This is the accepted pre-print version of a paper submitted for the ECEEE 2019 Sumer Study. The Proceedings version of this paper can be found here.

4-114-19 McKenna

Analysis of international residential solar PV self-consumption

Eoghan McKenna
UCL Energy Institute. The Bartlett School of Environment, Energy and Resources. University College London (UCL). Central House, 14 Upper Woburn Place, London WC1H 0NN, UK
Email: e.mckenna@ucl.ac.uk

Ellen Webborn
UCL Energy Institute. The Bartlett School of Environment, Energy and Resources. University College London (UCL). Central House, 14 Upper Woburn Place, London WC1H 0NN, UK
Email: e.webborn@ucl.ac.uk

Philip Leicester
Centre for Renewable Energy Systems Technology, The Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Wolfson Building, Loughborough University, LE11 3TU, UK
Email: p.a.leicester@lboro.ac.uk

Simon Elam
UCL Energy Institute. The Bartlett School of Environment, Energy and Resources. University College London (UCL). Central House, 14 Upper Woburn Place, London WC1H 0NN, UK
Email: s.elam@ucl.ac.uk

Abstract

How much does installing solar photovoltaic ('PV') panels reduce domestic electricity bills? Despite over 800,000 households with solar panels in the UK today, we do not have a firm answer to this simple question. This is because the empirical data necessary to calculate solar self-consumption is not widely available; an issue common to many countries. Without electricity metering that records solar PV generation, imports and exports, it is impossible to know exactly how much solar power is being used directly in the household, and as a result, how much solar panels actually reduce electricity bills. As many countries reduce state subsidies, the revenue stream from avoided grid imports (i.e. self-consumption) becomes increasingly critical for the economic viability of PV. Quantifying this potential revenue is also vital for the solar industry to evaluate the potential benefits of battery storage and flexibility services. This paper will help to address this knowledge gap by analysing a previously unused dataset of electricity readings from over 1,300 households with solar panels located across the UK, USA, Australia, Germany, the Netherlands, and Belgium. The results show there is considerable variation in self-consumption between different households, but this is relatively well explained by differences in the amount of solar electricity generated and the amount of electricity consumed during the day (r-squared 0.915), and that the factors affecting self-consumption are similar between the different countries analysed here. The findings will be relevant to regulators for making evidence-based decisions about solar energy policy, provide better information about potential self-consumption to people interested in adopting solar panels, and highlight the importance of improving data availability for 'behind-the-meter' micro-generation.

Introduction

We are entering the era of subsidy-free solar photovoltaics ('PV'). Solar feed-in tariffs, which guarantee payment for solar electricity generated or exported to the grid, have been used internationally by regulators to encourage adoption of PV. With the rapid global deployment of PV, however, the cost of PV panels has fallen, and feed-in tariffs have been progressively reduced and, in places like Australia and the UK, even removed altogether. Less than ten years ago, households with PV received a premium rate for their solar electricity generation. Now, new adopters face having to negotiate a rate on the open market or give exported generation away for free.

The financial viability of adopting PV increasingly depends on the resulting reduction in electricity required from the grid due to 'self-consumption' of PV generation. Indeed, in the absence of any feed-in tariff or market based solutions for energy trading, self-consumption provides the only financial incentive for adopting PV.

This is the accepted pre-print version of a paper submitted for the ECEEE 2019 Sumer Study. The Proceedings version of this paper can be found here.

Error! Reference source not found. shows how feed-in tariff rates have consistently fallen for new PV installations in the UK, driving up the level of self-consumption[1] required to meet a 10% return on investment.

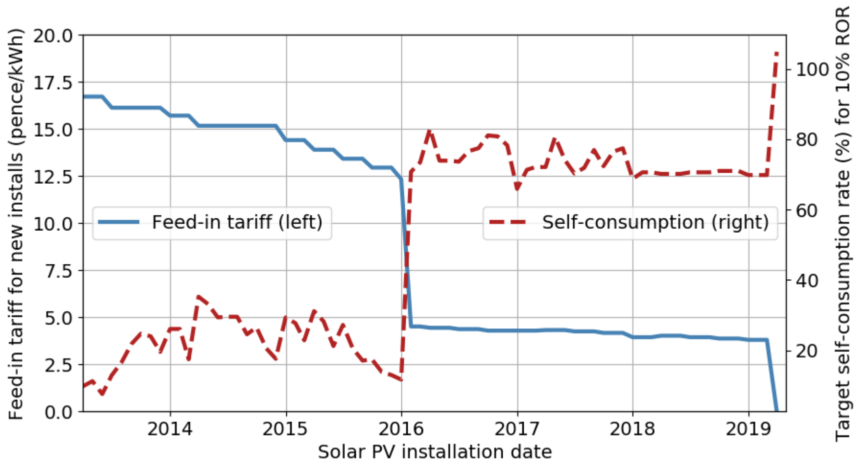

Figure 1: Timeline showing the progressive reduction of the feed-in tariff (generation) rate in the UK and the corresponding target self-consumption rate to achieve a rate of return of 10% of the installation cost. Based on a typical UK PV system of 2.9kW with 10.2% load factor. Feed-in tariff rates from (Ofgem, 2018), installation costs from (BEIS, 2018a), and electricity prices[2] from (BEIS, 2018b). We assume here no payments for exports once feed-in tariffs are removed, though we note that the Government is consulting on this whether an export price of some sort should be guaranteed.

While understanding self-consumption becomes increasingly important for the continued adoption of PV, there is a lack of empirical data available about it. Although smart meters measuring electricity imports and exports are being widely rolled out, they do not measure any on-site generation such as PV that is connected 'behind the meter'. If households with PV are not 'fully metered' (i.e. metering of imports, exports and generation) then it is impossible to calculate their gross electricity demand and self-consumption. Even where generation metering does exist, such as in the UK, these tend to be non-smart meters and are not integrated into the national smart metering communications infrastructure. As a result, even when smart meters are fully rolled out, we will still not know how much electricity households with PV[3] actually consume (their gross demand) or their self-consumption, unless generation data is also collected.

