

NZ GREEN Grid Household Electricity Demand Study

Research data overview (version 1.0)

Anderson, B., Eyers, D., Ford, R., Giraldo Ocampo, D., Peniamina, R., Stephenson, J., Suomalainen, K., Wilcocks, L. and Jack, M.

Last run at: 2018-08-30 14:45:40 (Pacific/Auckland)

1 About

1.1 Report circulation:

- Public – this report is intended to accompany the data release.

1.2 License

This work is made available under the Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) License (<https://creativecommons.org/licenses/by-sa/4.0/>).

This means you are free to:

- *Share* – copy and redistribute the material in any medium or format
- *Adapt* – remix, transform, and build upon the material for any purpose, even commercially.

Under the following terms:

- *Attribution* – You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- *ShareAlike* – If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.
- *No additional restrictions* – You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

- You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.
- No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material. #YMMV

For the avoidance of doubt and explanation of terms please refer to the full license notice (<https://creativecommons.org/licenses/by-sa/4.0/>) and legal code (<https://creativecommons.org/licenses/by-sa/4.0/legalcode>).

1.3 Citation

If you wish to use any of the material from this report please cite as:

- Anderson, B., Eyers, D., Ford, R., Giraldo Ocampo, D., Peniamina, R., Stephenson, J., Suomalainen, K., Wilcocks, L. and Jack, M. (2018) NZ GREEN Grid Household Electricity Demand Study: Research data overview (version 1.0) , Centre for Sustainability (<http://www.otago.ac.nz/centre-sustainability/>), University of Otago: Dunedin.

This work is (c) 2018 the authors.

1.4 History

You may not be reading the most recent version of this report. Please check:

- the overall package documentation (<https://cfsotago.github.io/GREENGridData/>);
- this report's edit history (<https://github.com/CfSOtago/GREENGridData/commits/master/makeDocs/buildOverviewReport.Rmd>)

1.5 Support

This work was supported by:

- The University of Otago (<https://www.otago.ac.nz/>);
- The University of Southampton (<https://www.southampton.ac.uk/>);
- The New Zealand Ministry of Business, Innovation and Employment (MBIE) (<http://www.mbie.govt.nz/>) through the NZ GREEN Grid (<https://www.otago.ac.nz/centre-sustainability/research/energy/otago050285.html>) grant (Contract ID: UOCX1203);
- SPATIALEC (<http://www.energy.soton.ac.uk/tag/spatialec/>) - a Marie Skłodowska-Curie Global Fellowship (http://ec.europa.eu/research/mariecurieactions/about-msca/actions/if/index_en.htm) based at the University of Otago's Centre for Sustainability (<http://www.otago.ac.nz/centre-sustainability/staff/otago673896.html>) (2017-2019) & the University of Southampton's Sustainable Energy Research Group (2019-2022).

2 Introduction

The NZ GREEN Grid household electricity demand study (<https://cfsotago.github.io/GREENGridData/>) recruited a sample of c 25 households in each of two regions of New Zealand (Stephenson et al. 2017). The first sample was recruited in early 2014 and the second in early 2015.

Research data includes:

- 1 minute electricity power (W) data was collected for each dwelling circuit using GridSpy (<https://gridspy.com/>) monitors on each power circuit (and the incoming power). The power values represent mean(W) over the minute preceeding the observation timestamp;
- Dwelling & appliance surveys;
- Occupant time-use diaries (focused on energy use).

NB: Version 1 of the data package does not include the time-use diaries.

This report provides an overview of the GREEN Grid project (Stephenson et al. 2017) research data.

3 Data Package

Version 1.0 of the data package contains:

- powerData.zip: 1 minute power demand data for each circuit in each household. One file per household;
- ggHouseholdAttributesSafe.csv.zip: anonymised household attribute data;
- checkPlots.zip:
 - simple line charts of mean power per month per year for each circuit monitored for each household. These are a useful check;
 - tile plots (heat maps/carpet plots) of the number of observations per hour per day. Also a useful check...

4 Study recruitment

The project research sample comprises 44 households who were recruited via the local power lines companies in two areas: New Plymouth starting in May 2014 and Hawkes Bay starting in November 2014.

Recruitment was via a non-random sampling method and a number of households were intentionally selected for their 'complex' electricity consumption (and embedded generation) patterns and appliances (Giraldo Ocampo 2015, Stephenson et al. (2017), Jack et al. (2018), Suomalainen et al. (2017)).

