Methodology for IEEE-CIS Technical Challenge on Predict + Optimize for Renewable Energy Scheduling

Dr Richard Bean

Centre for Energy Data Innovation

UQ School of Information Technology and Electrical Engineering

17 February 2022, 11am

Room 49-301

https://uqz.zoom.us/j/9357420213

The University of Queensland
Australia

Create Change

Rationale for competition (comp webpage)

- Tackle climate change by decarbonisation of energy production with the use of renewable energy sources such as wind and solar.

Why should we forecast demand and renewable production

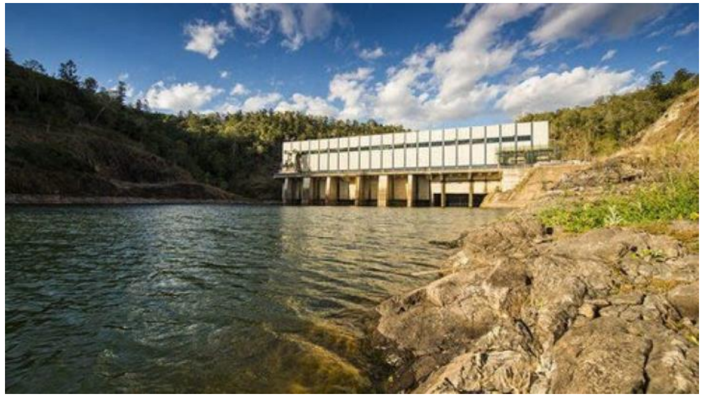
- Renewable energy cannot be produced on demand but the production depends on when wind blows and sun shines
- Storing energy is costly and normally associated with loss of energy
- With more renewable energy in grid, increasingly important to **accurately forecast**
  - **energy demand**
  - **energy production from renewables**

"…the reality is that the technology is not there at the moment to store energy when the sun's not shining or the wind's not blowing…" (April 2018)

FORECASTS ARE THEN USED TO …

- To be able to produce power from on-demand-sources (e.g., gas plants) if needed
- To shed loads and schedule demand to certain times where possible
- To optimally schedule energy storage solutions such as batteries.

Oakey Gas Turbine, Qld



Wivenhoe pumped hydro, near Brisbane

Hornsdale Battery, 220 km N of Adelaide

Abstract

- In particular, a nowadays common setup is a rooftop solar installation and a battery, together with certain demand flexibilities.
- Here, we need to forecast
  - the electricity demand,
  - the renewable energy production,
  - the wholesale electricity price, to be able to then **optimally schedule the charging and discharging of the battery**, and to schedule the schedulable parts of the demand (when to put the washing machine, when to use the pool pump, etc.).
- In this way, we can charge the battery with overproduction of solar energy, and use power from the battery instead of power from the grid when energy prices are highest, as well as schedule demand according to energy availability.

redback technologies

Battery scheduling at Redback (2016-2019)

The University of Queensland
Australia

Using solar and load predictions in battery scheduling at the residential level

Richard Bean
Redback Technologies
Brisbane, Australia

Hina Khan
School of ITEE
The University of Queensland, Australia

8v1 [cs.SY] 26 Oct 2018

***Abstract*—Smart solar inverters can be used to store, monitor and manage a home's solar energy. We describe a smart solar inverter system with battery which can either operate in an automatic mode or receive commands over a network to charge and discharge at a given rate. In order to make battery storage financially viable and advantageous to the consumers, effective battery scheduling algorithms can be employed. Particularly, when time-of-use tariffs are in effect in the region of the inverter, it is possible in some cases to schedule the battery to save money for the individual customer, compared to the "automatic" mode. Hence, this paper presents and evaluates the performance of a novel battery scheduling algorithm for residential consumers of solar energy. The proposed battery scheduling algorithm optimizes the cost of electricity over next 24 hours for residential consumers. The cost minimization is realized by controlling the charging/discharging of battery storage system based on the predictions for load and solar power generation values. The scheduling problem is formulated as a linear programming problem. We performed computer simulations over 83 inverters using several months of hourly load and PV data. The simulation results indicate that key factors affecting the viability of optimization are the tariffs and the PV to Load ratio at each inverter. Depending on the tariff, savings of between 1% and 10% can be expected over the automatic approach. The prediction approach used in this paper is also shown to out-perform basic persistence forecasting approaches. We have also examined the approaches for improving the prediction accuracy and optimization effectiveness.**

