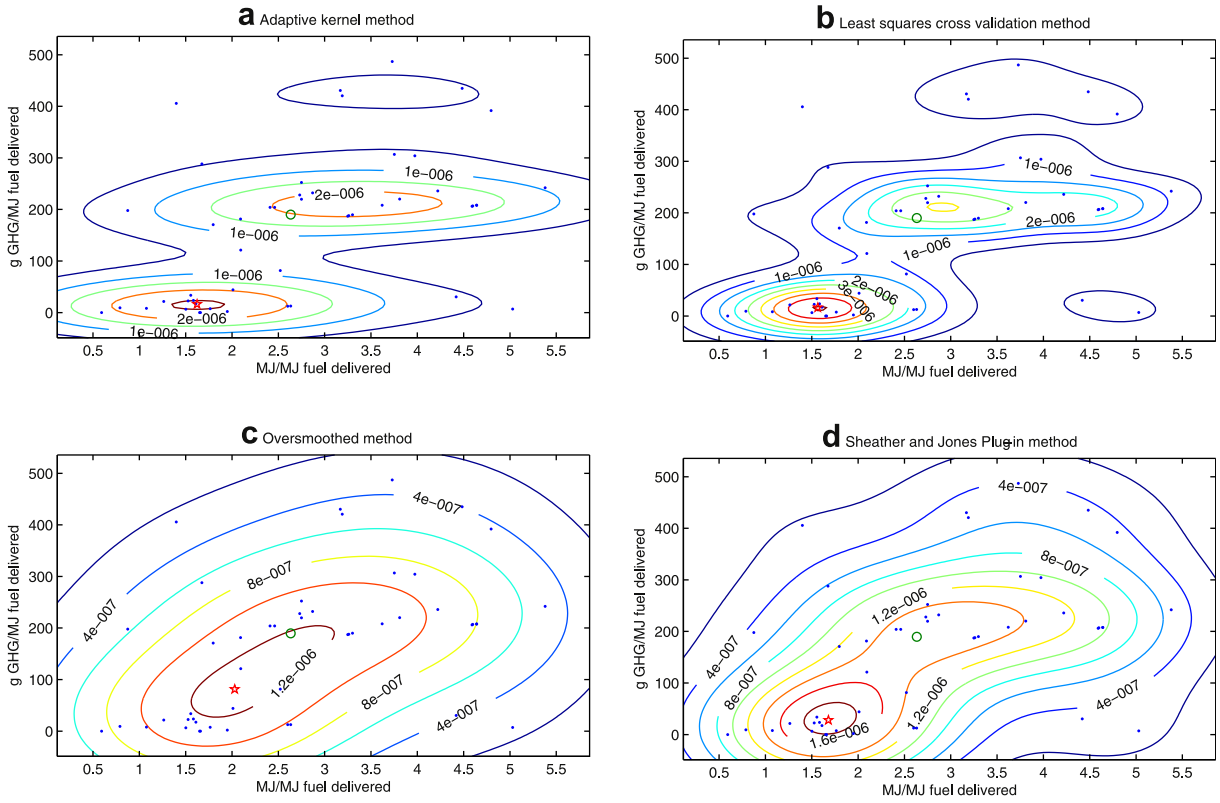


Fig. 4 – WTT estimates (blue dots) of supplying hydrogen from electrolysis with contour plots (lines) of probability distribution of WTT MJ/MJ and g GHG/MJ impacts across the a) adaptive kernel, b) least squares CV, c) oversmoothed and d) SJPI methods, respectively. The distribution peak, indicating the best estimate, is denoted with a green circle and the median of the distribution given by a red star. Iso-contours are labelled with the corresponding distribution probability. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

stratum correspond to electrolysis from renewables; those in the intermediate band represent electrolysis using natural gas and country mix primarily; and WTT estimates in the upper band utilize coal and country mix. The country mixes in the intermediate band are from the European studies, which are largely natural gas based, while those in the upper *stratum* are the US mixes, of which coal is the major constituent fuel.

The probability density function for hydrogen from steam reforming of natural gas shows two main trend lines, largely divided by the year of WTT estimate. Both subsets of WTT estimates span the range of energy uses. The larger, longer group accounts for intermediate to high energy GHG impacts, while the smaller set covers low to medium GHG impacts. The two groups are completely separable by a line $y = 69.5x + 19.6$.[4] Both groups contain estimates for EU and non-EU natural gas, reformed centrally and onsite, with compression and liquefaction. The larger group contains 43 of the 55 estimates, or 78.2% and comprises the newest WTT studies from 2005 [22] and 2007 [2,25], save one estimate from 2000 [24]. Conversely, the smaller group is largely represented by 2002 estimates [23], with two estimates from 2006 [28,29].

[4] This separating line joins the median points between the two extremes of each subset.

Three further observations (Figs. 2–5) justify the use of the kernel density estimator method: inappropriateness of the median as a best estimate approximation; multimodal distributions; and large bias of the CV approach. The median (green circle) is a statistically robust measure of the centre of the data. However, it is not coincident with the distribution peak (red star). Therefore, this measure of central tendency does not offer an accurate WTT best estimate based on the data.

