Data Article

Optimization instances for deterministic and stochastic problems on energy efficient investments planning at the building level

Emilio L. Cano*, Javier M. Moguerza, Antonio Alonso-Ayuso

Department of Computer Science and Statistics, Rey Juan Carlos University, Spain

A R T I C L E   I N F O

Article history:
Received 19 October 2015
Accepted 24 October 2015
Available online 31 October 2015

A B S T R A C T

Optimization instances relate to the input and output data stemming from optimization problems in general. Typically, an optimization problem consists of an objective function to be optimized (either minimized or maximized) and a set of constraints. Thus, objective and constraints are jointly a set of equations in the optimization model. Such equations are a combination of decision variables and known parameters, which are usually related to a set domain. When this combination is a linear combination, we are facing a classical Linear Programming (LP) problem. An optimization instance is related to an optimization model. We refer to that model as the Symbolic Model Specification (SMS) containing all the sets, variables, and parameters symbols and relations. Thus, a whole instance is composed by the SMS, the elements in each set, the data values for all the parameters, and, eventually, the optimal decisions resulting from the optimization solution. This data article contains several optimization instances from a real-world optimization problem relating to investment planning on energy efficient technologies at the building level.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

DOI of original article: http://dx.doi.org/10.1016/j.enbuild.2015.10.020
* Corresponding author.
E-mail addresses: emilio.lopez@urjc.es (E.L. Cano), javier.moguerza@urjc.es (J.M. Moguerza), antonio.alonso@urjc.es (A. Alonso-Ayuso).

http://dx.doi.org/10.1016/j.dib.2015.10.021
2352-3409/© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
### Specifications Table

| Subject area                  | Statistics and Operations Research |
|-------------------------------|-------------------------------------|
| More specific subject area    | Stochastic Optimization             |
| Type of data                  | Binary R data files (.RData) containing R objects. |
| How data was acquired         | Data from research project and simulation |
| Data format                   | Scenario generation from initial values and uncertainty structure |
| Experimental factors          | Not applicable                      |
| Experimental features         | Not applicable                      |
| Data source location          | Not applicable                      |
| Data accessibility            | Data is with this article            |

### Value of the data

- The Symbolic Model Specification (SMS) within the data can be used to generate new instances, with different building configurations and parameter values.
- The optimization instance input data, i.e., parameter values and set elements, can be exported in order to be used by different modeling and/or optimization software.
- The optimization instance output data, i.e., optimal decision variable values, can be used to make different data analysis and visualizations.

### 1. Data

The data shared in this data article consist of six RData files. Five of them contain an R object each whose class is optimInstance. The latter contains a vector of investment costs. Table 1 shows the name and contents of each file. A complete description with details about the model and the instances can be found in [1,2].

The contents of each optimInstance object is in the end a bunch of lists with data.frame R objects, i.e., data tables with the data of each relevant entity in the instance (sets, parameters, variables) and the SMS itself. These objects can be easily managed with the R package optimr as described below.

### Table 1

| File       | Object | Description                                                         |
|------------|--------|---------------------------------------------------------------------|
| EV.instance. | EV instance | Instance of the deterministic problem (Expected Values)          |
| RData      |         |                                                                     |
| RN.instance. | RN instance | Instance of the stochastic, Risk Neutral (RN) problem; $\beta=0$ |
| RData      |         |                                                                     |
| RAL.instance. | RAL instance | Instance of the stochastic, Risk Averse – Low (RAL) problem; $\beta=0.1$ |
| RData      |         |                                                                     |
| RAM.instance. | RAM instance | Instance of the stochastic, Risk Averse – Medium (RAM) problem; $\beta=0.5$ |
| RData      |         |                                                                     |
| RAH.instance. | RAH instance | Instance of the stochastic, Risk Averse – High (RAH) problem; $\beta=0.95$ |
| RData      |         |                                                                     |
| EEVcosts.i. | EEV.costs. | Vector of size 125 with the total investment costs for each scenario of the Expectation of the Expected Values problem (EEV) |
| RData      | i       |                                                                     |
2. Experimental design, materials and methods

The deterministic optimization problem was first described in [1] and then extended to the stochastic version, both risk neutral and with risk management in [2]. We include here a summary of the problem for completeness. This is a decision making problem regarding which investments to make at the beginning of a planning horizon (long term) by a building manager or operator. Such decisions are, in turn, affected by the short-term building performance and decisions, i.e., how to use technologies and market options in the short time. Thus, a holistic model taking into account short- and long-term decisions was built in the frame of the EnRiMa project [3]. The base parameter values for the model were obtained from the project work itself. This base information was basically related to energy demand and price (both purchasing and selling), investment costs, and subsidies. The building under study was constructed in 1975 and it is located in a rural area in the north of Spain. It is set on a plot of approximately 31,363 m², and the ground area of the building is approximately 10,300 m². It has a cellar, a ground floor, and a first floor, though in some areas there is no cellar and/or first floor. Due to its age, the building does not have any ventilation system for assuring the air quality. The heating system provides heating, domestic hot water, and heats the swimming pool of the building. The lighting of the building is done by fluorescent lamps, with a total installed lighting power of 60 kW. In terms of the model, the energy types on the demand side are electricity and heat, including heating and domestic hot water (DHW). On the supply side, they can purchase electricity and natural gas, and sell electricity under regulated tariffs. In addition, a Combined Heat and Power (CHP) unit is used to generate heat and electricity: the primer is used to fulfill the heating demand along with two boilers; the latter is sold to the grid.