The lines companies invited their own employees and those of other local companies to participate in the research and ~80 interested potential participants completed short or long forms of the Energy Cultures 2 household survey (Wooliscroft 2015). Households were then selected from this pool by the project team based on selection criteria relevant to the GREEN Grid project. These included:

- having the majority of their energy supply from electricity (i.e. not gas heating);
- household size;
- types of appliances owned.

After informed consent was obtained from each household, an electrician contracted by the two lines companies completed an appliance survey to record detailed information about the appliances in each house. This survey contained information about the number of appliances owned, brand, model number, efficiency and age. The electrician also installed the GridSpy units which recorded electricity power demand at a circuit level. The GridSpy units automatically upload the monitoring data to the GridSpy company's secure database from where it was downloaded by the GREEN Grid research team.

As a result of this process the sample cannot be assumed to represent the population of customers (or employees) of any of the companies involved, nor the populations in each location (Stephenson et al. 2017).

Table 4.1 shows the number in each sample.

Table 4.1: Sample location

Location	nHouseholds
Hawkes Bay	20
New Plymouth	24

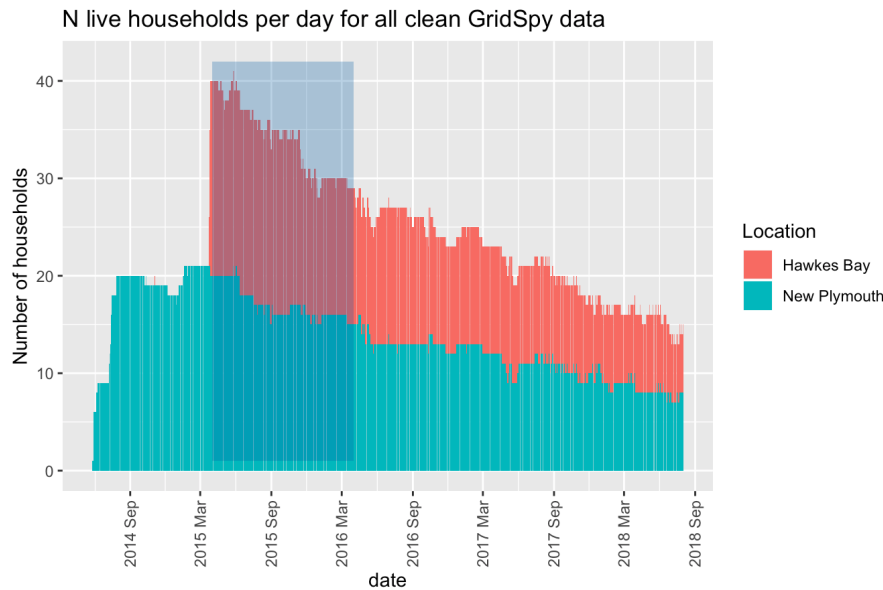
Table 4.2 shows the number for whom valid appliance and survey data is available in this data package. Note that even those which appear to lack appliance data may have sufficient survey data to deduce appliance ownership (see question numbers Q19_* and Q40_*).

Table 4.2: Sample information

Location	hasShortSurvey	hasLongSurvey	hasApplianceSummary	nHouseholds
Hawkes Bay	NA	NA	NA	1
Hawkes Bay	NA	NA	Yes	1
Hawkes Bay	NA	Yes	Yes	5
Hawkes Bay	Yes	NA	Yes	13
New Plymouth	NA	Yes	NA	12
New Plymouth	NA	Yes	Yes	12

5 Data collection duration

Figure 5.1 shows the total number of households for whom GridSpy data exists on a given date by sample. The plot includes any data, including partial data and suggests that for analytic purposes the period from April 2015 to March 2016 (indicated) would offer the maximum number of households.



Source: GREEN Grid Project data 2014-05-26 to 2018-08-01
 Shaded area indicates 12 month period with largest number of households

Figure 5.1: Number of households sending GridSpy data by date

6 Key attributes

Table 6.1 shows key attributes for the recruited sample. Note that two GridSpy monitors were re-used and so require new hhIDs to be set from the date of re-use using the linkID variable. This is explained in more detail in the GridSpy processing (gridSpy1mProcessingReport_v1.0.html#reallocation) report. Linkage between the survey and GridSpy data should therefore always use linkID to avoid errors.