All the methods agree on which peak is the largest, although only the distributions which derive from the least squares CV and adaptive kernel methods reflect the multimodal features. The oversmoothed method draws the distribution into a single large mountain and associated lone peak, while the SJPI approach resolves individual peaks upon the same large mountain. This contrasts the output from least squares CV and adaptive kernel which shows almost-completely distinct hills. A simple histogram, based on a unimodal Gaussian distribution would fail to account for the multiple features of these datasets. Moreover, the simple oversmoothed and more advanced SJPI fail to capture the finer distribution details.

The least squares CV method can produce a distribution which is too variable – true and potentially spurious features – due to large bias [7]. The distribution derived

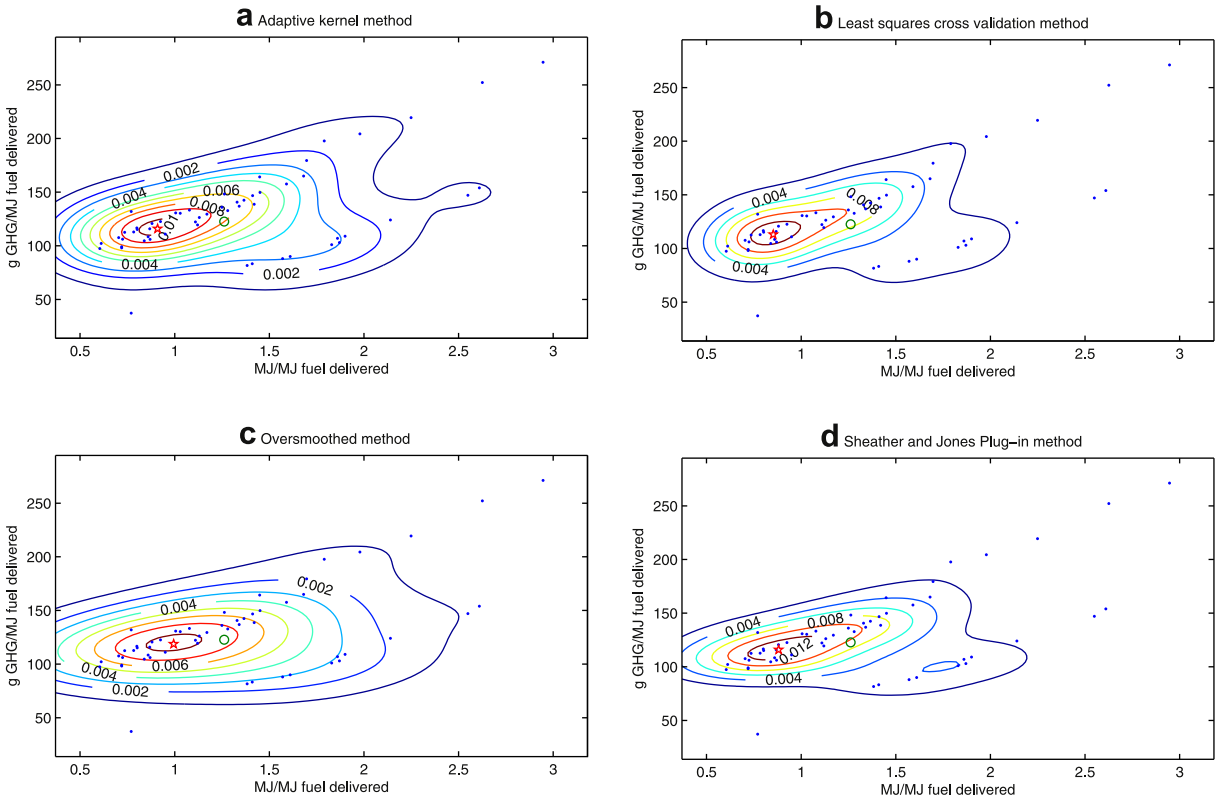


Fig. 5 – WTT estimates (blue dots) of supplying hydrogen from steam reforming with contour plots (lines) of probability distribution of WTT MJ/MJ and g GHG/MJ impacts across the a) adaptive kernel, b) least squares CV, c) oversmoothed and d) SJPI methods, respectively. The distribution peak, indicating the best estimate, is denoted with a green circle and the median of the distribution given by a red star. Iso-contours are labelled with the corresponding distribution probability. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3 – Best estimates and best-in-class pathways for hydrogen supply WTT impacts overall and by pathway. The error is the absolute Euclidean difference between the best estimate and the nearest actual pathway from the WTT estimates set. The units for the Euclidean distance are g GHG per MJ used to produce the hydrogen (input). A "-" indicates that there is no error for the best-in-class pathways since they represent actual WTT estimates.

