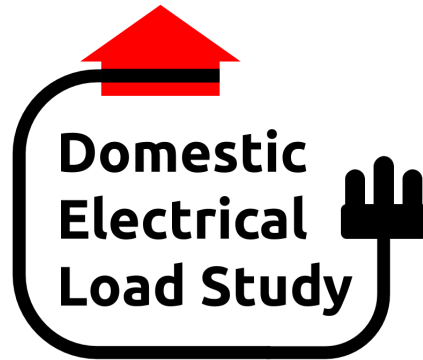


Big Data for Household Energy Insights



Project Report

7 February 2020

Submitted by Wiebke Toussaint
on behalf of the Energy Research Centre
at the University of Cape Town



ERC

ENERGY RESEARCH CENTRE
University of Cape Town

Project Overview

The goals of the Big Data for Household Energy Insights project were to provide strategic support and leadership to convene and mobilise key stakeholders towards building big data capabilities for domestic load research and to deliver technical data science and data stewardship activities in this regard.

The project has delivered South Africa's first online, citeable datasets of Domestic Electrical Load (DEL) studies in South Africa. The published DEL datasets capture valuable household energy consumption and survey data that covers over two decades of research gathered under the NRS Load Research Programme. Hosted securely at DataFirst, Africa's only internationally accredited data archive, the datasets are now available to researchers for non-commercial use. The DEL datasets cover the period of major electrification in South Africa from 1994 to 2014 and span across five climatic zones, two time zones and various population groups. They are unique on the African continent and their release marks the pioneering step that key stakeholders - namely Eskom, a group of municipal stakeholders, the University of Cape Town, Stellenbosch University, professionals and specialists working in the load research space - have taken to advance data-driven energy research through digitisation and open data. Together with the five DEL datasets that have been published, the project has released two open source python packages to facilitate data access and processing.

Beyond delivering the technical milestones, the project has ensured rigorous and frequent industry engagement. In August 2018 the project convened an initial stakeholder workshop to bring together stakeholders from across the country in order to formulate a vision for load research in South Africa. At the workshop a task team was appointed to work towards the release of the datasets. Less than a year later in June 2019 an expanded group of stakeholders was invited to attend a feedback workshop that celebrated the successful launch of both the DEL datasets and a new Load Research Chapter at the South African Institute of Electrical Engineering (SAIEE). The project delivered a webinar to the SAIEE Load Research Chapter to showcase how machine learning and artificial intelligence can be applied to the DEL datasets to better understand typical household demand patterns in South Africa. The webinar was attended by an international audience of over 45 people and a paper on this work was presented at the South African Forum for Artificial Intelligence Research in December 2019.

To further the usage of the dataset and the software that has been created, 7 training videos have been created to help incoming researchers understand and work with the DEL datasets. A data descriptor for the datasets has been written and is in the process of being submitted for publication. Links to all the deliverables of the project are provided in the next section.

Deliverables

Strategic Task Team Support

Strategic support was delivered to the NRS Interim Task Team by convening, participating in and where applicable recording the minutes of meetings on the following dates:

14 November 2018

28 January 2019

4 April 2019 (Load Research Chapter Launch)

9 May 2019 (DataFirst Site Acceptance Test)

Datasets

The following five datasets have been published:

Toussaint, Wiebke. *Domestic Electrical Load Metering, Hourly Data 1994-2014* [dataset]. version 1. Johannesburg: SANEDI [funders]. Cape Town: UCT [producers], 2014. Cape Town: DataFirst [distributor], 2019. <https://doi.org/10.25828/56nh-fw77>

Toussaint, Wiebke. *Domestic Electrical Load Survey - Key Variables 1994-2014* [dataset]. version 1. Johannesburg: SANEDI [funders]. Cape Town: UCT Energy Research Centre [producers], 2014. Cape Town: DataFirst [distributor], 2019. <https://doi.org/10.25828/mf8s-hh79>

Eskom, Stellenbosch University, University of Cape Town. *Domestic Electrical Load Metering Data 1994-2014* [dataset]. version 1. Johannesburg: Eskom, Cape Town: UCT, Stellenbosch: US [producers], 2014. Cape Town: DataFirst [distributor], 2019. <https://doi.org/10.25828/p3k7-r965>

Eskom, Stellenbosch University, University of Cape Town. *Domestic Electrical Load Survey 1994-2014* [dataset]. version 1. Johannesburg: Eskom, Cape Town: UCT, Stellenbosch: US [producers], 2014. Cape Town: DataFirst [distributor], 2019. doi:[10.25828/kzer-gd88](https://doi.org/10.25828/kzer-gd88)

Eskom, Stellenbosch University, University of Cape Town. *Domestic Electrical Load Survey-Secure Data 1994-2014* [dataset]. version 1. Johannesburg: Eskom, Cape Town: UCT, Stellenbosch: US [producers], 2014. Cape Town: DataFirst [distributor], 2019. doi:[10.25828/7jtv-mc31](https://doi.org/10.25828/7jtv-mc31)

Publications

Toussaint, W. *The Domestic Electrical Load (DEL) Study Datasets for South Africa*. 2020. <https://doi.org/10.25375/uct.11774691.v1>

Toussaint, W. and Moodley, D. *Comparison of Clustering Techniques for Residential Load Profiles in South Africa*. Workshop Proceedings of the South African Forum for AI Research. CEUR-WS, Vol. 2540, 2019. http://ceur-ws.org/Vol-2540/FAIR2019_paper_55.pdf

Code

delretrieve facilitates data retrieval from a SQL Server installation of the South African Domestic Electrical Load database.

Toussaint, Wiebke. *delretrieve*: Data Retrieval of the South African Domestic Electrical Load Study, version 1.01. Zenodo. <https://doi.org/10.5281/zenodo.3605425> (2019).

delprocess helps researchers extract, clean, resample and enumerate load profile and survey data from a local file hierarchy retrieved from South Africa's Domestic Load Research database.

Toussaint, Wiebke. *delprocess*: Data Processing of the South African Domestic Electrical Load Study, version 1.01. Zenodo. <https://doi.org/10.5281/zenodo.3605422> (2019).

Public Engagement

SAIEE Load Research Chapter Webinar, 30 July 2019. *Shedding light on residential energy consumer behaviour in South Africa*. <https://youtu.be/ULBfbPN5i7I>

UCT e-Research report 2019: <https://www.paperturn-view.com/?pid=NjY66984&v=1.1&p=24>
Plot.ly Dash for good blog: <https://plot.ly/resources/erc/>

Training

The DEL training videos provide an introduction to accessing and using the DEL datasets. In particular, they are a step to step guide for installing and using the two python packages that have been developed to facilitate data access. The training videos can be accessed on Zivahub, UCT's open data repository:

<https://doi.org/10.25375/uct.11774607.v1>