Informing energy consumption uncertainty: an analysis of energy data revisions

Rachel M Hoesly1 and Steven J Smith1,2
1 Joint Global Change Research Institute, Pacific Northwest National Lab, College Park, MD 20740, United States of America
2 Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD 20742, United States of America

Abstract
Quality energy consumption data are important for many types of analysis, and global data sets estimate trends of county level energy consumption, derived from country-reported data and regional reports. We present a novel basis for informing uncertainty in energy data by quantifying the changes in reported energy consumption as countries update their previously reported data. We use 17 editions of the British Petroleum World Energy Statistics (2001–2017) to evaluate how reported energy consumption is revised over time in aggregate coal, oil, and natural gas consumption data. We find that 70% of non-zero data points are adjusted by an average of 1.3% of a country’s total fossil fuel use in the year after their first publication. Earlier data points are revised less often, but almost half of historical trends contain some revisions in later years. The size and rate of data revisions vary over countries and fuels: coal data points have larger, less frequent revisions while oil data points have smaller more frequent revisions, with natural gas in between. A k-means cluster analysis was performed to group together countries with similar revision patterns. These groups span income, economy classification, OECD membership, and regions. Standard country groupings, therefore, do not predict the extent to which a country’s energy data has undergone revisions in the past.

1. Introduction
Quality energy data is an important input to various policy and scientific analysis including climate modeling, human/earth system modeling, air pollution modeling, energy analysis, economic modeling and analysis, and many others (Friedlingstein et al 2006, Amann et al 2013, Riahi et al 2017, Hoesly et al 2018). Several organizations provide high quality global energy estimates and while there has been some work to estimate uncertainties (Macknick 2011), there is little quantitative basis for estimating energy data uncertainty. Past work has, therefore, often used expert judgment for energy data uncertainty (Bond et al 2004, Smith et al 2011).

Widely used global energy data sets include the International Energy Agency (IEA) Energy Balances and Statistics (IEA 2015) and the United Nations (UN) Energy Statistics Yearbook/Database (UN 2016). Both are proprietary energy statistics (UN data is freely available after 1990), released every year, with energy data by country, sector and fuel. They include data on supply, trade, stocks, transformation and consumption and widely used for science and policy analysis. Private corporations also collect and publish energy statistics, such as the widely used British Petroleum (BP) Statistical Review of World Energy (BP 2015a), which is compiled annually and includes data on production, consumption, and trade of oil, gas and coal. Lastly, national governmental organizations, such as the US Energy Information Administration (EIA) (US Energy Information Administration 2018), track and publish global energy statistics.

While global estimates of primary energy use across these data sets are quite similar, aggregate totals often mask significant differences in estimates at sector and fuel level. Global primary energy use in 2007 estimated by EIA, IEA, and UN differ by less than 5%, and BP estimates are within 9% (Macknick 2011). However, comparisons at the sector level vary widely, which presents challenges, as sector and fuel level data are often the estimates used in models and analyses.
The definition of specific sectors and fuel vary across data sets; reported values are often not comparable, making uncertainty estimates and validation difficult. Estimates for some sectors and fuels, such as residential biomass, lack comprehensive data, even in developed economies with well-established energy statistics. Organizations rely on assumptions to fill data gaps, which can lead to a range of estimates across data sets. Lastly, reported energy estimates change over time. Historical energy use estimates may become better as methodologies evolve, assumptions are refined, and more data is collected. For example, energy statistics (as well as carbon dioxide emissions) in China have recently undergone extensive adjustments as Chinese statistical agencies collect and report more comprehensive data (Guan et al 2012, Liu et al 2015).

Often the only measure of uncertainty in energy data is a comparison across these different data sets. This range is caused by unique data collection, assumptions and methodology across these organizations. These organizations estimate country and global energy data through surveys and reporting forms from individual countries. While data is ultimately drawn from the same country-level data, each organization relies on unique survey questions or reports. IEA and UN use surveys distributed to member nations, BP and EIA rely on regional reports to derive energy data (Macknick 2011, BP 2018). Fundamental differences in assumptions and methodology also contribute to the range of energy consumption estimates. For example, the conversions from physical quantities to energy units requires a caloric or heating value, a value which varies across regions and sectors; organizations may estimate these values differently. While documentation and guidance are provided with these data sets, detailed methods and assumptions are often not readily available. While traditional uncertainty analysis might rely on sampling data multiple times, or systematically varying the assumptions for a given methodology, this is impossible for energy statistics and other socio-economic data, which makes traditional uncertainty analyses difficult.