Best estimate with weighting factor	MJ/MJ	g GHG/MJ	Specific route	Error g GHG/MJ input
Overall	1.09	92.97	EU natural gas mix, central reforming [23]	3.02
Gasification	1.97	7.80	Residual woody biomass, onsite gasification [23]	0.51
Electrolysis	1.62	16.23	Electricity from solar photovoltaics, central electrolysis [25]	1.06
Steam reforming	0.85	114.91	Natural gas in North America, central reforming [22]	0.26
Best-in-class				
Overall	0.59	0.00	Electricity from renewables in USA, onsite electrolysis [22]	–
Gasification	1.84	6.60	Residual woody biomass, onsite gasification [23]	–
Electrolysis	0.59	0.00	Electricity from renewables in USA, onsite electrolysis [22]	–
Steam reforming	1.20	35.71	Thermal cracking of natural gas [33]	–
Overall	1.09	92.97	EU natural gas mix, central reforming [23]	3.02
Gasification	1.97	7.80	Residual woody biomass, onsite gasification [23]	0.51
Electrolysis	1.62	16.23	Electricity from solar photovoltaics, central electrolysis [25]	1.06
Steam reforming	0.85	114.91	Natural gas in North America, central reforming [22]	0.26

from the adaptive kernel method is less variable (noticeable in Fig. 4) and may represent a better blend of bias and variance.

The overall best estimate for providing 1 MJ of hydrogen (Table 3) uses natural gas from the EU with centralized reforming [23], resulting in WTT impacts of 1.1 MJ energy required and 93 g GHG emitted. The overall best-in-class hydrogen supply pathway is from the electrolysis using electricity from renewables in the USA [22], with impacts of 0.6 MJ/MJ and 0 g GHG/MJ. The emissions-free supply of hydrogen using the best-in-class pathway disregards the pollution which arises during the other stages of the wind farm life cycle [34]. The exclusion of externalities highlights the difficulty of comparing WTT estimates across studies on account of the varying assumptions and system boundaries which the non-parametric, adaptive kernel density approach addresses.

There is a trade-off between the pathways which minimize energy and GHG impacts, respectively. The highest WTT MJ/MJ impact pathway for supplying hydrogen is by gasification of residual woody biomass (2.0 MJ/MJ, 7.8 g GHG/MJ) [23]. Steam (central) reforming of natural gas in North America carries the largest WTT GHG burden (0.9 MJ/MJ, 114.9 g GHG/MJ) [2]. If the priority is to reduce emissions, the aforementioned gasification pathway is best. Conversely, an objective to reduce the energy used to make the hydrogen suggests that steam reforming of natural gas should be chosen. In the limit, using the state-of-the-art, deriving hydrogen from electricity generated from renewable sources gives the lowest energy and GHG WTT impacts.

When a quasi-static timing is introduced, MJ/MJ energy efficiency improves for the best estimate at the expense of higher g GHG/MJ, while both energy and emissions efficiency increase through time for the best-in-class pathway. The time trajectories of the overall best estimates and best-in-class pathways from the oldest estimate to the most recent one (Fig. 6) are not linear. In general, natural gas pathways constitute the overall best estimates by year, whereas electricity from renewables is the best-in-class pathway for delivering low GHG-emitting hydrogen through from 2000 to 2010 (Tables 4 and 5). Viewing energy impacts alone, the best estimates and best-in-class pathways show improvements of 5.4% and 5.1% per year. Conversely, annual g GHG/MJ impacts worsen for the best estimate at 1.8%, while improving for the best-in-class pathway at 4.7%. By pathway, the trend of increasing MJ/MJ energy efficiency with higher g GHG/MJ impacts is observed over the best estimates and best-in-class pathways. Exceptionally, both the electrolysis best estimate and its equivalent best-in-class pathway show decreasing MJ/MJ and g GHG/MJ efficiencies from 2000 to 2010. The Euclidean distance between the first and last years in the time series is used to combine the two WTT impacts into an overall change in impacts. The units for the Euclidean distance are g GHG per MJ used to produce the hydrogen