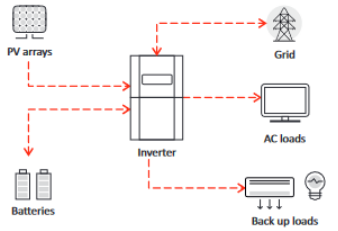
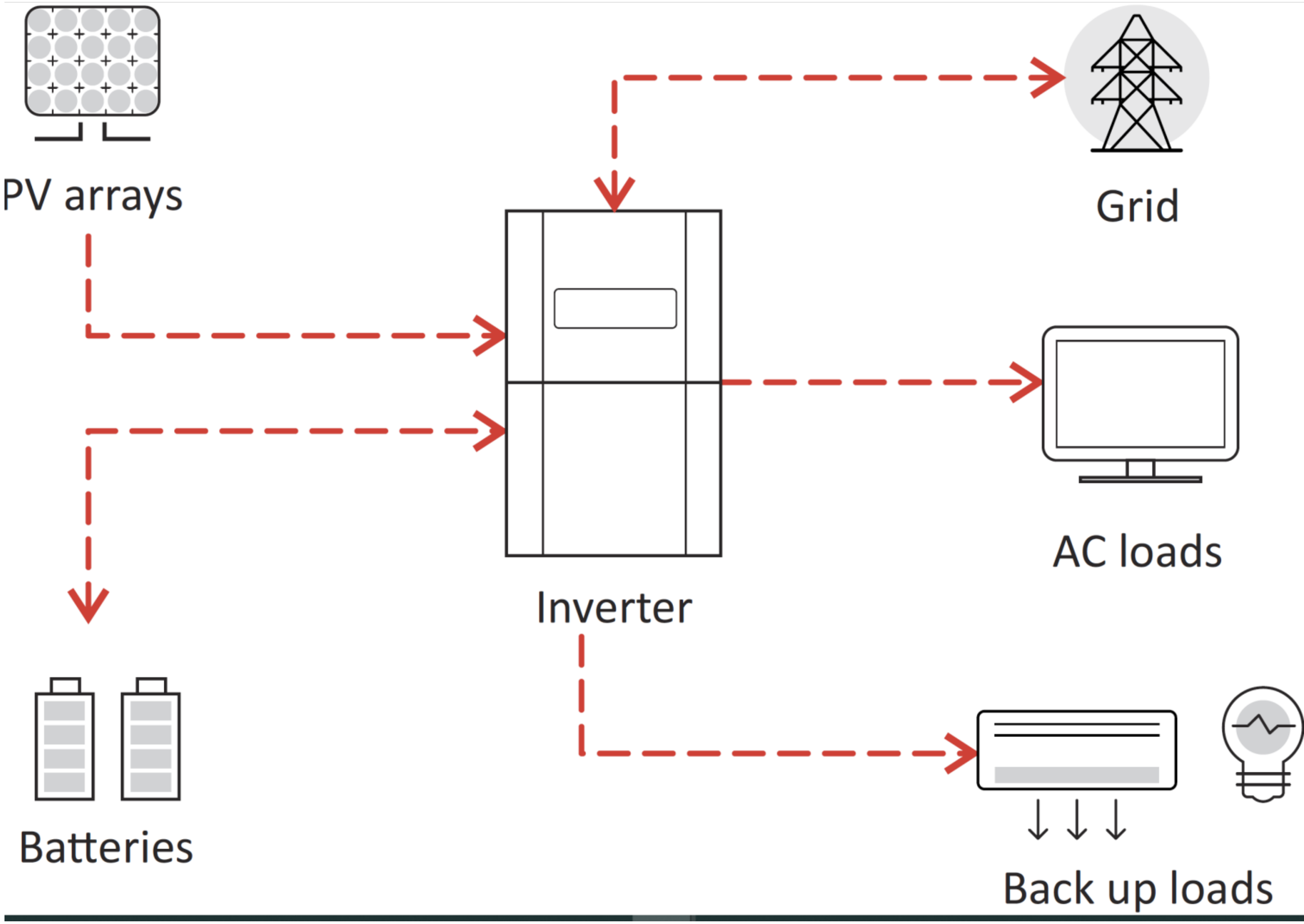


Fig. 1. Schematic of inverter with associated electrical loads, battery and grid connections

of 5 kW. If the inverter is located in Australia with 5 kW of solar panels, this corresponds to an average daily output of between 17.5 kWh (Hobart) and 25 kWh (Alice Springs) ([5]). The batteries attached to the inverter have an associated state of charge value which must be kept between a range of values (e.g. 20%-100%) to avoid adverse effects or battery failure. The information related to the load, PV, state of



THE PREDICT+OPTIMIZE PROBLEM

- Researchers from the IEEE Computational Intelligence Society (IEEE-CIS) want to improve solutions to this complex problem of “predict + optimize”, in this particular application of scheduling in the context of renewable energy.