From the base values described above (see [1] for the figures), a prediction about the future evolution was made via experts opinion. For the deterministic problem, an average yearly percentage of change for each parameter is needed. The deterministic instance is build just projecting this evolution. Thus, the deterministic instance is composed by one scenario for the 16 years considered. As for the stochastic problem, in addition to average values, knowledge about the distribution of the stochastic parameters is needed. In this case, to generate possible scenarios, an estimation of the evolution standard deviation of each stochastic parameter was also done, and a total of 125 scenarios were generated using the approach in [4], branching the multi-stage scenario tree at years 1, 6, and 11 as described in [2].

In order to manage the instances, we use the R statistical software and programming language [5]. In particular, the optimr package [6] has a class of objects named optimInstance defined to contain optimization instances, including their SMS. See [7] for a complete description of this approach in the framework of Decision Support Systems (DSS). Moreover, the package was also used to generate instances for two-stage stochastic optimization problems in [8].

In what follows, simple examples on how to retrieve and use the data files are provided. For the reader convenience, in addition to the data files enumerated above, the files in Table 2 are provided with this data article. You need to have installed R¹ and the optimr² package. If you intend to generate GAMS files in order to solve instances (either as they are or new/modified ones), you need also the gdxrwrw³ package (version 0.3.0). Assuming that all files are in the R working directory, we can load

---

1 Available at http://cran.r-project.com.
2 Available at http://github.com/emilopezcano/optimr.
3 Available at http://support.gams.com/gdxrwrw:interfacing_gams_and_r.

---

Table 2

| Code files                      | Type      | Description                                |
|--------------------------------|-----------|--------------------------------------------|
| equations.gms                  | GAMS code | Contains the equations of the whole model  |
| functions.R                   | R Script  | Contains convenient functions to get and visualize instance data |

---
any of the instances in Table 1 as follows:

```r
load("SV.instance.RData")

EV.instance$pars$CI
##  i   v  value
## 1 PV-S245 68502 1327.0000
## 2 PV-S245 68503 1281.2890
## 3 PV-S245 68504 1240.8131
## 4 PV-S245 68505 1199.9146
## ... ... ... ...
## 29 CHP-Dachs-G55 68514 2956.0773
## 30 CHP-Dachs-G55 68515 2374.9591
## 31 CHP-Dachs-G55 68516 2206.5854
## 32 CHP-Dachs-G55 68517 2229.7981

names(EV.instance$pars)
## [1] "G" "dG" "dG" "dG" "dG" "dG" "dG" "dG" "dG" "dG" "dG" "dG" "dG" "dG" "dG"
## [15] "B" "dF" "dH" "dH" "dH" "dH" "dH" "dH" "dH" "dH"

tapply(EV.instance$pars$CI$value, EV.instance$pars$CI$i, summary)
## $"CHP-Dachs-G55"
##   Min. 1st Qu.  Median     Mean  3rd Qu.     Max.  
## 2221 2519   2857   2901     2941 3674
## $"PV-S245"
##   Min. 1st Qu. Median  3rd Qu.     Max.  
## 792.2 989.8 1032.0 1084.0 1170.0 1327.0

invest <- merge(EV.instance$pars$x1, EV.instance$pars$PT, by = "v", suffixes = c(".x1", ".x2"))
library(lattice)
barchart(value.x1 ~ as.factor(value.Year) | i, data = invest, horizontal=F)
```
Acknowledgments

This work has been partially supported by the EC’s Seventh Framework Program via the “Energy Efficiency and Risk Management in Public Buildings” (EnRiMa) project (Number 260041). We acknowledge national projects OPTIMOS 3 (MTM2012-36163-C06-06), Content & Intelligence (IPT-2012-0912-430000), PPI (RTC-2015-3580-7), and UNIKO (RTC-2015-3521-7).

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2015.10.021.
References

[1] E.L. Cano, M. Groissböck, J.M. Moguerza, M. Stadler, A strategic optimization model for energy systems planning, Energy Build. 81 (0) (2014) 416–423. http://dx.doi.org/10.1016/j.enbuild.2014.06.030.
[2] E.L. Cano, J.M. Moguerza, A. Alonso-Ayuso, A multi-stage stochastic optimization model for energy systems planning and risk management, Energy Build. 110 (2016) 49–56. http://dx.doi.org/10.1016/j.enbuild.2015.10.020.
[3] EnRiMa consortium, EnRiMa – energy efficiency and risk management, FP7 EC Project, grant agreement 260041 (2010–2012). URL (http://www.enrima-project.eu).
[4] M. Kaut, K. Midthun, A. Werner, A. Tomasgard, L. Hellemo, M. Fodstad, Multi-horizon stochastic programming, Comput. Manag. Sci. 11 (1–2) (2014) 179–193. http://dx.doi.org/10.1007/s10287-013-0182-6.
[5] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2015. URL: (https://www.R-project.org).
[6] E.L. Cano, J.M. Moguerza, optimr: Integrated framework for modeling, representation, solution and reporting of optimization problems, R package v0.0-10, 2014, (https://github.com/emilopezcano/optimr).
[7] E.L. Cano, An Integrated Framework for the Representation and Solution of Stochastic Energy Optimization Problems (Ph.D. thesis), June 2014 Doctor Europeus, Universidad Rey Juan Carlos, Móstoles (Madrid), Thesis advisors: Javier M. Moguerza and Yuriy Ermolieva.
[8] E. Cano, J. Moguerza, T. Ermolieva, Y. Ermoliev, Energy efficiency and risk management in public buildings: strategic model for robust planning, Comput. Manag. Sci. 11 (1–2) (2014) 25–44. http://dx.doi.org/10.1007/s10287-013-0177-3.