---

[1] We define 'self-consumption (%)' to be the percentage of generation that is consumed on-site (i.e. not exported to the grid).

[2] 1 pence is worth approximately 1.11 Euro cents.

[3] This problem is not limited to PV but applies to any 'behind the meter' micro-generation e.g. micro-CHP, home battery systems.

This is the accepted pre-print version of a paper submitted for the ECEEE 2019 Sumer Study. The Proceedings version of this paper can be found here.

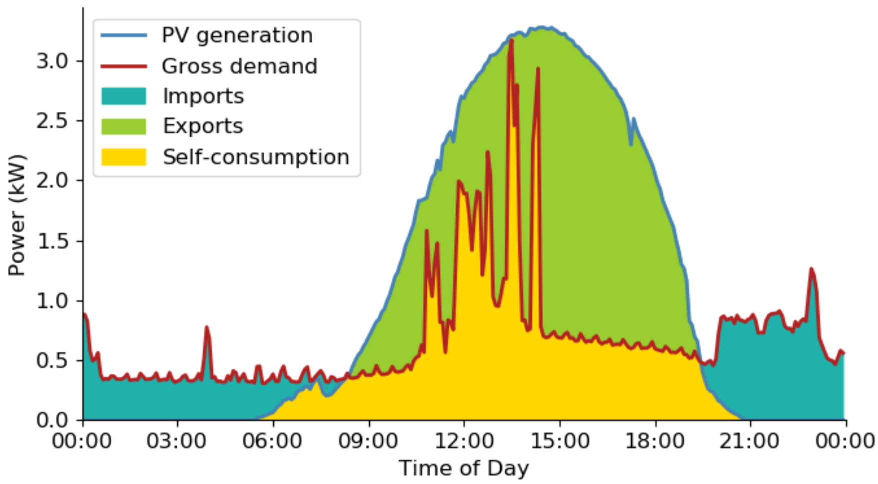

Figure 2: Electrical power flows for a summer day for an example residential household with PV in the UK.

Self-consumption is a subject of international policy importance. The revised EU Renewable Energy Directive makes it a legal requirement that all member states allow 'self-consumers to generate, store, consume, and sell electricity without facing disproportionate burdens' (European Parliament, 2018). Governments and regulators need accurate data on self-consumption to calculate fair regulations, tariffs and subsidies. Currently, this is being done in the absence of reliable data on self-consumption, and this reduces the effectiveness of policies.

In the UK, solar power is the most popular electricity generating technology (BEIS, 2017). While households adopting PV are still a minority in the UK, the majority of households would install PV if they could (Vaughan, 2018). Prospective adopters of PV need reliable, impartial advice to make informed decisions. Moreover, reliable estimates of self-consumption are needed to underpin the development of markets for a number of innovative technologies that are being driven, primarily or in part, by the opportunity of making use of cheap solar power: electric vehicles, home battery systems, smart home systems, and peer to peer energy trading. Current analyses of these technologies, often heavily reliant on modelling, are limited by the fundamental lack of empirical data on a critical resource for their development and adoption: solar self-consumption.

What is needed are monitoring campaigns of large, representative samples of households with PV, combined with good contextual data (socio-demographics, building characteristics, presence and use of appliances, etc.). Unfortunately, these don't exist, in the UK or elsewhere. Furthermore, although access to large-scale smart meter and linked contextual data resources are starting to become available (Webborn, Elam and McKenna, 2019), while distributed generation continues to be un-metered and un-integrated with national smart metering communications infrastructures, the goal of reliable data on households with PV will remain expensive and challenging to achieve.

To help address this evidence gap and improve understanding of self-consumption, this paper presents an analysis of empirical data from over 1,300 'fully metered' households with PV located across the UK, USA, Australia, Germany, the Netherlands, and Belgium. We show how self-consumption varies across households and countries and we show results for a simple linear regression model which can be used to estimate self-consumption for households with PV with a reasonable level of accuracy given a few simple inputs.

Capturing self-consumption

Self-consumption is the amount of energy generated by a dwelling's solar PV and consumed at that dwelling (see Figure 2 for reference). Accurate self-consumption calculations require real-time monitoring of generation and export to the grid. Accurate gross demand calculations require, in addition, monitoring of imports from the grid.

Estimation

Unfortunately, the vast majority of residential PV systems are not configured *by regulation* to allow the calculation of self-consumption, or indeed, gross demand[4]. Half-hourly monitoring (for example by a smart meter) of imports and exports from/to the grid is insufficient to determine self-consumption and gross demand.

---

[4] A description of PV metering configurations is presented in (McKenna and Thomson, 2013).

Early PV systems in Germany were connected directly to the grid and not 'behind the meter' with the result that these households cannot meter self-consumption, as everything that is generated is exported to the grid. The only option in the absence of full monitoring is to model self-consumption, typically involving superimposing simulated PV generation and simulated gross demand, such as in (Leicester, Goodier and Rowley, 2016).

Calculation

The alternative to modelling and simulation is to install additional monitoring equipment in a (typically small) number of households with PV, such as in (McKenna, Pless and Darby, 2018). The data used in this paper comes from more than 1,300 fully-monitored dwellings in six countries.

Common calculation challenges

A common problem for calculating self-consumption arises when net demand (import minus export) readings are taken rather than separate import and export readings. For example, a meter may be configured to store a single net demand reading for each half-hour, which could be positive (a net import in that half-hour) or negative (a net export in that half-hour). However, this may mask the fact that the household both imported and exported during periods within the half-hour. 'Netting off' like this and using the result to calculate self-consumption at half-hourly intervals will underestimate actual imports and exports and therefore overestimate self-consumption. The longer the netting off period, the greater the overestimation. For example, netting over 30 minutes can underestimate imports and exports by ~15% compared to 1-minute data (McKenna, Pless and Darby, 2018). This illustrates the importance of using high-resolution data if using net demand data, or for low-resolution data ensuring use of actual exports and imports and rather than values derived from net demand.