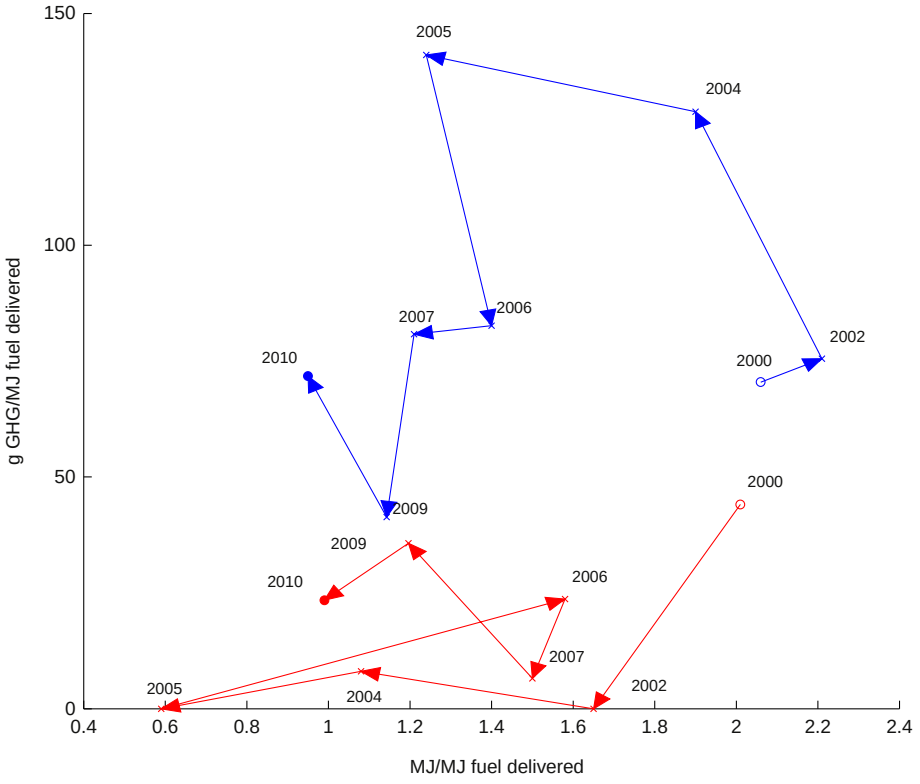


Fig. 6 – Trajectory of overall best estimates (blue) and best-in-class pathways (red) WTT impacts. The first year of the time series is 2000 (open circle) and arrows indicate the direction of time to the most recent WTT estimate in 2010 (closed circle). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4 – Best estimates and best-in-class pathways for hydrogen supply WTT impacts overall and by pathway, by year, when data is available. The number of estimates used for each year is given by *n*. The error is the absolute Euclidean difference between the best estimate and the nearest actual pathway from the WTT estimates set. The units for the Euclidean distance are g GHG per MJ used to produce the hydrogen (input). Where there is no error, either because there is only one estimate or the best-in-class pathway is being represented, a "-" is used (Table 1 of 2).

Overall best estimate	*n*	MJ/MJ	g GHG/MJ	Specific route	Error g GHG/MJ input
2000	3	2.06	70.46	Electricity from solar photovoltaics in Algeria, central electrolysis, liquefied [26]	10.98
2002	29	2.21	75.53	Natural gas in the EU, central reforming [23]	12.48
2004	3	1.90	128.81	Electricity from wind, electrolysis [30]	120.73
2005	29	1.24	141.05	Natural gas in North America, central reforming [22]	1.48
2006	4	1.40	82.66	Natural gas, reformed [28]	0.54
2007	50	1.21	80.80	Natural gas, piped 4000 km, central reforming [2]	17.40
2009	3	1.14	41.37	Autocatalytic decomposition of 50% methane [33]	3.27
2010	3	0.95	71.75	Compressed wood, gasified [27]	48.32
Best-in-class					
2000	3	2.02	44.88	Electricity from solar photovoltaics in Canada, central electrolysis, liquefied [26]	–
2002	29	1.65	0.00	Electricity from wind, central electrolysis [23]	–
2004	3	1.90	128.81	Electricity from wind, electrolysis [30]	–
2005	29	0.59	0.00	Electricity from renewables in USA, onsite electrolysis [22]	–
2006	4	1.58	23.68	Electricity from wind energy [29]	–
2007	50	1.50	6.55	Electricity from nuclear energy, central electrolysis [25]	–
2009	3	1.20	35.71	Thermal cracking of natural gas [33],	–
2010	3	0.99	23.43	Gasification of wood [27]	–
Gasification best estimate					
2002	8	1.90	21.48	Biomass from poplar, central gasification [23]	0.89
2007	8	1.12	13.32	Farmed wood, central gasification [2]	0.18
2010	1	0.99	23.43	Non-North American natural gas, central gasification [22]	–
Best-in-class					
2002	8	1.84	6.90	Biomass from residual wood, central gasification [23]	–
2007	8	1.50	8.10	Farmed wood, central gasification, liquefied [2]	–
2010	1	0.99	23.43	Non-North American natural gas, central gasification [22]	–

(input). For the best estimates, there is a 0.2% annual decrease in impacts. For the best-in-class pathways, WTT impacts decrease by 4.7% annually. It may be inferred that the estimates reflect the focus of the literature of the day and by extension, the pathway(s) which are receiving the most attention. However, it is acknowledged that the trajectories will be affected by the small datasets used to find the best estimate and best-in-class pathway for each year. Therefore, the quasi-static analysis presented should be viewed as an exercise in the power of the adaptive kernel density estimator method instead of the basis for drawing firm conclusions on the progress of WTT hydrogen production pathways through time.

3.1. *Error analysis*

The absolute error between the adaptive kernel density method and other non-parametric approaches decreases from the least squares CV to the SJPI, oversmoothed and median. The median absolute error across the alternative methods is 16.3 g GHG/MJ input (Table 6). The least-squares CV method offered results which were closest to those of the adaptive kernel, consistent with expectations the adaptive kernel method is based on least squares CV. At the other end of the scale, using the median gives errors up to 173.8 g GHG/MJ input, consistent with the large differences between the red star and green circle (Figs. 2-5).