Table 6.1: Sample details

hhID	linkID	Location	surveyStartDate	nAdults	nChildren0_12	nTeenagers13_18	notes	r_stopDate	hasApplianceSumma
rf_06	rf_06	New Plymouth	2014-05-19 09:49:00	2	0	0	NA	NA	NA
rf_07	rf_07	New Plymouth	2014-06-23 21:25:00	2	2	0	NA	NA	NA
rf_08	rf_08	New Plymouth	2014-05-14 12:21:00	2	0	0	NA	NA	NA
rf_09	rf_09	New Plymouth	2014-06-19 11:33:00	2	1	0	NA	NA	NA
rf_10	rf_10	New Plymouth	2014-05-20 17:01:00	2	1	0	NA	NA	Yes
rf_11	rf_11	New Plymouth	2014-06-06 12:16:00	2	NA	NA	NA	NA	Yes
rf_12	rf_12	New Plymouth	2014-06-16 07:34:00	1	0	0	NA	NA	NA
rf_13	rf_13	New Plymouth	2014-05-14 12:07:00	2	1	1	NA	NA	Yes
rf_14	rf_14	New Plymouth	2014-06-10 11:51:00	1	1	0	NA	NA	Yes
rf_15	rf_15a	New Plymouth	2014-06-17 15:38:00	1	0	0	Disconnected 15/01/2015. Re-used	2015-01-15	NA

hhID	linkID	Location	surveyStartDate	nAdults	nChildren0_12	nTeenagers13_18	notes	r_stopDate	hasApplianceSumma
rf_15	rf_15b	New Plymouth	2014-05-16 17:36:00	2	0	0	Re-user. Then disconnected 02/04/2016	2016-04-02	NA
rf_16	rf_16	New Plymouth	2014-06-10 15:29:00	2	0	0	NA	NA	NA
rf_17	rf_17a	New Plymouth	2014-05-14 20:04:00	2	3	1	Unusual & specialist energy tech configuration. Disconnected 28/03/2016. Re-used.	2016-03-28	NA
rf_17	rf_17b	New Plymouth	2014-05-22 09:16:00	NA	NA	NA	Re-user	NA	NA
rf_18	rf_18	New Plymouth	2014-05-14 11:20:00	2	1	0	NA	NA	NA
rf_19	rf_19	New Plymouth	2014-05-22 13:37:00	1	0	0	NA	NA	Yes
rf_20	rf_20	New Plymouth	2014-05-14 11:46:00	2	2	0	NA	NA	NA
rf_21	rf_21	New Plymouth	2014-05-20 16:30:00	2	0	0	NA	NA	Yes
rf_22	rf_22	New Plymouth	2014-05-14 11:39:00	2	0	0	NA	NA	Yes
rf_23	rf_23	New Plymouth	2014-05-15 15:51:00	1	0	0	NA	NA	Yes
rf_24	rf_24	New Plymouth	2014-05-14 11:36:00	2	2	0	NA	NA	Yes
rf_25	rf_25	New Plymouth	2014-06-18 13:57:00	1	1	0	NA	NA	Yes
rf_26	rf_26	New Plymouth	2014-06-11 13:34:00	2	0	0	NA	NA	Yes
rf_27	rf_27	New Plymouth	2014-07-03 15:37:00	2	1	1	NA	NA	Yes
rf_28	rf_28	Hawkes Bay	2015-01-20 12:15:00	2	2	0	NA	NA	Yes
rf_29	rf_29	Hawkes Bay	2015-02-10 11:39:00	2	1	0	NA	NA	Yes
rf_30	rf_30	Hawkes Bay	2015-02-03 10:58:00	2	0	2	NA	NA	Yes
rf_31	rf_31	Hawkes Bay	2015-02-09 08:05:00	3	2	0	NA	NA	Yes
rf_32	rf_32	Hawkes Bay	2015-02-09 08:35:00	2	2	0	NA	NA	Yes
rf_33	rf_33	Hawkes Bay	2015-02-09 16:05:00	2	1	1	NA	NA	Yes
rf_34	rf_34	Hawkes Bay	2015-01-06 10:50:00	3	0	0	NA	NA	Yes
rf_35	rf_35	Hawkes Bay	2015-02-05 16:00:00	2	2	0	NA	NA	Yes