In the UK smart meters are specified so that they can record both exports and imports for a given half-hour (BEIS, 2014). As a result, they are appropriately configured to be able to accurately calculate self-consumption (provided generation data is also available). In contrast, in the USA, 'net metering' is the standard. In simple terms, net metering is the equivalent of the meter 'going backwards' when exports occur. If a household with PV and net metering exports 1 kWh during the day, then they can import 1 kWh later and the meter will go back to where it was before the exports occurred. This means that there is no financial incentive for households on net metering to use electricity at the exact time when they generate it. We would expect households with PV and net metering to have different demand profiles to those with PV and a feed-in tariff as the latter only benefits from self-consumption when consumption occurs at the same time as generation.

The next challenge is the requirement for longitudinal data from many households. Longitudinal data is required because self-consumption is a function of generation, demand, and their time coincidence. All of these vary seasonally and so it is important to have data that covers a sufficient breadth of time to capture this variation. Data from many households are required to capture the variability of residential demand both in terms of overall consumption and time-of-use, as well as to capture the diversity of PV systems (smaller systems will tend to have higher levels of self-consumption while larger systems will tend to export more). We note that households with PV have demand profiles and consumption practices that are significantly different from households without PV because of the financial incentive and desire to maximise self-consumption (McKenna and Thomson, 2014; McKenna, Pless and Darby, 2018). This is another reason why it is important to use data from actual households with PV, as we do not know how to correct either simulated data or data from households without PV to account for this apparent behavioural difference.

Contextual data are important, in particular metering configuration and tariff details and these are generally determined by where and when the PV system was installed. Metering configurations can be different across countries or states (McKenna and Thomson, 2013) and this can affect self-consumption, as described above in relation to net metering. Feed-in tariffs incentivise different kinds of consumption behaviour. Early adopters of PV received large premiums for solar electricity compared to the average retail price of electricity, and were therefore incentivised to export as much as possible. Feed-in (export) tariff reduction below the import price incentivises maximising self-consumption (and minimising exports), as demonstrated in Figure 1. In New South Wales, Australia, this transition happened overnight, with households going from a subsidised feed-in tariff payment of 60c/kWh to a voluntary market-led rate of 4.7 to 6.1c/kWh[5] on the 31 December 2016 (NSW Government, no date). We would expect this change in financial incentive to impact consumption behaviour.

In the UK, the feed-in tariff premium is paid for generated electricity, with a smaller payment received for anything exported (in the absence of smart meters, exports are 'deemed' to be 50% of generation). However, although there has always been a financial incentive to maximise self-consumption, for early adopters the

[5] 1 AUD is about 0.63 EURO

financial benefit of self-consumption made up a relatively small proportion of the overall return on investment for a PV system (compared to the payments received for generation), whereas for present adopters self-consumption is the major, or only component (see Figure 1).

Contextual data about each household, such as socio-demographic and building information, as well as details of how the households were recruited, and sampling method used are also important for assessing self-consumption and understanding consumer behaviour(Webborn, Elam and McKenna, 2019). Without this it is impossible to generalise the results of a study and this limits the insights that can be drawn from the data analysis. Unfortunately, this type of contextual data is generally missing from the few empirical datasets available.

Methods

Data

The data we analyse here consist of gross electricity demand and generation readings at 5-minute resolution. These are from sensors that have been self-installed by owners of PV systems. The data were sourced from PVOutput.org, a website which offers a free service for comparing and sharing PV generation and demand data among interested users. Generation and demand data were available for a total of 1,315 unique dwellings in Australia, Netherlands, UK, USA, Germany and Belgium. Figure 3 shows the availability of data over time for each country.

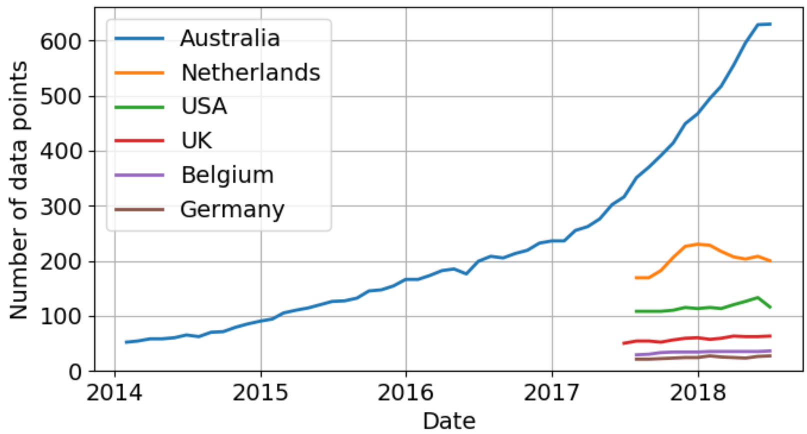

Figure 3: Timeline showing number of concurrent monthly data points by country. The legend entries are in same order as the countries appear on the figure.

In this paper we calculate monthly totals of generation, gross demand, and self-consumption for each dwelling. Annual totals are calculated in (McKenna, Pless and Darby, 2018). We use monthly values here because it allows us to explicitly capture the seasonal variability in the data and include data from systems where a full year of data is unavailable.

The 5-minute generation and gross demand data are used to calculate derived variables of 5-minute imports, exports and self-consumption. The resulting files are down-sampled to 30-minute resolution to speed computation. Missing data is often an issue with time series energy data and this data is no different. We adopt a similar approach to (McKenna, Pless and Darby, 2018) and use 'thresholds' to determine what data to discard, starting by discarding all half-hours with any missing data. Each day's data for each dwelling is discarded if less than 75% is available. Daily totals are then calculated (or estimated by scaling up totals from the available half-hours). The remaining are used to compute daily totals which are scaled in proportion to their availability so that each PV system-day represents a complete estimate. The same is done for each system-month using the daily data and a 60% threshold. A sensitivity analysis varying these thresholds revealed that increasing the threshold had a negligible impact on estimates but considerably reduced the sample size. Any set of readings for a dwelling for a given month with any monthly total greater than 5000 kWh was excluded (53 sets of readings).