IEEE Computational Intelligence Society

Smart “Predict, then Optimize”

Adam N. Elmachtoub,[a] Paul Grigas[b]

[a] Department of Industrial Engineering and Operations Research and Data Science Institute, Columbia University, New York, New York 10027; [b] Department of Industrial Engineering and Operations Research, University of California, Berkeley, Berkeley, California 94720

Contact: adam@ieor.columbia.edu, https://orcid.org/0000-0003-0729-4999 (ANE); pgrigas@berkeley.edu, https://orcid.org/0000-0002-5617-1058 (PG)

Received: December 14, 2017
Revised: July 18, 2019; July 8, 2020
Accepted: November 3, 2020
Published Online in Articles in Advance: March 12, 2021

https://doi.org/10.1287/mnsc.2020.3922

Copyright: © 2021 INFORMS

Abstract. Many real-world analytics problems involve two significant challenges: prediction and optimization. Because of the typically complex nature of each challenge, the standard paradigm is predict-then-optimize. By and large, machine learning tools are intended to minimize prediction error and do not account for how the predictions will be used in the downstream optimization problem. In contrast, we propose a new and very general framework, called Smart “Predict, then Optimize” (SPO), which directly leverages the optimization problem structure—that is, its objective and constraints—for designing better prediction models. A key component of our framework is the SPO loss function, which measures the decision error induced by a prediction. Training a prediction model with respect to the SPO loss is computationally challenging, and, thus, we derive, using duality theory, a convex surrogate loss function, which we call the SPO+ loss. Most importantly, we prove that the SPO+ loss is statistically consistent with respect to the SPO loss under mild conditions. Our SPO+ loss function can tractably handle any polyhedral, convex, or even mixed-integer optimization problem with a linear objective. Numerical experiments on shortest-path and portfolio-optimization problems show that the SPO framework can lead to significant improvement under the predict-then-optimize paradigm, in particular, when the prediction model being trained is misspecified. We find that linear models trained using SPO+ loss tend to dominate random-forest algorithms, even when the ground truth is highly nonlinear.

Smart Predict-and-Optimize for Hard Combinatorial Optimization Problems

Jayanta Mandi,[1] Emir Demirović,[2] Peter. J Stuckey,[2] Tias Guns[1]
[1] Data Analytics Laboratory, Vrije Universiteit Brussel {jayanta.mandi,tias.guns}@vub.be
[2] University of Melbourne {emir.demirovic,pstuckey}@unimelb.edu.au

Abstract

Combinatorial optimization assumes that all parameters of the optimization problem, e.g. the weights in the objective function, are fixed. Often, these weights are mere estimates and increasingly machine learning techniques are used to for their estimation. Recently, Smart Predict and Optimize (SPO) has been proposed for problems with a linear objective function over the predictions, more specifically linear programming problems. It takes the regret of the predictions on the linear problem into account, by repeatedly solving it during learning. We investigate the use of SPO to solve more realistic discrete optimization problems. The main challenge is the repeated solving of the optimization problem. To this end, we investigate ways to relax the problem as well as warm-starting the learning and the solving. Our results show that even for discrete problems it often suffices to train by solving the relaxation in the SPO loss. Furthermore, this approach outperforms the state-of-the-art approach of Wilder, Dilkina, and Tambe. We experiment with weighted knapsack problems as well as complex scheduling problems, and show for the first time that a predict-and-optimize approach can successfully be used on large-scale combinatorial optimization problems.

execution and predictive ML models can be used for estimation of those parameters from historical data. For instance, Cohen et al. first predicted future demand of products using an ML model and then use the predicted demand to compute the optimal promotion pricing scheme over the products through non-linear integer programming.

When predictive ML is followed by optimization, it is often assumed that improvements in the quality of the predictions (with respect to some suitable evaluation metric) will result in better optimization outcomes. However, ML models make errors and the impact of prediction errors is not uniform throughout the underlying solution space, for example, overestimating the highest-valued prediction might not change a maximization problem outcome, while underestimating it can. Hence, a better prediction model may not ensure a better outcome in the optimization stage. In this regard, Ifrim, O’Sullivan, and Simonis (2012) experienced that a better predictive model does not always translate to optimized energy-saving schedules.

The alternative is to take the effect of the errors on the optimization outcome into account during learning. In the

Monash Microgrid

Smart Energy City

We're building an on-site microgrid at Clayton campus. In the next 10 years, we'll eliminate our dependence on coal-fired energy sources.

Our microgrid will be versatile enough to receive and store energy from various renewable energy sources. We'll be able to control when and how we use our energy, which means we can reduce demand and strain on the network during peak times.

Our microgrid will also help stabilise the wider grid, making it more resilient. This will benefit the broader community, especially during extreme weather events.