We present a unique basis for informing energy data uncertainty estimates. This study examines multiple editions of the BP Statistical Review of World Energy (henceforth referred to as BP data) to examine how energy statistics change over time. We do not attempt to estimate the total uncertainty of energy data with this analysis. We aim to understand how reporting of energy data changes over time, thereby providing some insight into the known quantities of uncertainty in energy data. Revisions in energy data cannot replace comprehensive uncertainty estimates; however, they do reflect the underlying uncertainty and can be seen as one aspect of uncertainty in energy data.

2. Methods and data overview

The BP data report aggregate fuel consumption, production, and other market metrics for select countries and regions. We use 17 published BP reports (BP 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015b, 2016, 2017) to track energy consumption trends of coal, gas, and petroleum for 63 countries (available in the data as Mtoe). BP data includes trends from 1965 through the most recent year (i.e. the 2017 report provides estimates from 1965–2016). We use energy consumption estimates measured in energy units (Mtoe), as reported by BP, to have consistent units over all fuels as well as complete geographic coverage. Energy consumption is perhaps the most widely used energy statistic and is central for air pollutant and greenhouse gas emission estimates, making consumption a particularly relevant quantity for analysis. Consumption analysis in terms of energy units also facilitates comparison across fuels. BP data was selected for this analysis because these data are revised and released annually; and are publicly and freely available, making this analysis replicable by other researchers. Past versions of other energy data sets, such as EIA, IEA and UN are not publicly available.

Figure 1 shows the evolution of published trends of coal consumption in Turkey. Some newly published trends rewrite previous estimates; for example,
significant rewrites occur in 2012, and smaller adjustments occur in 2015 and 2017, shown in figure 1(a).

Figure 1(b) includes all 17 years of BP data and shows the revisions undergone by each data point. The black triangle indicates points that were in the future at the time of publication. The most recently published data points within a published trend are on the right side of the figure. The majority of data points undergo revisions in their second publication with progressively fewer and smaller revisions in subsequent publications. Older data points are revised more sporadically and occur as a shift in many consecutive data points within a trend, as shown in 2005, 2007, 2012, 2015, and 2017 in figure 1(b). Additional figures similar to figure 1(b) for select countries and fuels are shown in section 2 of the supplemental information (SI) is available online at stacks.iop.org/ERL/13/124023/mmmedia.

Energy estimates are organized for analysis along with identifying data including region, fuel, the value of the estimate, the year for which the estimate describes, and the edition (i.e. or year) in which the data point was published. Additionally, $n$, is defined as the number of years an estimate has been published. For example, $n = 1$ for Turkey coal consumption in 2007, published in 2008. While $n = 2$ for Turkey coal consumption in 2007 published in 2009. The first appearance of the data point occurs at $n = 1$, while the first possible revision can occur when $n = 2$. These metrics are described further in section 1 in the SI.

An energy estimate, defined by country, fuel, estimate value, estimate year, and edition are compared to the previously published version of that estimate. Changes in energy consumption estimates are evaluated as the difference between an estimate value and its previously published version. Unless otherwise noted, these differences are described as the absolute difference as a percent of a country’s total fossil fuel use, where a country’s total fossil fuel use is estimated as the sum of coal, oil, and gas consumption in the estimate year. This metric emphasizes changes that are most impactful to energy systems. Statistics measured as the percent of each fuel can result in accentuating large changes to small values. Equations showing how these values were calculated are shown in section 3.1 of the SI. Additional statistics showing values as a percent of the previously published value are shown throughout the SI.

3. Results

3.1. Historical versus recent changes in the data

As energy consumption trends evolve, all data points in a trend do not follow the same patterns; recent and historical data points within trends evolve with different behaviors. Figure 2(a) shows the occurrence of revisions as $n$ increases (through successive publications), where solid lines indicate all non-zero revisions and dotted lines indicate revisions larger than 0.5% of the country’s total fossil fuel use. Recently published data points are revised at a higher rate than historical data points. Revisions rates flatten around $n = 3–7$; however, the transition from a ‘recently published’ to a ‘historical’ data point is continuous and the definition of a ‘recent year revision’ versus a ‘historical revisions’ are somewhat arbitrary. This analysis will define a historical revision as any revision in a data point where $n \geq 5$ (i.e. is five or more years old at the time of publication of the trend). Results below do not vary substantially when that value is varied between three and six years.