The difference between the real-world pathways chosen by different methods is more important than the peak in their respective distributions. In considering pathways, no single method yields a real-world pathway that either consistently matches or is consistently different from that chosen by the adaptive kernel density approach. However, at least one of the estimates using an alternative method yields a pathway which equals that chosen by the adaptive approach.

Across all methods and studies, the best estimate pathways to supply hydrogen use steam reforming of natural gas. The least squares CV approach chooses the same WTT pathway as the adaptive kernel, while each of the other methods select different pathways each.

For gasification, the least squares CV and SJPI approaches select the same real-world pathway as the adaptive method, while the median and oversmoothed choose the same pathway as each other, but different from the adaptive kernel method. The electrolysis production process sees different methods across each of the four methods, with least squares CV proposing the same one as the adaptive method. For steam reforming, the least squares CV and SJPI choose the same natural gas pathway as the adaptive methods. The median

Table 5 – Best estimates and best-in-class pathways for hydrogen supply WTT impacts overall and by pathway, by year, when data is available. The number of estimates used for each year is given by *n*. The error is the absolute Euclidean difference between the best estimate and the nearest actual pathway from the WTT estimates set. The units for the Euclidean distance are g GHG per MJ used to produce the hydrogen (input). Where there is no error, either because there is only one estimate or the best-in-class pathway is being represented, a "-" is used (Table 2 of 2).

Electrolysis best estimate	*n*	MJ/MJ	g GHG/MJ	Specific route	Error g GHG/MJ input
2000	2	2.02	44.88	Electricity from solar photovoltaics in Canada, central electrolysis, liquefied [26]	0.79
2002	13	4.62	207.53	Electricity from the EU mix, central electrolysis [23]	0.47
2004	3	1.90	128.81	Electricity from wind, electrolysis [30]	120.73
2005	5	1.76	146.33	Electricity from natural gas in a CCGT, onsite electrolysis [22]	24.26
2006	2	1.56	25.11	Electricity from wind [29]	1.43
2007	24	1.50	16.74	Electricity from solar photovoltaics, central electrolysis [25]	0.56
2010	1	2.09	121.26	Electrolysis [27]	–
Best-in-class					
2000	2	2.02	44.88	Electricity from solar photovoltaics in Canada, central electrolysis, liquefied [26]	–
2002	13	1.65	0.00	Electricity from wind, central electrolysis [23]	–
2004	3	1.90	128.81	Electricity from wind, electrolysis [30]	–
2005	5	2.09	181.21	Electricity from USA mix in California, onsite electrolysis [22]	–
2006	2	1.58	23.68	Electricity from wind energy [29]	–
2007	24	1.50	6.55	Electricity from nuclear energy, central electrolysis [25]	–
2010	1	0.99	23.43	Gasification of wood [27]	–
Steam reforming best estimate					
2000	1	0.77	132.00	Natural gas, onsite reforming [24]	–
2002	9	1.86	107.00	Natural gas from Russia, central reforming [23]	1.60
2005	24	1.13	134.77	Non-North American natural gas, central reforming [22]	1.19
2006	2	1.38	81.67	Natural gas from North America, central reforming [29]	0.03
2007	18	0.84	108.72	Liquified natural gas, onsite reforming [2]	1.22
2009	3	1.14	41.37	Autocatalytic decomposition of 50% methane [33]	3.27
2010	1	0.88	120.60	Natural gas [27]	–
Best-in-class					
2000	1	0.77	132.00	Natural gas, reformed onsite [24]	–
2002	9	1.57	88.00	Natural gas from the EU, central reforming [23]	–
2005	24	0.60	97.39	Natural gas from North America, central reforming [22]	–
2006	2	1.38	81.67	Natural gas from North America, central reforming [29]	–
2007	18	0.77	37.20	Natural gas, piped 4000 km, central reforming with CCS [2]	–
2009	3	1.20	35.71	Thermal cracking of natural gas [33],	–
2010	1	0.88	120.60	Natural gas [27]	–

and oversmoothed methods yield different natural gas pathways each (Table 7).

There is no difference in results – overall and by pathway – when the weighting factor is present in the adaptive kernel method (Table 3). However, the oversmoothed and least squares CV methods yield different results without the weights. The adaptive kernel method uses different bandwidths depending on the density of the data: when the data is clustered, the bandwidth is small to capture the features; and when the data is sparse, the bandwidth is large to avoid spurious peaks. Conversely, the first and second generation methods to which the adaptive kernel method is compared use a single bandwidth across the entire data set. Therefore, it is inferred that the adaptive kernel method is biased by the relative density of data points which overshadows the resource potential weighting factor. In contrast, the fixed bandwidth oversmoothed, least squares CV and SJPI methods do not introduce a data clustering bias, leaving the weighting factor to influence the peak of the probability density function.

Table 6 – Best WTT estimates for hydrogen based on median, oversmoothed, least squares CV and SJPI methods. The error to adaptive kernel density best estimate is the absolute Euclidean distance to the best estimates found by the other methods. The error to real-world pathway is the absolute Euclidean distance from real-world pathway corresponding to the adaptive kernel density best estimate to the best estimates found by the other methods.