Abstract

- From a machine learning point of view, the provided data poses an interesting time series prediction problem, with
  - *multiple seasonality,*
  - *use of external data sources (weather, electricity price)*
  - *the opportunity for cross-learning across time series on two different prediction problems (energy demand and solar production).*
- Then, from an optimization point of view, uncertainty in the inputs needs to be addressed together with a couple of constraints, to achieve a good solution.
- RB note: need to use R, Python and Java quite a bit
  - R for forecasting
  - Python for optimisation (works best with Gurobi)
  - Java for the schedule evaluation program
- *If successful, you will not only help making renewable energy more reliable and affordable, thus playing your part in the fight against climate change, but the proposed technical challenge may be applicable in many other fields facing similar problems of optimal decision-making under uncertain predictions*

Requirements

- Develop an
  - optimal battery schedule
  - an optimal lecture schedule – *recurring activities and once-off activities*

based on predictions of future values of energy demand and production.

- With input data:
  - **Energy consumption data** every 15 minutes from 6 buildings on the Monash Clayton campus, to September 2020
  - **Solar production data** every 15 minutes from 6 rooftop solar installations from the Clayton campus, to September 2020
  - **Daily weather data** from Australian Bureau of Meteorology (daily solar, max and min temperatures, rain) – three sites
  - **Hourly weather data** from European Centre for Medium Range Weather Forecasting (ECMWF) – one point (from 11 August 2021)
  - **Electricity price data** from Australian Energy Market Operator (30 minute)

ECMWF EUROPEAN CENTRE FOR MEDIUM RANGE WEATHER FORECASTS
BOM
MONASH University
AEMO AUSTRALIAN ENERGY MARKET OPERATOR

Phase 1

- Optimally schedule a battery and timetabled activities (lectures) for the month of October 2020.
- In real life, the battery scheduling would usually happen on a daily basis, with day-ahead forecasting.
- For the competition the test set cannot be disclosed, so that a whole month needs to be forecasted.
- However, with the availability of weather data, this task is still close to the real world application, with the assumption of having perfect 1-day-ahead weather forecasting and having perfect electricity price forecasting.
- Phase 1 - public leaderboard where participants submit forecasts and the leaderboard shows the evaluation of the forecasts (MASE error rate and cost)

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

LEADERBOARD

Title	Submitted by	Mean MASE	Energy Cost	Last Submission
nov3x	Richard Bean	0.646022	335107.248311086	2021-11-03 08:59:38
final_submission	Rasul Esmaeilbeigi	0.744052	328359.202909373	2021-11-01 22:08:38
SZU-PolyU-Team	Qingling ZHU	0.774996	342810.01676205	2021-11-03 00:12:18
final_submission_	Xu YaoJian	0.774996	342838.193792926	2021-11-02 00:30:26
EVERGi team final submission	Julian Ruddick	0.807299	340725.940843299	2021-11-02 13:37:39
Base solution	Akylas Stratigakos	0.847391	363168.136647165	2021-11-03 03:47:14
Submission19	Steffen Limmer	0.855737	339160.427284564	2021-10-22 21:53:30
Standard	Nils Einecke	0.903562	589356.97587725	2021-10-16 01:29:10
FRESNOB	Rui YUAN	1.002641	395581.33876972	2021-10-28 21:25:08
Final_submission	Tomas Ochoa	1.012309	589356.97587725	2021-11-02 12:04:19