hhID	linkID	Location	surveyStartDate	nAdults	nChildren0_12	nTeenagers13_18	notes	r_stopDate	hasApplianceSumma
rf_36	rf_36	Hawkes Bay	2015-02-10 20:25:00	1	0	2	NA	NA	Yes
rf_37	rf_37	Hawkes Bay	2015-02-09 18:49:00	2	2	0	NA	NA	Yes
rf_38	rf_38	Hawkes Bay	2015-02-05 15:30:00	2	2	0	very large number of circuits	NA	Yes
rf_39	rf_39	Hawkes Bay	2015-02-05 15:43:00	3	0	1	NA	NA	Yes
rf_40	rf_40	Hawkes Bay	NA	2	0	0	NA	NA	Yes
rf_41	rf_41	Hawkes Bay	2015-01-12 13:16:00	2	3	0	NA	NA	Yes
rf_42	rf_42	Hawkes Bay	2015-02-10 18:04:00	2	3	0	NA	NA	Yes
rf_43	rf_43	Hawkes Bay	NA	2	1	0	NA	NA	NA
rf_44	rf_44	Hawkes Bay	2015-02-04 20:47:00	2	2	1	NA	NA	Yes
rf_45	rf_45	Hawkes Bay	2015-02-09 13:26:00	2	3	0	NA	NA	Yes
rf_46	rf_46	Hawkes Bay	2014-12-19 08:40:00	2	1	0	NA	NA	Yes
rf_47	rf_47	Hawkes Bay	2015-01-06 09:01:00	3	0	1	NA	NA	Yes

7 Code examples

We have provided a number of code examples (<https://github.com/CfSotago/GREENGridData/tree/master/examples>) for suggestions on how to load, further process and analyse the data.

8 Known issues

We maintain a known data issues list (<https://github.com/CfSotago/GREENGridData/labels/datalssue>) via our GitHub repository. If you think there is a data issue please check the repo list first and then add a new one if appropriate.

9 Runtime

Analysis completed in 5.34 seconds (0.09 minutes) using knitr (<https://cran.r-project.org/package=knitr>) in RStudio (<http://www.rstudio.com>) with R version 3.5.1 (2018-07-02) running on x86_64-apple-darwin15.6.0.

10 R environment

10.1 R packages used

- base R (R Core Team 2016)
- bookdown (Xie 2016a)
- GREENGridData (Anderson and Evers 2018) which depends on:
 - data.table (Dowle et al. 2015)
 - dplyr (Wickham and Francois 2016)
 - hms (Müller 2018)
 - lubridate (Grolemund and Wickham 2011)
 - progress (Csárdi and FitzJohn 2016)
 - readr (Wickham, Hester, and Francois 2016)
 - readxl (Wickham and Bryan 2017)
 - reshape2 (Wickham 2007)
- ggplot2 (Wickham 2009)
- kableExtra (Zhu 2018)
- knitr (Xie 2016b)
- rmarkdown (Allaire et al. 2018)

10.2 Session info

```
## R version 3.5.1 (2018-07-02)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS High Sierra 10.13.6
##
## Matrix products: default
## BLAS: /System/Library/Frameworks/Accelerate.framework/Versions/A/Frameworks/vecLib.framework/Versions/A/libBLAS.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_GB.UTF-8/en_GB.UTF-8/en_GB.UTF-8/C/en_GB.UTF-8/en_GB.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] stringr_1.3.1      bindrcpp_0.2.2     kableExtra_0.9.0
## [4] bookdown_0.7       rmarkdown_1.10    knitr_1.20.13
## [7] skimr_1.0.3        readxl_1.1.0      readr_1.1.1
## [10] lubridate_1.7.4    ggplot2_3.0.0     data.table_1.11.4
## [13] GREENGridData_1.0
##
## loaded via a namespace (and not attached):
## [1] Rcpp_0.12.18      highr_0.7         cellranger_1.1.0
## [4] pillar_1.3.0     compiler_3.5.1    plyr_1.8.4
## [7] bindr_0.1.1      prettyunits_1.0.2 tools_3.5.1
## [10] progress_1.2.0   digest_0.6.15     viridisLite_0.3.0
## [13] evaluate_0.11    tibble_1.4.2      gtable_0.2.0
## [16] pkgconfig_2.0.2  rlang_0.2.2       rstudioapi_0.7
## [19] yaml_2.2.0       xfun_0.3          xml2_1.2.0
## [22] httr_1.3.1       withr_2.1.2       dplyr_0.7.6
## [25] hms_0.4.2        rprojroot_1.3-2   grid_3.5.1
## [28] tidyselect_0.2.4 glue_1.3.0        R6_2.2.2
## [31] reshape2_1.4.3   purrr_0.2.5       magrittr_1.5
## [34] backports_1.1.2  scales_1.0.0      htmltools_0.3.6
## [37] rvest_0.3.2      assertthat_0.2.0  colorspace_1.3-2
## [40] labeling_0.3     stringi_1.2.4     lazyeval_0.2.1
## [43] munsell_0.5.0    crayon_1.3.4
```