Pathway	Median		Oversmoothed		Least squares CV		SJPI	
	MJ/MJ	g GHG/MJ	MJ/MJ	g GHG/MJ	MJ/MJ	g GHG/MJ	MJ/MJ	g GHG/MJ
Overall	1.61	118.04	1.26	98.88	1.09	92.97	1.09	116.58
Gasification	1.84	14.20	1.60	14.32	1.97	7.80	1.89	9.10
Electrolysis	2.63	190.00	2.03	81.17	1.56	16.23	1.68	28.04
Steam reforming	1.22	122.40	0.97	120.61				
Error to adaptive kernel density best estimate								
Minimum	6.40		5.71		0.00		0.03	
Median	16.29		6.22		0.01		6.56	
Maximum	173.77		64.93		0.06		23.61	
	62.25	1070.43	25.03	400.00	0.00	0.00	3.57	72.73
Error to real-world pathway								
Minimum	5.90		4.87		0.51		0.82	
Median	61.35		34.95		0.95		5.79	
Maximum	172.71		96.88		90.98		114.59	

Table 7 – Hydrogen production pathways corresponding to the best WTT estimates using the median, oversmoothed, least squares CV and SJPI methods.

Method	Pathway			
	All	Gasification	Electrolysis	Steam reforming
Median	Liquefied natural gas (LNG), Onsite reforming [2]	Farmed wood, central gasification [2]	Electricity from natural gas in CCGT, onsite electrolysis [23]	Natural gas, piped 7000 km, onsite reforming [2]
Oversmoothed	Natural gas, piped 4000 km central reforming, trucked to station [2]	Farmed wood, central gasification [2]	Electricity from solar photovoltaics in Algeria, electrolysis, liquefied [26]	Compressed natural gas [27]
Least squares CV	EU natural gas mix, Central reforming [23]	Residual woody biomass, onsite gasification [23]	Electricity from solar photovoltaics, central electrolysis [25]	North American natural gas, onsite reforming [22]
SJPI	Non-North American natural gas, Onsite reforming via LNG [22]	Residual woody biomass, onsite gasification [23]	Electricity from farmed wood, in conventional power plant, onsite electrolysis [2]	North American natural gas, onsite reforming [22]

4. Conclusions

This research has applied non-parametric statistical methods to understand the distribution of WTT estimates for producing hydrogen through combined and single pathways (gasification, electrolysis and steam reforming). This is justified by the expanding set of WTT estimates which are based on different assumptions and boundary conditions. The weighted adaptive kernel density estimate is proposed as a flexible, expandable method for extracting the unknown probability density function using a data-driven method, on both static and quasi-static bases.

The adaptive kernel density method offers a better estimate than the first and second generation non-parametric methods (oversmoothed, least squares CV, and SJPI) and the median. The adaptive method captures the multimodal features of the distribution better than the oversmoothed and SJPI without the same degree of variability as the least squares CV. The median is the least appropriate measure for finding the peak of the underlying distribution of the WTT estimates. In the case of hydrogen production, adding the resource weighting had no effect on the best estimate output of the adaptive method. This is attributed to the adaptive kernel already applying a weight to the bandwidth based on the density of the data. Conversely, the fixed bandwidth methods exhibit changes when weighting is present versus when it is absent. These methods use fixed bandwidths such that the only biasing present in the distribution arises from the applied resource potential weight. However, it is suggested that a resource weighting will be a suitable addition for other WTT estimate distributions.

Overall, the best estimate pathway for supplying hydrogen is by natural gas from the EU mix with centralized reforming. Onsite gasification of woody biomass, electricity from renewables in the EU and North American natural gas centrally reformed are the best estimates pathways for providing hydrogen from gasification, electrolysis and steam reforming, respectively. There is an observed trade-off between the lowest GHG-emitting pathway and that requiring the least energy per MJ hydrogen delivered. A quasi-static analysis of the hydrogen supply overall and by pathways indicates that energy efficiency has increased at the expense of emissions efficiency through time.

The best-in-class pathway is that with the lowest GHG per MJ hydrogen supplied and represents the state-of-the-art. In the limit of the externalities associated with the remainder of the power plant life cycle and supply chains being disregarded, electrolysis from renewables in the USA potentially provides a zero GHG/MJ fuel option. This paper does not address the technical, economic, or political barriers to implementing hydrogen as a transport fuel and the proposed method should not be seen in isolation from other elements of advice for decision-makers.

Acknowledgements

The authors acknowledge the funding provided by the Oxford Martin School for this work.