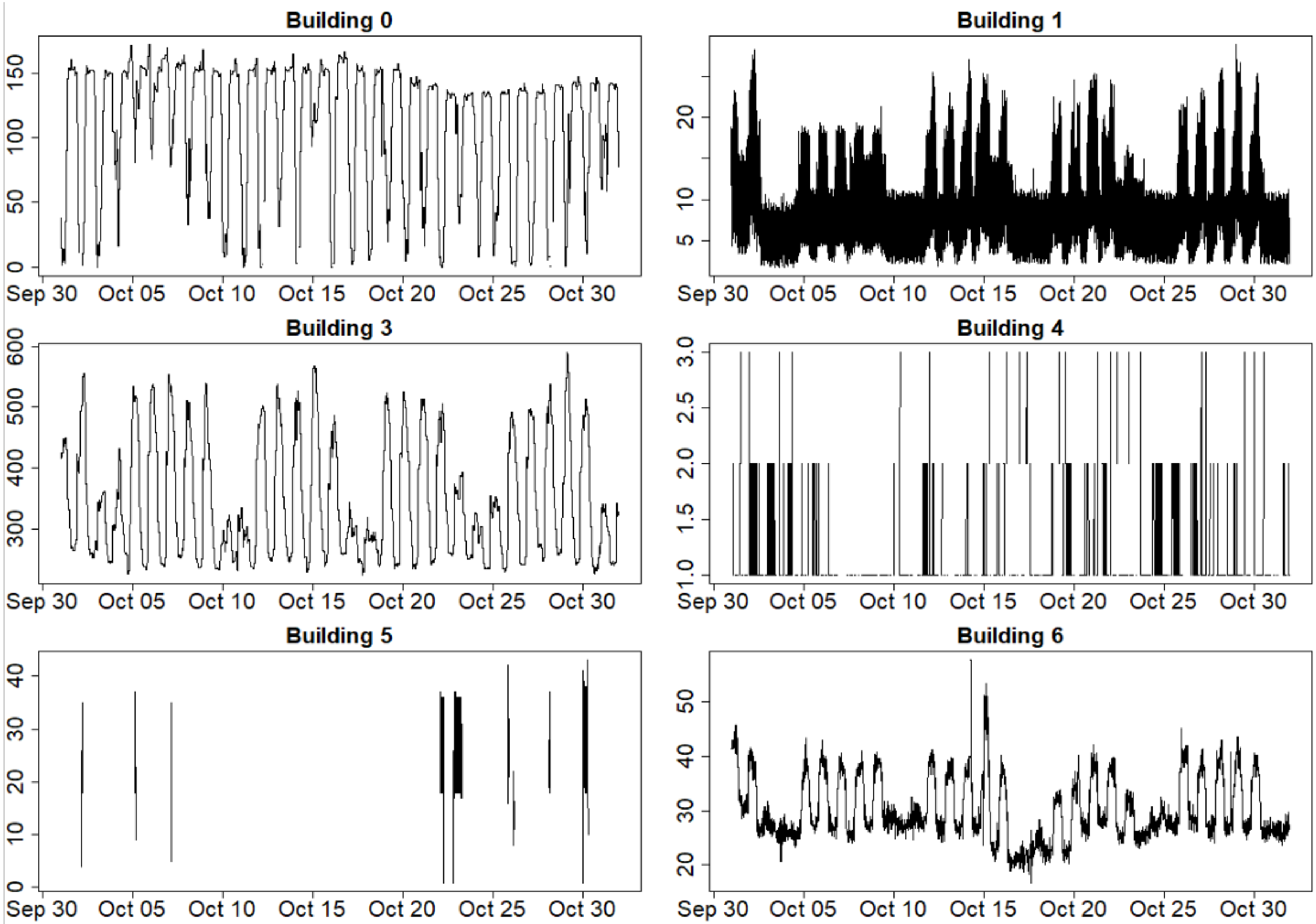
Requirements and Assessment

Phase 2

- Data for October 2020 is released to the participants, and they are now asked to perform the same forecasting and optimisation exercise for November 2020.
- Now, only minimal feedback is provided to the participants about the quality of their submissions. Solely Phase 2 of the competition is relevant to determine competition winners and prizes.
- The 3 main competition prizes will be awarded to the schedules that lead to the **lowest cost on the Phase 2 test set. ($US7000, $US5000, $US3000)**
- An additional prize will be awarded to the team that achieves the most accurate forecasts on Phase 2. **($US2000)**