3.2. Revisions in recent data points

70% of non-zero data points (45% of all data points) are revised in the first revision ($n = 2$), while only 26% and 13% of non-zero data points are revised in subsequent years ($n = 3, 4$). Figure 2(b) shows the cumulative distribution of the size of data revisions by $n$, for all data points (including those where size of revisions is 0). In the first revision ($n = 2$), data points change by an average of 1.3% of the total fossil fuel use for non-zero revisions, but many changes that occur are small; 54% of changes are less than 1%. Revisions when $n = 3$ and 4 follow similar distributions as $n = 2$, with progressively smaller and fewer adjustments. Non-zero changes average 1.1% and 0.64% of total fossil fuel use, with 87% and 95% of non-zero changes reflecting less than 0.1% of total fossil fuel use.

While changes in the first year are dominated by small adjustments, the largest 3% of changes reflect ±8% to 36% of total fossil use. Many large changes occur in the data of non-OECD member countries such as Algeria, Qatar, or Belarus; however, large revisions occur in high income, IEA member, developed economies like Sweden, where two, $n = 2$ revisions were +8.6% and −12% of total fossil fuel use.

For $n = 2–4$, the average of all revisions skews slightly toward revisions that represent a decrease in fuel use. Meaning that the average revised data point is smaller than originally estimated. However, the median adjustments represents an increase. While increases occur more frequently, decreases tend to be larger.

Revisions of oil, coal, and gas data points occur at different rates (figure 2(a)) and sizes. Oil data undergo larger revisions, more often than coal and gas data. While roughly 65% of non-zero coal and gas data points undergo adjustments in the first revision ($n = 2$), 79% of oil data points are adjusted. Median non-zero changes for coal, gas, and oil data points are 0.14%, 0.13%, and 0.19% of total fossil fuel use respectively. Through subsequent revisions, oil data points continue to be adjusted at a higher rate than coal and gas. Table S1-2, in the SI, shows similar statistics as a percent relative to the select fuel rather than as a percent of total fossil fuel. More complete summary
statistics on distribution of data adjustments by fuel and $n$ are available in the section 3 of the SI.

3.3. Revisions in historical data points
This section discusses adjustments in historical data points, which are defined as data points where $n \geq 5$. The term ‘historical trend’ here, refers to the sequence of historical points ($n \geq 5$) within any trend, regardless of edition. A trend is said to have a historical adjustment if any historical data point ($n \geq 5$) within the trend is revised. A historical revision is made up of at least one revision in a historical data point, although it is usually made up of revisions in many consecutive historical points. While revisions in recently published data points were evaluated as adjustments in single data point, historical revisions were evaluated as a summary of the revisions in a published trend.

44% of all published trends (including zero data points) contain historical data adjustments. While trends do not undergo historical revisions every year, all countries experience at least one (often more than one) historical adjustment for each reported fuel over the 17 revisions of BP data in this analysis. On average, 46% of data points in a published trend are adjusted, but sometimes only a single historical data point or all historical estimates in a trend are revised. Often, historical revisions include a few moderate adjustments (more than 1% of total fossil fuel use) and many smaller non-zero adjustments.

Average historical revisions are 0.8% of total fossil fuel use, measured as the average of all non-zero adjustments in a trend; however, the average maximum adjustment within each trend is 2.7%. Figure 2(c) shows the cumulative density function of the size of historical revisions, shown as the size of maximum revision in each trend adjustment as a percent of total fossil fuel. While 78% of historical data points do not change, 44% trends have some historical adjustment, and 20% of those trends have adjustments that are larger than 1% of total fossil fuel use.

The frequency and size of historical changes varies with fuel type with a similar pattern to revisions in recent data points. Coal data points undergo less frequent, smaller adjustments, while oil data points undergo more frequent larger adjustments with revisions in gas data points in between. 35%, 43%, and 56% of coal, gas, and oil trends undergo adjustments; over all trends the average of the average (max) revision within the trend are 0.7 (2.3)%, 0.9 (2.7)%, 0.7 (2.9)% for coal oil and gas. The frequency of data adjustments by fuel reflect the differences between the fuels themselves and how they are tracked and accounted for. For example, coal is shipped in discrete quantities from fewer coal mines to fewer countries; while oil has more trade in a larger number of products with different characteristics, more import and exports, and thus more complicated data collection and manipulation. Historical revisions may reflect changes in methodology or assumptions during data collection and processing such as a change in heat content, or changes in the data collected.