Notations and Abbreviations

- WTT: Well-to-tank;
- WTW: Well-to-wheel;
- MISE: Mean integrated squared error;
- CV: Cross-validation;
- h: optimum bandwidth of the data set used to create probability density estimate which balances bias and variability;
- m: the number of bins, where $m = 1/h$;
- MJ/MJ: The energy expended, excluding that transferred to the final fuel, per MJ of energy content in the final fuel [1];
- GHG: greenhouse gas;
- g GHG/MJ: The emissions (g GHG) associated with each stage of energy use in the WTT pathway;
- Best estimate pathway: The pathway which is closest to peak of the probability distribution which results from the non-parametric estimation method;
- Best-in-class pathway: The WTT pathway with the lowest GHG emissions per MJ of hydrogen delivered;
- SJPI: Sheather and Jones plug-in method, a second generation bandwidth estimation method which offers better performance than first generation rules-of-thumb and CV approaches;
- n: number of entries in the data set;
- h_{binned}: The bandwidth h corresponding to the minimum CV value; and
- h_{adap}: The vector of optimum h corresponding to the density of data in each bin, m.

Math formulae

$$\text{weight overall}_i = \frac{\text{primary resource capacity}_i}{\max(\text{primary resource capacity})_i} \times \forall i \text{ pathways} \quad (1)$$

$$lsqcv(i) = f_1(i) + f_2(i) + f_3(i); \quad (2)$$

$$f_1(i) = \frac{1}{n^2}\sum_{i=1}^{m}\sum_{j=1}^{n}\text{freq1}(i)^2.[G(t(i),h(i),c) * G(t(i),h(i),c)]; \quad (3)$$

where:

- n is the number of estimates in the data vector;
- freq1(i) is the frequency count of estimates in bin i;
- c=[min_x−2*max_h, max_x+2*max_h], incremented by (*max_x*+2**max_h*−min_*x*−2**max_h*)/*vec_length* where vec_length = 50;
- max_x, min_x, max_h, min_h are the maximum values of data and h, where $h(i)$=*data_range*/i and *data_range*=*max_x*−*min_x*;
- * represents a two dimensional convolution;
- t(i) is the mid-point of the bin i; and
- G is a univariate Gaussian function with $\mu = 0$, $\sigma = h$.

$$f_2(i) = \frac{1}{n^2}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{m} n_i.n_k.[G(t(i),h(i),c) * G(t(k),h(k),c)]; \quad (4)$$

where:

- $n_i.n_k = \text{freq1}(i).\text{freq1}(k)\,\forall i \neq k$, 0 otherwise.

$$f_3(i) = \frac{-2}{n(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{m} n_{ik}.G(t(i),h(i),c); \quad (5)$$

where:

- $n_{ik} = \text{freq1}(i)\,\forall i \neq k$, $n_{ik} = \text{freq1}(i) - 1$ otherwise.

$$f_{adap}(i) = \sum_{i=1}^{\text{optimum\_bin}}\sum_{j=1}^{freq(i)} G(t(i),h(i),c); \quad (6)$$

REFERENCES

[1] Edwards R, Larivé JF, Mahieu V, Rouveirolles P. Well-to-wheels analysis of future automotive fuels and powetrains in the European context: well-to-tank report Appendix 2. Version 2c. Brussels: European Commission Joint Research Centre. URL, http://ies.jrc.ec.europa.eu/uploads/media/WTW_Report_010307.pdf; 2007.
[2] Edwards R, Larivé JF, Mahieu V, Rouveirolles P. Well-to-wheels analysis of future automotive fuels and powetrains in the European context. Version 2c. Brussels: European Commission Joint Research Centre. URL, http://ies.jrc.ec.europa.eu/uploads/media/WTW_Report_010307.pdf; 2007.
[3] Loader CR. Bandwidth selection: classical or plug-in? Ann Stat 1999;27:415–38.
[4] Sain SR. Multivariate locally adaptive density estimation. Comput Stat Data An 2002;39:165–86.
[5] Bura E, Zhumrov A, Barsegov V. Nonparametric density estimation and optimal bandwidth selection for protein unfolding and unbinding data. J Chem Phys 2009;130:1–15.
[6] Turlach BA. Bandwidth selection in kernel density estimation: a review. Berlin: Humboldt Universitaet Berlin; 1993.
[7] Sheather SJ. Density estimation. Stat Sci 2004;19:588–97.
[8] Knuth KH. Optimal data-based binning for histograms. Ithaca: Cornell University Library. URL, http://arxiv.org/abs/physics/0605197; 2006.