Unfortunately, we do not know what sensors were used to measure generation and gross demand. They will have been purchased from a range by the PV owners. The cheapest devices are clip-on current clamps which measure current flow through a cable. Some of these devices estimate voltage and this can introduce errors to power and energy readings. The resolution of these devices is usually quite high (10 seconds), so the 'net demand' issue with calculating gross demand and self-consumption is unlikely to be an issue.

This is the accepted pre-print version of a paper submitted for the ECEEE 2019 Sumer Study. The Proceedings version of this paper can be found here.

One disadvantage of this dataset is the absence of contextual data. We assume that the systems being monitored are residential households with PV, but some may be commercial consumers. Systems are discarded with particularly large or small annual generation (> 10,000kWh or <1,000kWh) or gross demand (>20,000 kWh or <1,000 kWh). We do not know if some systems have technologies installed to increase self-consumption e.g. home battery systems. Systems are discarded with particularly high demand during the day (>60% of total demand). We observe electricity demand only, and do not know if the systems have other fuel sources e.g. gas. We do not know when the systems were installed, only their country. We cannot be sure of their metering configuration or tariff although systems in the USA are likely to be on net metering and the other systems will be on feed-in tariffs. Finally, we assume the demand measurements are for a whole dwelling and not part of a dwelling or a single load.

Model

We know the data represents people who are self-motivated to monitor their energy generation and consumption. As a result, the data will not be representative of typical households with PV. We try to account for some of this bias by conducting a simple regression analysis on the data using the same method as described in (McKenna, Pless and Darby, 2018), but using monthly data rather than annual data. The regression equation includes a dependent variable for fraction of gross electricity demand consumed during the day-time (taken to be between 10am and 4pm). This variable will partially capture the difference between 'highly engaged' households who consume more during the day and those that are more like typical households. The results of the regression are described in the next section.

Results

Data features

Figure 4 shows the distributions of key variables for the combined data from all countries i.e. data from Australia is shown alongside data from Netherlands, etc., even though they have different climates. Table 1 presents the key characteristics of the distributions including the number and percentage of data points that are not shown in the plots. Naturally the mean and median daytime demand are significantly lower than total demand, and daytime demand is approximately 29% of total demand on average. On average self-consumption is approximately 49% of total generation, and meets approximately 37% of total demand.

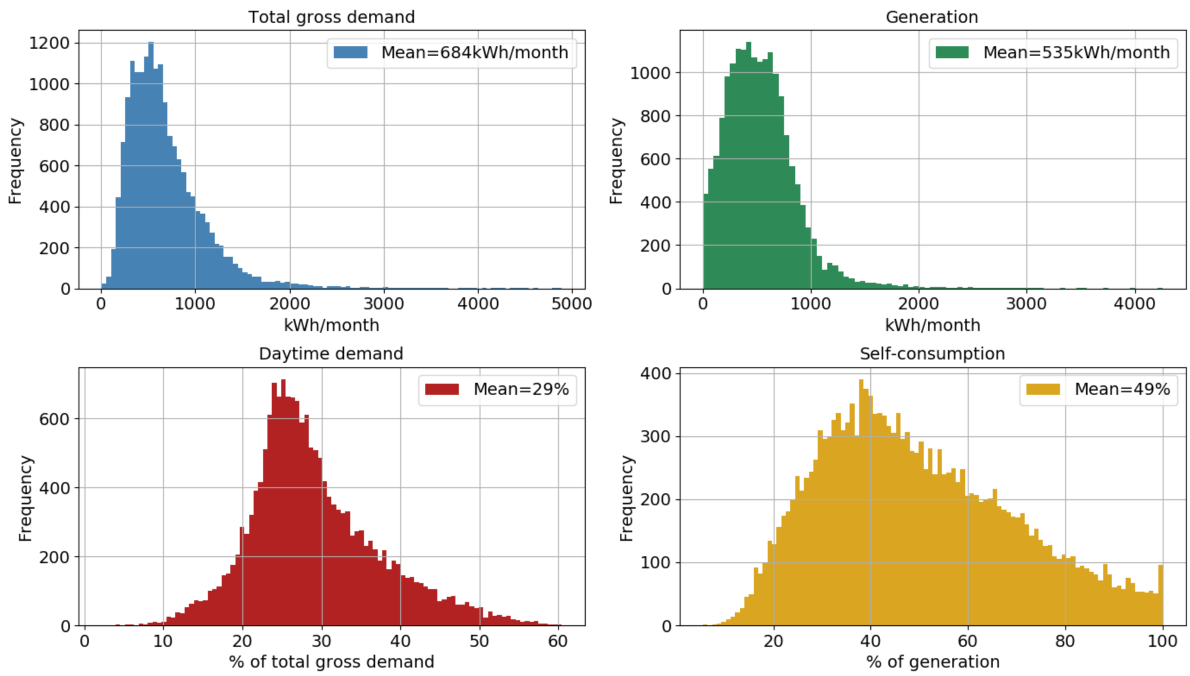

Figure 4: Histograms of total demand, PV generation, daytime demand fraction and self-consumption (as a percentage of PV generation) for each household each month. The legend indicates the mean values of each distribution.

Table 1: Basic statistics for each variable in the data set.

This is the accepted pre-print version of a paper submitted for the ECEEE 2019 Sumer Study. The Proceedings version of this paper can be found here.

