

Domestic Load Research Database Handbook

Wiebke Toussaint

19 December 2016

This handbook assumes that you have access to a live instance of the DLR database on SQL Server. The development of the DLR programme was driven by the need to measure individuals' load profiles as well as understanding **who** the individuals were.

Data Overview

Parameters monitored

In designing the programme, the cause and effect correlations of previously performed load research studies were mapped and condensed to the parameters that were ultimately monitored in the study.

Data collected with questionnaires

For a full list of survey questions asked, view the supporting document `NLRquestionnaires.Rmd`

Socio-demographic data: information about the people

- Number of people living in house
- Age of home owner
- Gender of home owner
- Level of education of home owner
- Income of home owner
- Employment status of home owner

Information about the dwelling

- Number of rooms
- Street address
- GPS coordinates
- Structure (ceiling or not)
- Floor area
- Building materials

Information about appliances

- Number of light bulbs in dwelling (does not distinguish energy saving bulbs)
- Number of people being cooked for
- Fuel source for cooking (quantities used not captured)
- Frequency of cooking

Load data collected

Project Timeline 1993 - 2014

Sites & number of people surveyed over the history of the project

On average the DLR project has monitored:

- 800 households per year
- with +- 60 households per site
- over a spread of 13 sites
- monitoring period 5 years/site for Eskom customers, shorter for municipal customers

Year | Number of Sites Monitored ———| ————— 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000
| 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 |

```
DECLARE @ProjectYears TABLE (GroupID int, GroupName varchar(50), Description varchar(100))
INSERT INTO @ProjectYears(GroupID, GroupName, Description)
    SELECT GroupID, GroupName, Description
        FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IN
            (SELECT GroupID FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IN
                (SELECT GroupID FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IS NULL))
```

```
DECLARE @SiteSummary TABLE (ParentID int, SiteCount int)
INSERT INTO @SiteSummary
    SELECT [General_LR4].[dbo].[Groups].ParentID, COUNT([General_LR4].[dbo].[Groups].ParentID) AS SiteC
    FROM [General_LR4].[dbo].[Groups]
    GROUP BY ParentID
```

```
DECLARE @YearSummary TABLE (GroupID int, GroupName varchar(50), Description varchar(50), SiteCount int)
INSERT INTO @YearSummary
    SELECT GroupID, GroupName, Description, SiteCount FROM @ProjectYears p
    LEFT JOIN @SiteSummary
    ON p.GroupID=[@SiteSummary].ParentID
    WHERE GroupID != 89 AND GroupID != 90 AND GroupID != 91
    ORDER BY GroupName, GroupID
SELECT SUM(SiteCount) FROM @YearSummary
```

Project Evolution over time

1993

| Project start

2000

| added rural households | added GPS coordinates (taken at front gate/door of dwelling)
| combined Eskom & municipal data sets

| consolidated the two existing versions of socio-demographic data collected in the questionnaires

2004

| matched prepaid transactions with meter numbers to verify household energy consumption
| data loggers measured up to 16 channels current & one voltage across all channels
| > inferred power from measurement
| > errors at low current

2009

| data logger format changed | > wireless logging
| > 1 channel for up to 3 houses
| > voltage, current, real power and reactive power recorded
| > changed to cts that can handle low current error
| did not manage to close calibration loop (ie verifying calibration of individual appliances)

2012

| migrated to mobile data capture
| added some 'jump driven' loggers to separate out appliances

Data Checking

Socio-demographic verification

The DLR has heavily standardised data treatment. Socio-demographic data was collected in 2 phases. **Phase 1** (pre 2000) captured on paper was verified with a visual test. For **Phase 2** (post 2000) error checking has been implemented with rules checking on the database. The following validation process has been followed:

- verify if a data source or data capture problem exists
- correct demographic data if possible by checking the paper source
- remaining errors are residual and not possible to solve by inspection

Socio checking rules

Measurement verification

Site reference database

Every winter the connectivity has been traced from the data logger to each household to verify the recording channels. From this, a site reference database has been created.

Data Validation

Database Structure

Database Diagrams (on SQL)

Link Table

PQ Tables

Tables

Primary Tables

dbo.Context The context table provides a list of all the projects and project phases that were executed as part of the DLR Programme. A project is a load research study that was implemented outside the ‘normal’ South African residential electricity sector, which was the focus of the programme. A phase marks a point in time where the programme underwent a significant change of methodology, such as a change in questionnaire or logger used. The projects and phases are:

- NRS LR 2000
- NRS LR 1994 -1999
- NRS LR 2009
- NRS LR 2010
- Namibia 2000
- Non-domestic - Clinics (NonDom_clinics)
- Non-domestic - Shops (NonDom_shops)
- Non-domestic - Schools (NonDom_schools)

dbo.Groups A group is a set of households that was surveyed and whose load profiles were recorded as part of a project or phase within the DLR Programme in a specific year, at a specific location. The ParentID attribute can be used to navigate through the groups structure. Use the SQL code below to get a better understanding of how groups are set up.

```
# Top level (DLR/non-domestic)
```

```
SELECT *  
FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IS NULL
```

```
# Second level (grouped by PROJECT; lowest level for non-domestic)
```

```
SELECT *  
FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IN  
(SELECT GroupID FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IS NULL)
```

```
# Third level (grouped by YEAR; DLR only)
```

```
SELECT *  
FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IN  
(SELECT GroupID FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IN  
(SELECT GroupID FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IS NULL))  
ORDER BY ParentID, GroupName
```

```
# Fourth level (grouped by LOCATION)
SELECT *
FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IN
(SELECT GroupID FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IN
(SELECT GroupID FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IN
(SELECT GroupID FROM [General_LR4].[dbo].[Groups] WHERE [ParentID] IS NULL))
ORDER BY ParentID, GroupName
```

dbo.Questionnaires

- NLR Main form
- NRS Form 1994-1999 (pre-2000 questionnaire)
- Namibia Form
- Alternative Energies
- Clinic Questionnaire
- Shop Questionnaire
- School Questionnaire

dbo.Questions Questions and their answers are recorded in a one of 3 questions types:

1. long text (blob)
2. short text (char)
3. number

Supporting Tables

Views

Stored Procedures

Auxilliary Data

Photos

Geo-referencing

Data Access

Referencing the Dataset