3.4. Data adjustments across countries
The characteristics of historical data adjustments vary widely across countries. Some countries, like Germany and Japan, show frequent, but only very small adjustments. Other countries have few small adjustments or sporadic larger adjustments; and in contrast, countries such as Malaysia, Australia and China show larger, and more frequent changes over more than one fuel.

A $k$-means cluster analysis was performed to group countries by the characteristics of their energy statistics revisions with a separate analysis for each fuel. This cluster analysis included the following variables: the number of editions with historical revisions, the number of historical data points with large revisions, and the 5th and 95th percentile of adjustments in each historical revision. If there are revisions in more than one historical trend (as in most country-fuel combinations have) the averages of the 5th and 95th percentiles of each historical revision are used.

Countries were clustered into three groups, shown in figure 3, which are most explained by the number of large historical revisions. Large revisions were defined
4. Discussion

The evolution of the reported of country energy consumption data varies widely by country, fuel, and historical nature of the estimated data point. While changes in historical and recent data points have different behaviors, the transition between the two categories of revisions (recent and historical) is continuous and the distinction between the two is somewhat arbitrary. In the second year that an estimate is published, 70% of non-zero data points are revised by an average of 1.3% of total fossil fuel use (median 0.51%), but those revisions taper off quickly thereafter. In 44% of trends, some historical data points (5 years or older at the time of publication) are revised. These revisions mostly consist of minor adjustments, but the average max adjustment within a
revision is 2.7% of a country’s total fossil fuel use. Data revisions for different fuels also vary; coal data are revised less frequently than gas or oil. Analyses using energy data and incorporating uncertainty estimates should take the likelihood that data points undergo revisions based on fuel type.

This analysis has focused on revisions as a percentage of a country’s total fossil fuel use. This metric is robust to changes in small values, avoids spuriously large values, and allows for comparison across both fuel types within countries and across countries with varying energy systems. However, statistics relative to that specific fuel, rather than to total fossil fuel use, may be more appropriate in a number of applications, for example, estimating uncertainty in emission inventories. Changes relative to a specific fuel will almost always be larger than those relative to total fossil fuel use, sometimes substantially so.

Table 1 shows both metrics for recent and historical revisions (all non-zero revisions). To illustrate the robust quality of metrics as a percent of total fossil fuel, table 1 also shows values where fuel type makes up at least 5% of total fossil fuel use. For coal estimates that are revised the first year after publication ($n = 2$), the 95th percentile change is 3.8% of a country’s total fossil fuel use, but the 95th percentile for changes from the previously published estimate is 43.6%. Excluding countries with small coal use, the 95th revision percentile is 4.11% of total fossil fuel and 21.8% relative to the previously published estimate.

Percentage changes are larger when examined relative to the estimates themselves. The difference is particularly large for coal, where the relative revision size is ~5 times higher (for the 5% threshold) when expressed as a percentage of the previously estimated coal consumption value rather than as the fraction of total fossil fuel consumption. These changes would result in corresponding changes in calculations using these values, such as emissions inventories.

This analysis has mostly evaluated data adjustments as an absolute value, however the distribution of positive and negative adjustments is not symmetrical. Small negative data adjustments are more common, so we routinely overestimate energy use. However, there are more very large, positive adjustments (more than 10% of total fuel use). Very large adjustments more substantially affect the makeup of a country’s energy use, which could have impacts on energy analysis or air quality modeling. Over 18 published editions of the same energy statistics, many data points have been revised. However, it does not seem that estimates of energy use are improving over time; the rate of change of data estimates within reports does not follow a trend.