	Total demand	Daytime demand	Generation	Self-consumption
# > 2000 kWh, (%)	252 (1.44%)	1 (0.00%)	81 (0.46%)	4 (0.02%)
Maximum	4887.16	2021.91	4257.28	2417.70
Mean	684.90	200.81	535.39	246.99
Median	594.92	165.19	498.61	209.84
Standard deviation	424.84	138.92	336.63	172.89

Table 2 shows the mean, median and standard deviation of each of the variables used in the linear regression above as well as the number of data points for each country. Australian data accounts for approximately 70% of the dataset, whereas Belgium and the Netherlands only have data for a few hundred months and households. The US has the highest levels of generation, daytime demand and self-consumption. This is unsurprising due to the high levels of irradiance and solar PV penetration, driving up both demand from air-conditioning and solar generation. The Netherlands and Great Britain have the lowest mean and median generation, the Netherlands and Germany have the lowest daytime demand, and the Netherlands has significantly lower self-consumption than any of the other countries. Of the four European countries, Belgium has the highest levels of generation, daytime demand and self-consumption.

Table 2: Mean, median and standard deviation of each key variable for each country to the nearest integer. Dark blue shading indicates the country with the highest value in each column, light yellow indicates the lowest.

		Generation			Daytime demand			Self-consumption		
	Number of observations	Mean	Median	Standard deviation	Mean	Median	Standard deviation	Mean	Median	Standard deviation
AU	12255	568	549	276	215	184	137	269	237	158
BE	405	473	393	361	183	154	106	207	163	156
DE	287	433	289	471	130	114	93	158	125	147
NL	2446	277	224	236	112	98	72	113	94	86
GB	751	298	252	218	160	138	87	169	146	122
US	1385	870	786	528	278	228	192	362	301	262
All	**17529**	**535**	**499**	**337**	**201**	**165**	**139**	**247**	**210**	**173**

Figure 5 illustrates the main variables that influence self-consumption. Self-consumption (as a percentage of generation) is plotted against monthly generation over monthly demand ('cover factor'). The colour illustrates the proportion of demand that occurs during the day, averaged over the month corresponding to the data point. The broad trend is that when the cover factor is low (generation is much lower than demand), most generation is self-consumed (high self-consumption %). As the cover factor increases (i.e. as generation overtakes demand), self-consumption % decreases asymptotically. For a given cover factor, there is a wide variation of self-consumption which is partly explained by variation in daytime demand proportion. For example, a cover factor of 1 (indicating the same amount has been generated as consumed in the month) results in self-consumption levels of around 20% when most of the demand is not during the day (low daytime demand proportion (~10%)). This rises to around 80% self-consumption for high daytime demand proportions (~60% demand during the daytime). The variability of observed self-consumption decreases as the cover factor increases, particularly above 1.

This is the accepted pre-print version of a paper submitted for the ECEEE 2019 Sumer Study. The Proceedings version of this paper can be found here.

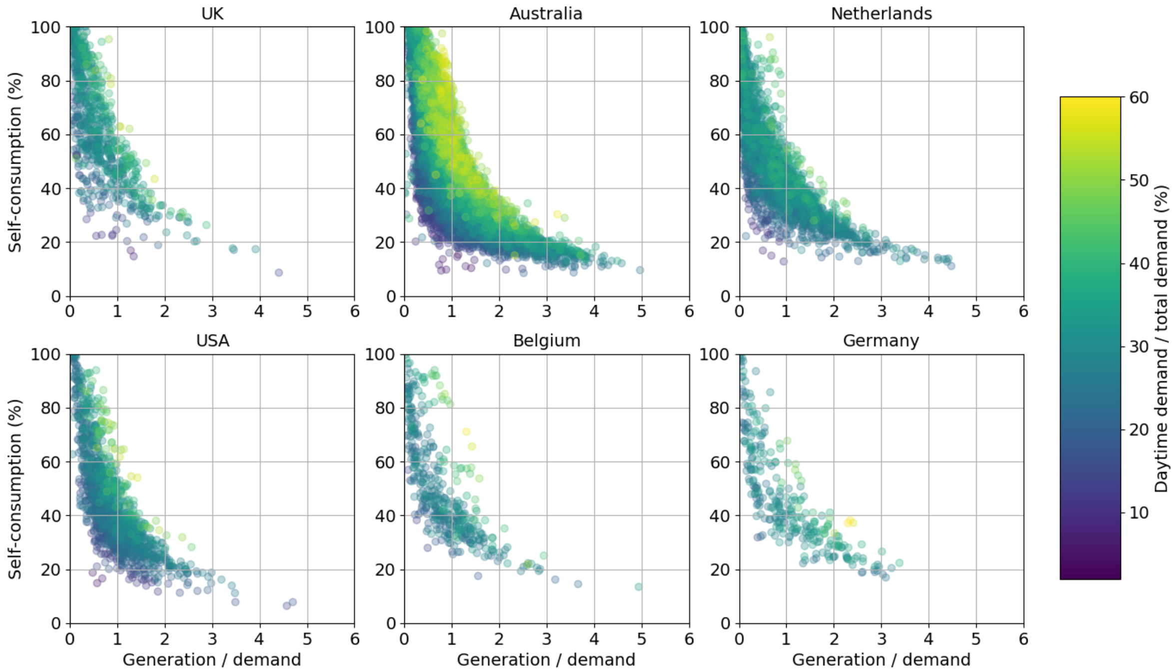

Figure 5: Self-consumption (as a percentage of generation) against monthly generation divided by monthly demand for each country. Each circle represents one month's totals for one PV system. Circle colour indicates the proportion of demand consumed during the daytime (10:00 – 16:00).

While not identical, the overall trend is similar for the different countries, particularly for cover factors around 1 and above. This indicates that although we are clearly not capturing all the explanatory variables, there is a fair agreement in overall trends between countries, and it is reasonable to apply a common approach to the analysis of self-consumption internationally.