[9] Chen SX. Probability density function estimation using gamma kernels. Ann Stat Math 2000;52:471–80.
[10] Jones MC, Henderson DA. Maximum likelihood kernel density estimation: on the potential of convolution sieves. Comput Stat Data An 2009;53:3726–33.
[11] Hall P, Sheather SJ, Jones MC, Marron JS. On optimal data-based bandwidth selection in kernel density estimation. Biometrika 1991;78:263–9.
[12] Jones MC, Marron JS, Sheather SJ. A brief survey of bandwidth selection for density estimation. J Am Stat Assoc 1996;91:401–7.
[13] Sheather SJ, Jones MC. A reliable data-based bandwidth selection method for kernel density estimation. J Roy Stat Soc B Met 1991;53:683–90.
[14] Brabanter KD, Brabanter JD, Suykens JAK, Moor BD. Optimized fixed-size kernel models for large data sets. Comput Stat Data An 2010;54:1484–504.
[15] Liao JG, Wu Y, Lin Y. Improving Sheather and Jones' bandwidth selector for difficult densities in kernel density estimation. J Nonparametr Stat 2009;22:105–14.
[16] Sain SR. Adaptive kernel density estimation. Ph.D. thesis. Houston: Rice University; 1994.
[17] Sain SR, Scott DW. On locally adaptive density estimation. J Am Stat Assoc 1996;91:1525–34.
[18] Huber PJ, Ronchetti EM. Robust statistics. 2nd ed. Hoboken: John Wiley & Sons; 2009.
[19] Wilcox RR. Fundamentals of modern statistical methods. 2nd ed. New York: Springer; 2010.
[20] Simonoff JS. Smoothing methods in statistics. New York: Springer; 1996.
[21] Moshtagh N. Minimum volume enclosing ellipsoids. Natick: MATLAB central file Exchange. URL, www.mathworks.com/matlabcentral/fileexchange/9542; 2009.
[22] Argonne National Laboratory. Well-to-wheels analysis of advanced fuel/vehicle systems - a North American study of energy use, greenhouse gas emissions, and criteria pollutant emissions. Chicago: Argonne National Laboratory. URL, http://www.transportation.anl.gov/pdfs/TA/339.pdf; 2005.
[23] Ludwig Bölkow Systemtechnik GmbH. GM well-to-wheel analysis of energy use and greenhouse gas emissions of advanced fuel/vehicle systems - a European study. Munich: Ludwig Bölkow Systemtechnik GmbH. URL, http://www.lbst.de/ressources/docs2002/TheReport_Euro-WTW_27092002.pdf; 2002.
[24] Weiss MA, Heywood JB, Drake EM, Schafer A, AuYeung FF. On the Road in 2020: a life-cycle analysis of new automobile technologies. MIT EL 00–003. Massachusetts: Massachusetts Institute of Technology. URL, http://lfee.mit.edu/public/el00-003.pdf; 2000.
[25] Argonne National Laboratory. Greenhouse gases, Regulated Emissions and Energy use in Transportation (GREET). Chicago: Argonne National Laboratory. URL, www.transportation.anl.gov/modeling_simulation/GREET/index.html; 2007.
[26] Oeko-Institut. Global Emission Model for Integrated Systems (GEMIS 4.5) for embodied greenhouse gas emissions in energy conversion systems and biofuels. Freiburg: Oeko-Institut. URL, http://www.oeko.de/service/gemis/en/; 2008.
[27] Natural Resources Canada. GHGenius 3.18: a model for lifecycle assessment of transportation fuels. Ottawa: Natural Resources Canada. URL, www.ghgenius.ca; 2010.
[28] Beer T, Grant T, Morgan G, Lapszewicz J, Anyon P, Edwards J, et al. Comparison of transport fuels: final report to the Australian Greenhouse Office on the stage 2 study of life-cycle emissions analysis of alternative fuels for heavy vehicles. EV45A/2/F3C. Clayton: Commonwealth Scientific and Industrial Research Organisation; 2006.
[29] Granovskii M, Dincer I, Rosen MA. Life cycle assessment of hydrogen fuel cell and gasoline vehicles. Int J Hydrogen Energ 2006;31:337–52.
[30] Spath PL, Mann MK. Life cycle assessment of renewable hydrogen production via wind/electrolysis. NREL/MP-560–35404. Golden: National Renewable Energy Laboratory. URL, www.nrel.gov/docs/fy04osti/35404.pdf; 2004.
[31] Simpson AG. Full-cycle assessment of alternative fuels for light-duty road vehicles in Australia, vol. 1. Sydney: World Energy Congress; 2004. pp. 1–10.
[32] Karlström M. Environmental technology assessment of introducing fuel cell city buses - a case study of fuel cell buses in Götenburg. 2002:10. Gothenburg: Chalmers University of Technology. URL, www.cpm.chalmers.se/CPMDatabase/DataReferences/ESA_2002–10.pdf; 2002.
[33] Dufour J, Serrano DP, Gálvez JL, Moreno García C. Life cycle assessment of processes for hydrogen production. Environmental feasibility and reduction of greenhouse gases emissions. Int J Hydrogen Energ 2009;34:1370–6.
[34] European Community. Externalities of energy: extension of accounting framework and policy applications. 2nd ed. Brussels: European Community; 2005 (Final Report on Work Package 6-New energy technologies).
[35] Food and Agriculture Organization of the United Nations. FAOSTAT: FAO statistical database. Rome: Food and Agriculture Organization of the United Nations. URL, faostat.fao.org; 2010.
[36] Bauen A, Berndes G, Junginger M, Londo M, Vuille F, Ball R, et al. Bioenergy – a sustainable and reliable energy source: a review of status and prospects (Main report). Paris: International Energy Agency. URL, http://www.ieabioenergy.com/LibItem.aspx?id=6479; 2009.
[37] International Energy Agency. World energy outlook 2008. Paris: International Energy Agency; 2008.
[38] British Petroleum. BP statisical review of world energy. London: British Petroleum. URL, www.bp.com/statisticalreview; 2010.