### Table 1. Contrasting revisions statistics as a percent of total fossil fuel and as a percent of the previous estimate for the same fuel. Table shows the 50th and 95th percentile from the distributions of recent and historical revisions by fuel.

|                    | Coal                     | Natural gas              | Oil                      |
|--------------------|--------------------------|--------------------------|--------------------------|
|                    | % total fossil | % previous value | % total fossil | % previous value | % total fossil | % previous value |
| No threshold—includes all non-zero revisions | n | P50 | P95 | P50 | P95 | P50 | P95 | P50 | P95 | P50 | P95 |
| Recent revisions   | 2 | 0.431 | 3.8 | 3.43 | 43.6 | 0.514 | 5.68 | 1.69 | 13.4 | 0.569 | 4.85 | 1.37 | 12.5 |
|                    | 3 | 0.198 | 2.36 | 12.8 | 46.4 | 0.304 | 5.39 | 0.72 | 11.6 | 0.304 | 4.53 | 0.767 | 12.6 |
| 5% threshold—includes non-zero revisions where fuel type makes up >5% of total fossil fuel use | | | | | | | | | | | | |
| Recent revisions   | 2 | 0.346 | 11.1 | 26.3 | 21.8 | 0.528 | 5.90 | 1.67 | 12.2 | 0.569 | 4.85 | 1.37 | 12.5 |
|                    | 3 | 0.202 | 4.72 | 0.952 | 20.4 | 0.309 | 4.47 | 0.716 | 11.2 | 0.304 | 4.53 | 0.767 | 12.6 |
| Historical revisions | n | P50 | P95 | P50 | P95 | P50 | P95 | P50 | P95 | P50 | P95 |
|                    | | | | | | | | | | | | |
| Historical revisions | n | P50 | P95 | P50 | P95 | P50 | P95 | P50 | P95 | P50 | P95 |
| No threshold—includes all non-zero revisions | | | | | | | | | | | | |
| | 0.220 | 2.92 | 2.34 | 55.9 | 0.158 | 4.39 | 0.909 | 23.8 | 0.209 | 3.56 | 0.457 | 6.97 |
| 5% threshold—includes non-zero revisions where fuel type makes up >5% of total fossil fuel use | | | | | | | | | | | | |
| | 0.232 | 3.13 | 1.14 | 18.2 | 0.175 | 4.82 | 0.656 | 12.0 | 0.209 | 3.56 | 0.457 | 6.97 |

Note. Values for recent revisions shown are the 50th and 95th percentile of the absolute values of revisions. Values for historical revisions shown are the 50th and 95th percentile of the mean value of each historical rewrite. Distributions reflect non-zero revisions.
In the k-means analysis, countries were broadly, placed in groups with a small, medium, and large number of large revisions in their energy statistics. All three groups contain countries with varied energy systems and economies. Even though countries with different income, OECD status, or economy classification show unique distributions of revisions (shown in the SI, figure SI-32), these classifications do not predict how their energy statistics have been revised in past years.

Global energy statistics data are derived from many sources: from nationally mandated reporting of individual refineries, to trade reports, and nationally compiled statistics, all of which have their own biases. There is uncertainty associated with the survey data itself, as well as the methods and parameters used to compile and process the data into energy consumption trends. Errors can stem from data gaps, structural mistakes such as categorization and aggregation errors, incorrect conversion factors, and misreporting.

In this analysis, countries may have characteristically similar revisions for different reasons, which may imply different underlying uncertainty. A country with good quality energy statistics, that are revised often may appear similar in this analysis to a country with energy statistics that are never revised. The characteristics of these revisions are quantitatively similar and would have similar impact on analyses using this data. However, a data set with no changes over time does not necessarily equate to a more certain data set. It may only indicate a country that does not have the institutional capacity to reevaluate or improve energy data. Countries face many challenges with collecting and processing energy statistics, but frequent, small revisions may indicate a country with an established energy data management system using quality assurance strategies (Liu et al 2017). This analysis provides one quantitative basis to make such qualitative evaluations of the integrity of a country’s energy statistics.

This analysis cannot identify why revisions occur. In the absence of transparency and comprehensive documentation on survey and processing methods and methodology changes, examining data revisions provides what can be the only insight available into the magnitude and extent of energy data changes. Better documentation of the reasons for data changes would provide further insights and allow improved estimates of underlying uncertainty. While we recognize that data confidentiality issues can limit options, reporting the general nature of data changes would be helpful (e.g. net imports, production, consumption survey, or conversion factors). Improvements in energy statistics governance and best practices are essential for improving transparency and the quality of energy data. Liu et al (2017) suggest five key areas to develop robust energy data management systems: improving energy consumption data, strengthening coordination between agencies, prioritizing energy data, ensuring data quality, and enhancing access to data. It is also essential to also encourage transparency when strategizing for improving energy statistics.