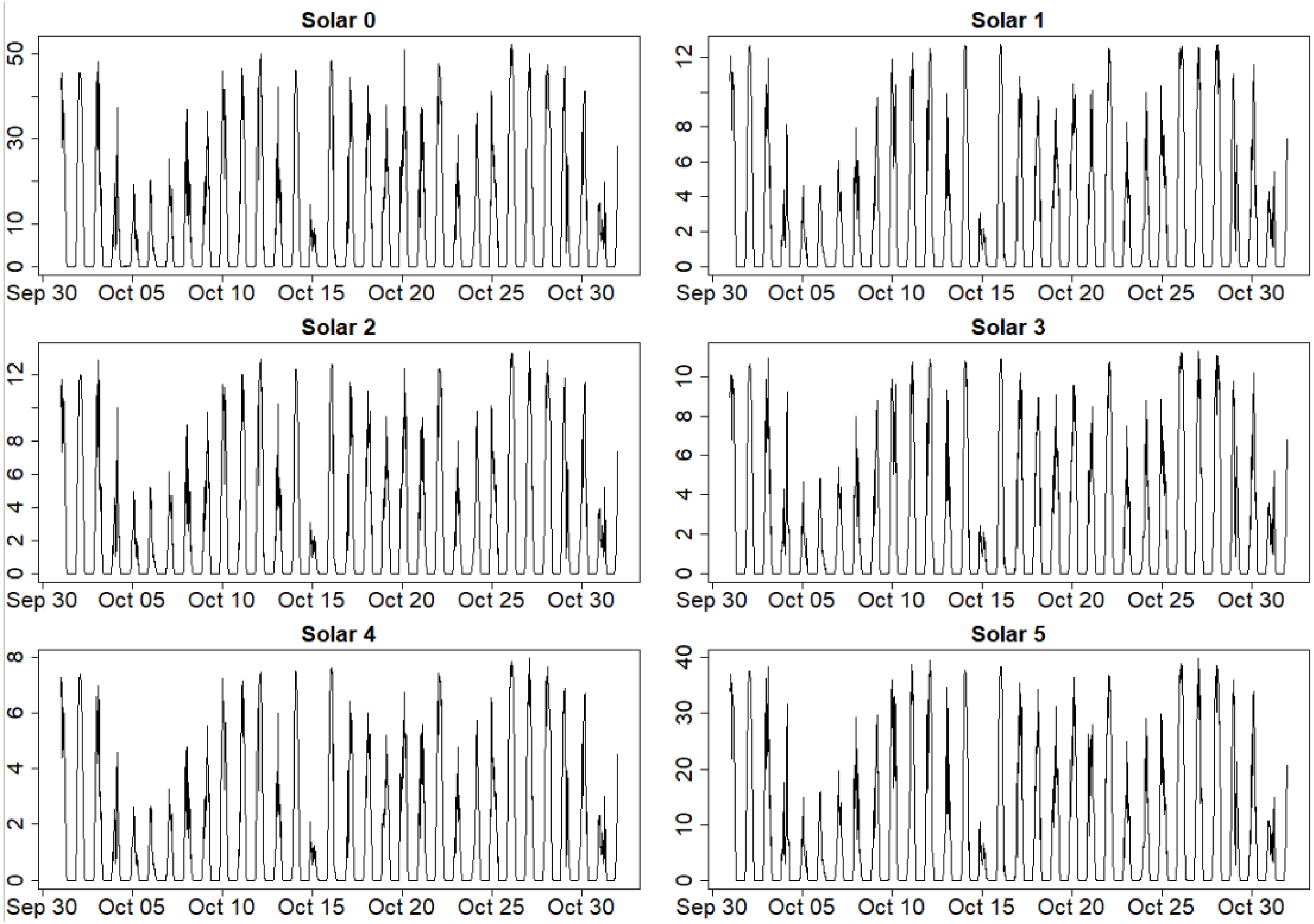
THE UNIVERSITY OF QUEENSLAND AUSTRALIA

BUILDING DATA IN OCTOBER 2020



The University of Queensland
Australia

Solar Data in October 2020



Objective function

Objective

For a feasible schedule, we compute the objective value in terms of the cost of the schedule, which is to be minimized. The cost of the schedule depends on three parts:

- The total energy cost computed against the wholesale price e_t,
- The peak load tariff taken over the whole month,
- The value of the once-off activities scheduled $d_i \in \{0, 1\}$, whether in or out of office $o_i \in \{0, 1\}$.

The objective O is computed as follows:

$$O = \sum_t \frac{0.25 l_t e_t}{1000} + 0.005(\max_t l_t)^2 - \sum_{a_i} (d_i \cdot (\text{value}_i - o_i \text{penalty}_i))$$

Becomes an MIQP (Mixed Integer Quadratic Program)

Scheduling file Examples

ppoi 6 6 2 200 100
b 0 13 2
b 1 1 1
b 3 5 5
b 4 4 3
b 5 7 3
b 6 1 1
s 0 0
s 1 1
s 2 3
s 3 4
s 4 5
s 5 6
c 0 1 150 75 0.85
c 1 3 420 60 0.60
r 0 1 S 31 5 5 22 86 91 98 137
r 1 2 S 51 6 5 21 52 63 126 149
...
a 0 2 S 54 7 30 28 3 11 34 93
a 1 1 S 51 2 4 4 5 12 15 42 63 69
...

Predict Plus Optimize Instance, **6** buildings/solar, **2** batteries, **200** recurring activities, **100** once-off activities

Building **0** has **13** small + **2** large rooms

Solar **0** is connected to Building **0** (irrelevant)