Regression

We are interested in predicting how much solar generation will be self-consumed, and so our next step is to perform linear regression against the key variables above. Table 3 shows the results of liner regression where self-consumption is the dependent variable and the independent variable(s) are shown in each column. The p-value is very low for all coefficients due to the large number of data points (17,582 monthly readings). Regressing against generation and daytime demand gives the highest R-squared value (0.915) and the lowest residual standard error (50.42). There is a significant benefit from using daytime demand rather than total demand, even with our basic assumption that 'daytime' is 10:00 – 16:00. Clearly a better fit could be achieved using a varying time window that reflected variation in actual sunlight hours.

Table 3: Regression coefficients and results. Self-consumption is the dependent variable against the variable(s) in each column. Note that * indicates p-value < 0.001. Standard errors shown in brackets.**

	Total demand	**Daytime demand**	**Generation**	**Generation & total demand**	**Generation & daytime demand**
Total Demand	0.304 (0.002)***	-	-	0.201 (0.002)***	-
Daytime demand	-	1.089 (0.005)***	-		0.806 (0.003)***
Generation	-	-	0.399 (0.002)***	0.282 (0.002)***	0.230 (0.001)***
Intercept	38.694 (1.646)***	28.335 (1.112)***	33.376 (1.545)***	-41.585 (1.089)***	-38.188 (0.769)***
R-Squared	0.556	0.765	0.604	0.796	0.915
Residual Standard Error	114.90	83.74	108.90	78.20	50.42

We take the linear regression coefficients from the final column of Table 3 to model the self-consumption for each available data point in the dataset. That is to say, for each household, for each month, we predict that

$$E_{\text{self}} = 0.806\ E_{\text{dayDem}} + 0.230\ E_{\text{gen}} - 38.188$$

where E_{self} is predicted self-consumption, E_{dayDem} is daytime demand, and E_{gen} is PV generation. Figure 6: Predicted self-consumption against actual self-consumption, derived from the linear regression using generation and daytime demand. Solid blue line indicates perfect prediction, dashed orange line is the line of best fit of these points. shows predicted E_{self} against actual self-consumption for all data points. The solid blue line represents perfect prediction; points above this line have had their self-consumption over-predicted, and those beneath the line have been under-predicted by the linear model. The dashed orange line is a linear regression for predictions against actual self-consumption. The 95% confidence interval of this line was so tight to the line that its indication made the plot less easy to understand. The regression indicated by the dashed line shows that as self-consumption increases the model is more likely to under-predict self-consumption. Therefore the relationship between generation and/or daytime demand and self-consumption is likely to be convex rather than linear. However, the dashed and solid lines are fairly close and so our choice of linear model is shown to be a fairly good fit, particularly for the vast majority of points which are less than 1000 kWh self-consumption/month.

Table 2 showed that there are some significant differences between the different countries' distributions of generation, daytime demand and self-consumption. Table 4 shows the results of performing linear regression of self-consumption against generation and daytime demand for each country, and the results for all countries (same as above) in the final column. Highest and lowest values out of the 6 countries are indicated with blue and yellow shading, respectively, to show the range of values. For example, generation plays the greatest role in Great Britain (coefficient 0.390) and the smallest role in Germany (0.183). Daytime demand plays the greatest role in Australia (0.836) compared to only 0.419 in the Netherlands, where daytime demand is the least significant relative to generation (generation over daytime demand coefficients is 0.672). This table highlights the variety of trends across the range of countries studied.

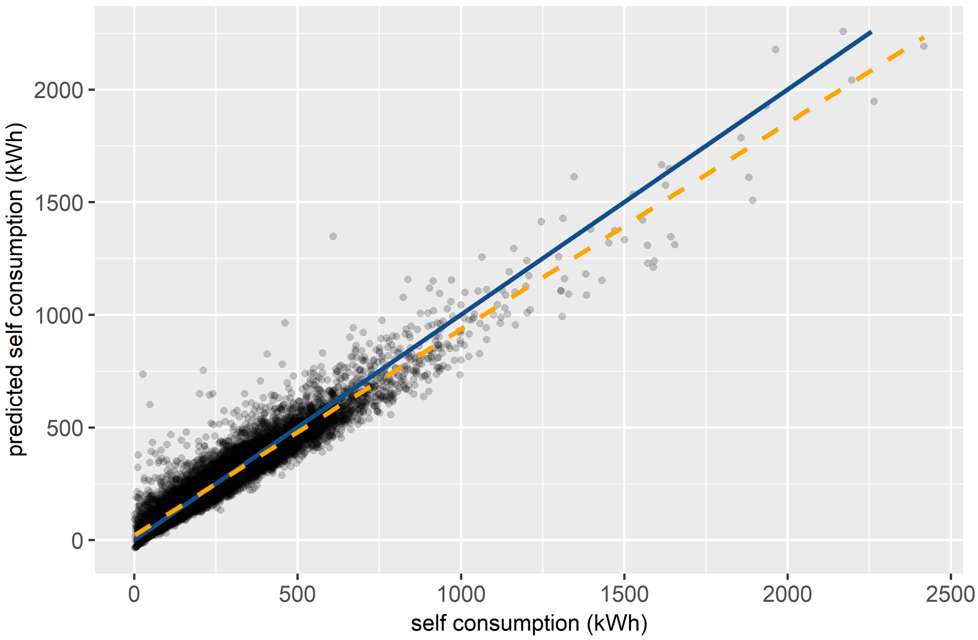

Figure 6: Predicted self-consumption against actual self-consumption, derived from the linear regression using generation and daytime demand. Solid blue line indicates perfect prediction, dashed orange line is the line of best fit of these points.

Table 4: Regression coefficients by country. The highest (lowest) values out of all countries are shown in bold (underline). The 5th row (Generation/daytime demand) is the calculated ratio of the coefficients.

	AU	BE	DE	NL	GB	US	All
Number of observations	12225	405	287	2446	751	1385	**17529**

This is the accepted pre-print version of a paper submitted for the ECEEE 2019 Sumer Study. The Proceedings version of this paper can be found here.