This analysis quantifies the magnitude of revisions in reported energy data. This reflects how BP’s understanding of energy trends changes over time, which is likely indicative of other organization and national agencies. However, this is not an estimate of total uncertainty, but is the magnitude of corrections over time and is an indication of uncertainty as reflected in what are often multiple corrections to historical and recent time series. We see logically consistent behavior where recent data is corrected more often, however historical data is revised not infrequently. We suggest that the analysis here can be interpreted as a minimum indicator of uncertainty, and could be used as such in energy data analysis. However, this is just one aspect of energy data uncertainty. Energy statistics certainly still contain mistakes or less than ideal assumptions which organizations have not found or have not attempted to fix. At this point, expert judgment is still necessary to bridge between these quantitative estimates and a fuller estimate of uncertainty, at least until more systematic analysis of energy data systems is conducted as suggested below.

Estimating uncertainty in energy data is fundamentally different than uncertainty in other kinds of scientific data. Rather than smooth distributions obtained by repeated sampling, energy statistics are derived from few sources, vary widely with quality. We have presented a novel way of analyzing energy data by examining the observed revisions in historically published energy statistics. This analysis does not attempt to quantify total uncertainty of energy estimates, nor explain why these changes occur, which would involve decomposing methods for estimating energy use, investigating survey methods, and ultimately comparing independent estimates. The scientific community would certainly benefit from such a study. In the absence of such comprehensive studies on uncertainty in energy data, or any guidance on how to address uncertainty in energy consumption data, this analysis provides a basis to begin quantifying uncertainty in these data.

Acknowledgments

This research was supported by the Office of Science of the US Department of Energy (DOE) as part of the Earth System Modeling Program. The Pacific Northwest National Laboratory is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830.

ORCID iDs

Rachel M Hoesly @ https://orcid.org/0000-0003-0965-2729
References

Amann M, Klimont Z. and Wagner F. 2013 Regional and global emissions of air pollutants: recent trends and future scenarios. Annu. Rev. Environ. Resour. 38 31–55

Bond T C, Streets D G, Yarber K F, Nelson S M, Woo J-H and Klimont Z. 2004 A technology-based global inventory of black and organic carbon emissions from combustion. J. Geophys. Res. 109 D14203

BP 2001 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2002 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2003 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2004 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2005 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2006 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2007 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2008 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2009 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2010 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2011 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2012 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2013 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2014 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2015a BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2015b BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2016 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2017 BP Statistical Review of World Energy (London: BP) (bp.com/statisticalreview)

BP 2018 Definitions and Explanatory Notes BP plc (https://bp.com/en/gLOBAL/corporate/energy-economics/statistical-review-of-world-energy/using-the-review/definitions-and-explanatory-notes.html)

Friedlingstein P et al 2006 Climate–carbon cycle feedback analysis: results from the C4MIP model intercomparison J. Clim. 19 3337–53

Guau D B, Liu Z, Geng Y, Lindner S and Hubacek K 2012 The gigatonne gap in China’s carbon dioxide inventories Nat. Clim. Change 2 672–5

Hoehly R M et al 2018 Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the community emissions data system (CEDS) Geosci. Model Dev. 11 369–408

IEA 2015 IEA Energy Statistics and Balances (http://iea.org/statistics/)

Liu B, Evans M, Yu S, Roshchanka V, Dukkipati S and Sreenivas A 2017 Effective energy data management for low-carbon growth planning: an analytical framework for assessment Energy Policy 107 32–42

Liu Z et al 2015 Reduced carbon emission estimates from fossil fuel combustion and cement production in China Nature 524 335–4

Macknick J 2011 Energy and CO₂ emissions data uncertainties Carbon Manage. 2 189–205

Riahi K et al 2017 The shared socioeconomic pathways and their energy, land use, and greenhouse gas emissions implications: an overview Glob. Environ. Change 42 153–68

Smith S J, van Aardenne J, Klimont Z, Andres R J, Volke A and Delgado Arias S 2011 Anthropogenic sulfur dioxide emissions 1850–2005 Atmos. Chem. Phys. 11 1101–16

UN 2016 UN Energy Statistics Database (UN) (http://unstats.un.org/unsd/energy/)

US Energy Information Administration 2018 International Energy Statistics (https://tia.gov/beta/international/data/browser/#/?pa=000000001&sc=41000000020000000000000000000000&dct=2&c=INTL.44-2-AFRC-QBTU.A&vs=0&v=H&rend=2015)