References

- Allaire, JJ, Yihui Xie, Jonathan McPherson, Javier Luraschi, Kevin Ushey, Aron Atkins, Hadley Wickham, Joe Cheng, and Winston Chang. 2018. *Rmarkdown: Dynamic Documents for R*. <https://CRAN.R-project.org/package=rmarkdown> (<https://CRAN.R-project.org/package=rmarkdown>).
- Anderson, Ben, and David Eyers. 2018. *GREENGridData: Processing Nz Green Grid Project Data to Create a 'Safe' Version for Data Archiving and Re-Use*. <https://github.com/CfSotago/GREENGridData> (<https://github.com/CfSotago/GREENGridData>).
- Csárdi, Gábor, and Rich FitzJohn. 2016. *Progress: Terminal Progress Bars*. <https://CRAN.R-project.org/package=progress> (<https://CRAN.R-project.org/package=progress>).
- Dowle, M, A Srinivasan, T Short, S Lianoglou with contributions from R Saporta, and E Antonyan. 2015. *Data.table: Extension of Data.frame*. <https://CRAN.R-project.org/package=data.table> (<https://CRAN.R-project.org/package=data.table>).
- Giraldo Ocampo, Diana. 2015. "Developing an Energy-Related Time-Use Diary for Gaining Insights into New Zealand Households' Electricity Consumption." Master's thesis, Centre for Sustainability: University of Otago. <http://hdl.handle.net/10523/5957> (<http://hdl.handle.net/10523/5957>).
- Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. <http://www.jstatsoft.org/v40/i03/> (<http://www.jstatsoft.org/v40/i03/>).
- Jack, M. W., K. Suomalainen, J. J. W. Dew, and D. Eyers. 2018. "A Minimal Simulation of the Electricity Demand of a Domestic Hot Water Cylinder for Smart Control." *Applied Energy* 211: 104–12.
- Müller, Kirill. 2018. *Hms: Pretty Time of Day*. <https://CRAN.R-project.org/package=hms> (<https://CRAN.R-project.org/package=hms>).
- R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/> (<https://www.R-project.org/>).
- Stephenson, Janet, Rebecca Ford, Nirmal-Kumar Nair, Neville Watson, Alan Wood, and Allan Miller. 2017. "Smart Grid Research in New Zealand—A Review from the GREEN Grid Research Programme." *Renewable and Sustainable Energy Reviews* 82 (1): 1636–45. <https://doi.org/10.1016/j.rser.2017.07.010> (<https://doi.org/10.1016/j.rser.2017.07.010>).
- Suomalainen, Kiti, Michael Jack, David Byers, Rebecca Ford, and Janet Stephenson. 2017. "Comparative Analysis of Monitored and Self-Reported Data on Electricity Use." In *Environment and Electrical Engineering and 2017 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I&CPS Europe), 2017 IEEE International Conference on*, 1–4. IEEE.
- Wickham, Hadley. 2007. "Reshaping Data with the reshape Package." *Journal of Statistical Software* 21 (12): 1–20. <http://www.jstatsoft.org/v21/i12/> (<http://www.jstatsoft.org/v21/i12/>).

- — —. 2009. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <http://ggplot2.org> (<http://ggplot2.org>).
- Wickham, Hadley, and Jennifer Bryan. 2017. *Readxl: Read Excel Files*. <https://CRAN.R-project.org/package=readxl> (<https://CRAN.R-project.org/package=readxl>).
- Wickham, Hadley, and Romain Francois. 2016. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr> (<https://CRAN.R-project.org/package=dplyr>).
- Wickham, Hadley, Jim Hester, and Romain Francois. 2016. *Readr: Read Tabular Data*. <https://CRAN.R-project.org/package=readr> (<https://CRAN.R-project.org/package=readr>).
- Wooliscroft, B. 2015. "National Household Survey of Energy and Transportation: Energy Cultures Two." Centre for Sustainability: University of Otago. <http://hdl.handle.net/10523/5634> (<http://hdl.handle.net/10523/5634>).
- Xie, Yihui. 2016a. *Bookdown: Authoring Books and Technical Documents with R Markdown*. Boca Raton, Florida: Chapman; Hall/CRC. <https://github.com/rstudio/bookdown> (<https://github.com/rstudio/bookdown>).
- — —. 2016b. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. <https://CRAN.R-project.org/package=knitr> (<https://CRAN.R-project.org/package=knitr>).
- Zhu, Hao. 2018. *KableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra> (<https://CRAN.R-project.org/package=kableExtra>).