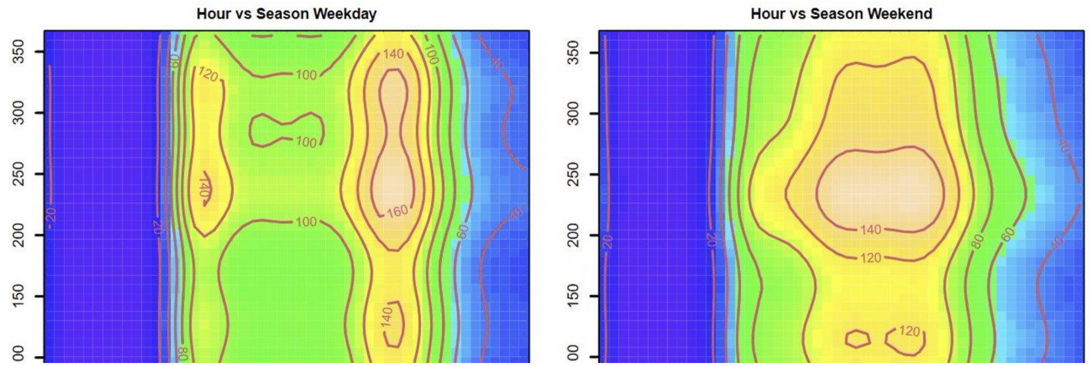
Battery **0** is **150** kWh, discharges at **75** kW, round trip efficiency **85%**

Recurring activity **0** requires **1** small room, uses **31** kW, is **5** periods long, must occur on weekdays between 9am-5pm, and recurring activities **22, 86, 91, 98, 137** must occur on *earlier* weekdays

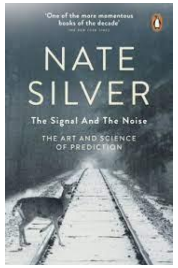
(optional) Once-off activity **0** requires **2** small rooms, uses **54** kW, **7** periods long, bonus **$30** if scheduled into working hours, penalty **$28** if scheduled outside working hours, once-off activities **11, 34, 93** must occur on *earlier* days of month (if included)

ppoi 6 6 2 200 100
sched 200 100
r 0 91 3 // period + buildings
r 1 492 2 2
...
a 0 2113 1 5 // period + buildings
a 1 1048 1
...
c 0 0 0
c 0 1 2
c 0 2 2
c 0 3 2
c 0 4 2 // hourly charging instructions

Personal Background



- Ph.D. mathematics (UQ 2001, combinatorics)
- ROAM Consulting (now EY) 2007-2012
- AEMO (Australian Energy Market Operator) 2013
- Redback Technologies (2016-2019)
- University of Queensland (2019-2022)
  - Centre for Energy Data Innovation https://cedi.uqcloud.net/
- Australian and NZ Electricity Market regional and sub-regional demand at ROAM/AEMO – “macro” forecasting
- Individual buildings/solar/distribution transformers at Redback/UQ from inverter or smart meter data - “micro” forecasting



- Cybersecurity – localization of houses with ERA5 solar / load data
- ROAM – simple quadratic programming for modelling NEM bidding (COIN-OR)
- Battery/inverter scheduling at Redback – linear programming
- Combinatorics / graph theory – 0-1 integer programming (CPLEX, BonsaiG, COIN-OR, Gurobi)
- Bike sharing forecasting with GAMs and ERA5 data emph. explainability >> error rate ~ energy
- Classical cryptanalysis – pattern recognition (closely connected)

WHAT WORKED IN THE PAST?

- Has anyone done this kind of thing before? What worked?
- Global Energy Forecasting Competition
  - GEFCOM – three editions 2012, 2014 and 2017
    - 2012 – hierarchical load forecasting and wind power
    - 2014 – hierarchical load, wind energy, price, and solar
    - 2017 – hierarchical probabilistic load
- Used to inform Redback model (2016-2019)

ELSEVIER

International Journal of Forecasting
Volume 30, Issue 2, April–June 2014, Pages 357-363

Global Energy Forecasting Competition 2012

Tao Hong [a], Pierre Pinson [b], Shu Fan [c]

ELSEVIER

International Journal of Forecasting
Volume 32, Issue 3, July–September 2016, Pages 896-913

Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond

Tao Hong [a], Pierre Pinson [b], Shu Fan [c], Hamidreza Zareipour [d], Alberto Troccoli [e], Rob J. Hyndman [c]

ELSEVIER

International Journal of Forecasting
Volume 35, Issue 4, October–December 2019, Pages 1389-1399

Global energy forecasting competition 2017: Hierarchical probabilistic load forecasting

Tao Hong [a], Jingrui Xie [b], Jonathan Black [c]

GEFCOM KAGGLE 2012

Table 2
Summary of methods in the hierarchical load forecasting track.