PV Generation	0.221 (0.001)***	0.276 (0.009)***	0.183 (0.006)***	0.282 (0.003)***	0.390 (0.007)***	0.238 (0.004)***	**0.230 (0.001)*** **
Daytime demand	0.836 (0.003)***	0.679 (0.031)***	0.753 (0.032)***	0.419 (0.010)***	0.626 (0.019)***	0.835 (0.012)***	**0.805 (0.003)*** **
Intercept	-36.357 (1.012)***	-47.686 (6.298)***	-19.438 (3.852)***	-11.792 (1.404)***	-47.501 (3.521)***	-76.857 (4.032)***	**-38.187 (0.769)*** **
Generation / daytime demand	0.265	0.406	0.243	0.672	0.622	0.285	**0.286**
R-Squared	0.914	0.852	0.934	0.842	0.877	0.922	**0.915**
Residual Standard Error	46.62	60.12	37.89	34.28	42.74	73.36	**50.42**

Checking the results

The coefficients above can be used to predict self-consumption for other households with PV. How accurate is the prediction? We check the results by applying the UK coefficients to an independent dataset of households with PV in the UK that were monitored as part of the Customer Led Network Revolution (CLNR). The data are described and analysed in (McKenna, Pless and Darby, 2018). We take the monthly generation and daytime demand totals for the CLNR households and predict their self-consumption using the UK coefficients above. Figure 7 shows the results. We find that many self-consumption data points are predicted to be more than the total generation that month, or predicted to be negative. If we require predictions to between zero and total generation, we get the result shown in Figure 8, which increases the r-squared value from 0.567 to 0.864.

Considering these are two completely independent groups of households with data from different time periods, the little we know about the households, and the data quality issues, the results are quite reasonable overall. It shows that it is possible to produce reasonable estimates of self-consumption given the coefficients produced by this type of analysis of empirical data and adding relatively simple input data: monthly generation and daytime demand totals.

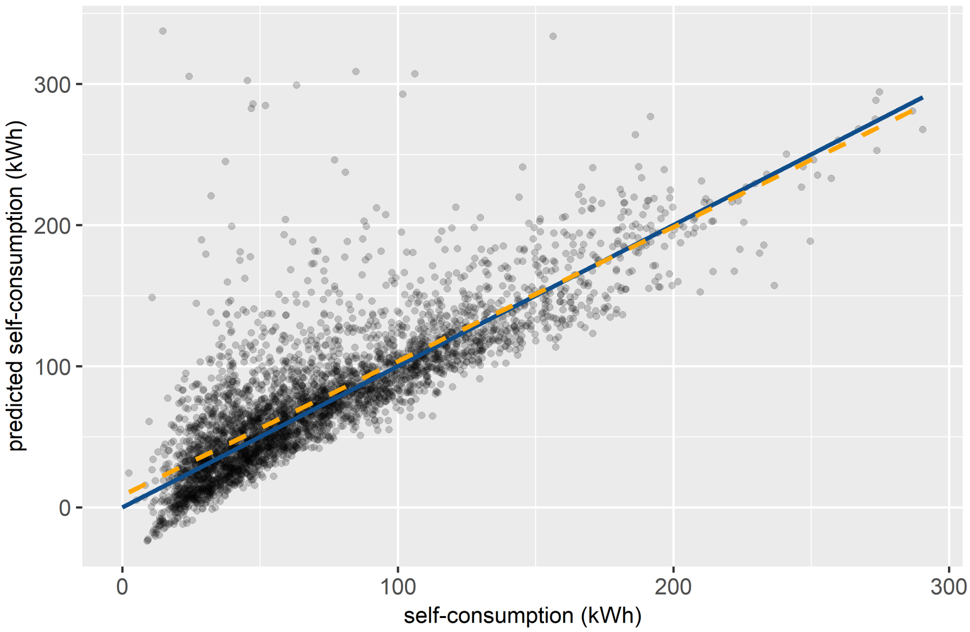

Figure 7: Actual self-consumption for households with PV in the UK (CLNR data) compared to their predicted self-consumption using coefficients derived from an independent group of households with PV in UK (PVOutput data). Solid blue line indicates perfect prediction, dashed yellow line is a linear regression of the points shown.

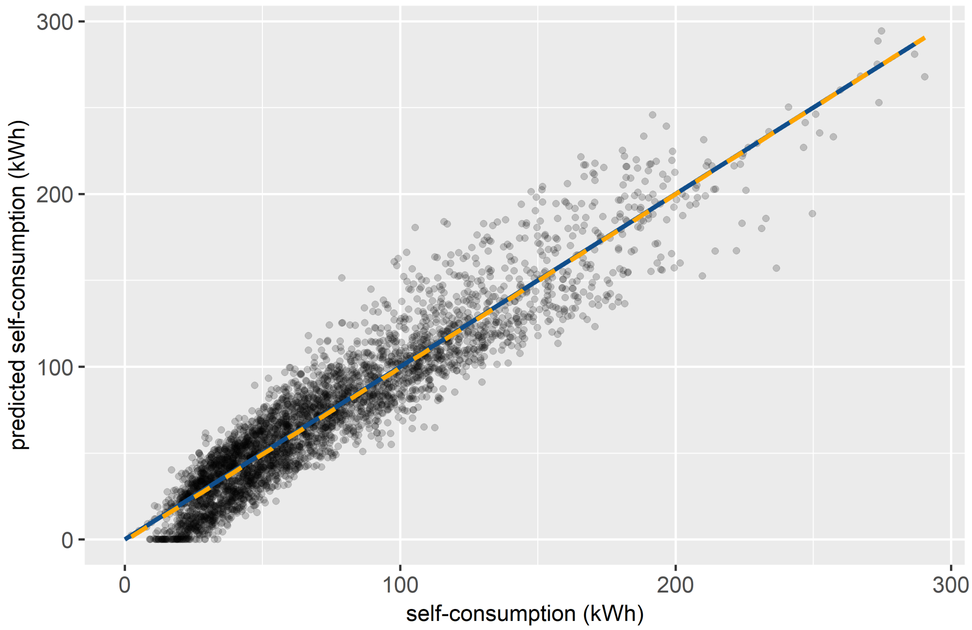

Figure 8: Same data as Figure 7 but restricting predictions to between zero and generation i.e. no impossible self-consumption predictions (negative or more than 100%).

Conclusion

This paper presents an analysis of international empirical data on self-consumption in households with PV. There is considerable variation in self-consumption between different households, but this is relatively well explained by differences in the amount of solar electricity generated and the amount of electricity consumed during the day (r-squared 0.915). The results show that the factors affecting self-consumption are similar between the different countries analysed here (Australia, Belgium, Germany, Netherlands, UK, USA).

We show how regression coefficients derived from one group of households with PV can be used to predict self-consumption for an independent group of households from the same country with a reasonable level of accuracy (r-squared 0.864) given relatively simple input values (monthly generation and daytime demand totals).

Empirical self-consumption data is still relatively rare and the development of methods for comparing and combining independent international data are still a matter for further research, in particular the need to account for environmental, structural, economic and behavioural differences between countries.

More broadly, the need within research and evidence-based policy for large representative samples of fully metered households with PV in different nations remains unresolved. Achieving this will remain challenging and expensive unless 'behind the meter' distributed micro-generation is metered, integrated into national smart meter communications infrastructures, and consumers are willing and able to share this data with trusted third-parties.

The lack of visibility of distributed generation is not only an issue for those interested in self-consumption. Accurate observations of generation are required to account for the carbon emissions impact of distributed generation. They are also required for peer to peer energy trading if the resource being traded is generation rather than exports.

Unmetered distributed generation also obscures observations of gross demand. As a result, anything that relies on accurate readings of gross (rather than net) demand will be increasingly affected as micro-generation continues to be deployed at scale. Accurate readings of gross demand are important. Without them we lose the ability to accurately estimate the impact of new technologies or services on demand such as energy efficiency technologies or, indeed, PV itself. Solar PV is highly popular, and costs will continue to fall; it is conceivable that it could become adopted by many more households around the world over the coming decades. If this happens it will pose a major issue for observations of gross demand and any metrics or instruments that rely on it.

Nations around the world are spending considerable funds on national metering infrastructures that do not meter distributed generation. In a world where we expect very large numbers of households to adopt micro-generation such as solar PV, this is an oversight that should be corrected in the interest of both consumers and the broader energy system. All micro-generation should require its own meter point and these readings should be integrated into national metering infrastructures, alongside import and export readings, and this will require changes to national metering legislation and regulation.

Acknowledgements

This work has been funded by the Engineering and Physical Sciences Research Council through the Smart Meter Research Portal (EP/P032761/1) and the Impact Acceleration Account Award (EP/R511742/1), and by the John Fell Oxford University Press (OUP) Research Fund (163/067).

References

BEIS (2014) *Smart metering equipment technical specifications: second version - GOV.UK*. Available at: https://www.gov.uk/government/consultations/smart-metering-equipment-technical-specifications-second-version (Accessed: 11 January 2019).

BEIS (2017) *Energy and Climate Change Public Attitudes Tracker: Wave 22 - GOV.UK*. Available at: https://www.gov.uk/government/statistics/energy-and-climate-change-public-attitudes-tracker-wave-22 (Accessed: 11 January 2019).

BEIS (2018a) *Solar photovoltaic (PV) cost data - GOV.UK*. Available at: https://www.gov.uk/government/statistics/solar-pv-cost-data (Accessed: 14 January 2019).

BEIS (2018b) *Updated energy and emissions projections: 2017 - GOV.UK. Annex M: Growth assumptions and prices*. Available at: https://www.gov.uk/government/publications/updated-energy-and-emissions-projections-2017 (Accessed: 11 January 2019).

European Parliament (2018) *Directive (EU) 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources (Text with EEA relevance.)*. Available at: http://data.europa.eu/eli/dir/2018/2001/oj (Accessed: 11 January 2019).

Leicester, P. A., Goodier, C. I. and Rowley, P. N. (2016) 'Probabilistic analysis of solar photovoltaic self-consumption using Bayesian network models', *IET Renewable Power Generation*, 10(4), pp. 448–455. doi: 10.1049/iet-rpg.2015.0360.

McKenna, E., Pless, J. and Darby, S. J. (2018) 'Solar photovoltaic self-consumption in the UK residential sector: New estimates from a smart grid demonstration project', *Energy Policy*, 118. doi: 10.1016/j.enpol.2018.04.006.

McKenna, E. and Thomson, M. (2013) 'Photovoltaic metering configurations, feed-in tariffs and the variable effective electricity prices that result', *IET Renewable Power Generation*, 7(3). doi: 10.1049/iet-rpg.2012.0268.

McKenna, E. and Thomson, M. (2014) 'Demand response behaviour of domestic consumers with photovoltaic systems in the UK: an exploratory analysis of an internet discussion forum', *Energy, Sustainability and Society*, 4(1). doi: 10.1186/2192-0567-4-13.

NSW Government (no date) *The solar bonus scheme is closed*. Available at: https://energysaver.nsw.gov.au/households/solar-and-battery-power/solar-bonus-scheme-closed (Accessed: 11 January 2019).

Ofgem (2018) *Feed-In Tariff (FIT) rates*. Available at: https://www.ofgem.gov.uk/environmental-programmes/fit/fit-tariff-rates (Accessed: 11 January 2019).

Vaughan, A. (2018) *Majority of UK public want to install solar panels, poll finds | Money | The Guardian*, *The Guardian*. Available at: https://www.theguardian.com/money/2018/aug/20/majority-of-uk-public-want-to-install-solar-panels-poll-finds (Accessed: 11 January 2019).

Webborn, E., Elam, S. and McKenna, E. (2019) 'Utilising Smart Meter Data for Research and Innovation in the UK (forthcoming)', in *Proceedings of the European Council for an Energy Efficient Economy Summer Study*.