Kaggle ID	Techniques	Data cleansing	Weather station selection	Holiday effect	Temperature forecast	Ensemble forecasting
CountingLab	MLR, singular value decomposition	Yes	11 models corresponding to the 11 weather stations were built	Yes	Using the average temperature of the same hour from similar days in the previous years	Combine forecasts from the 5-best fitted models
James Lloyd	Gradient boosting machines, Gaussian process regression, MLR	Not discussed	Temperatures from all stations were used	No	Estimating the smooth trend and daily periodicity of temperature separately	Combine forecasts from three models
Tololo	Semi-parametric regression, with B-splines or cubic regression splines as smooth function	Not discussed	A stepwise procedure was used for each zone to select the station that minimized forecasting error on a test set	Yes	Not discussed	No
TinTin	Nonparametric additive models with P-spline, component-wise gradient boosting	Yes	A testing week (the last week of the available data) was used to determine the station for each zone	Yes	Using the average temperatures at the same period across the previous years	No
Quadrivio	MLR	Yes	Load was fitted to temperature at each station separately, and the best three were used for each zone	No	Averaging the temperatures during the same days from previous years	No
Chaotic Experiments	Random forest, geometric Brownian motion models	Not discussed	Not discussed	Yes	Not discussed	Combine forecasts from three models
Andrew L.	additive model, spline, PCA	Not discussed	The first component of PCA was used as temperature variable for each hour	No	Using a generalized additive model	No
NHH	Wavelet decomposition, mutual information, neural networks	Not discussed	Temperatures from all stations were considered as input candidate	No	Not discussed	No
TheJellyTeam	Neural networks	Not discussed	Temperatures from all stations were considered	Yes	Using the mean of the same period from the previous years	No
Shooters Touch	Regression models and neural network	No	Weighted average of up to 3 stations, selected based on	Yes	Not discussed	No

Public Private

The private leaderboard is calculated with approximately 75% of the test data.
This competition has completed. This leaderboard reflects the final standings.

Prize Winners

#	△	Team	Members	Score	Entries	Last	Code
1	▲ 1	Tiberius Data Mining		60777.02849	4	9Y	
2	▲ 3	CountingLab		67214.64846	33	9Y	
3	▲ 3	James Lloyd		71467.03127	52	9Y	
4	▼ 3	Tololo		71779.61027	39	9Y	
5	▼ 2	TinTin		73307.05957	42	9Y	
6	▲ 3	yuenking		76904.61627	15	9Y	
7	—	Quadrivio		78195.94722	29	9Y	
8	▲ 2	Luxtorpeda		79791.43318	17	9Y	
9	▲ 5	Hugh Perkins		79850.76751	17	9Y	
10	▲ 15	Chaotic Experiments		80762.58779	19	9Y	

Congrats to CRW!

Posted in global-energy-forecasting-competition-2012-load-forecasting 9 years ago

Congrats to CRW! 55969.74840. Nice!

GEFCOM 2014 AND 2017

Table 7
Summary of final match methodologies.

Team	Ranking	Data cleansing	Weather station selection	Feature engineering: Load	Feature engineering: Weather	Feature engineering: Calendar	Temperature scenarios	Modeling techniques	Forecast combine	Meter grouping / hierarchy information
QUINKAN	1	Yes	Yes	No	Yes	Yes	No	Quantile regression and generalized additive model	No	Yes
dmlab	3	No / No discus-sion	No	No	Yes	Yes	Shifted-date (no discussion on the value k and n)	Quantile gradient boosted regression trees		No
Orbuculum	4	Yes	No	Yes (transformation)	No	Yes	No	Gradient boosting machine, quantile random forest, and naive forecast	Yes	No
GeertScholma	5	Yes	Yes	No	Yes	Yes	Shifted-date ($k = 7$, $n = 2$)	Multiple linear regression and autoregression	Yes	No
Cassandra	6	Yes	No	No	No	Yes	No	Neural network-based quantile forecasting model, and time series models	No	Yes
It Can Be Done	7	No / No discus-sion	No	No (transformation)	No	Yes		Gradient boosting machine, quantile random forest	Yes	No
Sim-ple_but_good	8	Yes	No	No	No	No	No	[illegible]	[illegible]	No
UC3M	9	Yes	No	No	Yes	Yes	No	Linear regression model, factorial model, profiling	Yes	Yes

Table 11
Summary of the methods used by the top five teams in the solar track of GEFCom2014.

Team	Parametric/nonparametric	Forecasting models and techniques	Generalization ability (preventing overfitting)	Input variables and features (most important)	Offsite information
Gang-gang	Nonparametric	Gradient Boosting (GB) and k-Nearest Neighbours (k-NN)	Cross-validation	Clear sky model, as well as all variables provided	Yes
dmlab	Nonparametric	Quantile Regression Forest (QRF) and Gradient Boosting Decision Trees (GBDT)	Cross validation	Variables provided, time of day and of year, differentiated variables (for the accumulated fields)	No
C3 Green Team	Nonparametric	Multiple Quantile Regression (MQR)	Feature selection algorithm, regularization when estimating, and cross-validation	Wealth of features based on all input variables, time of day and time of year	Yes
Giuseppe Casalicchio	Nonparametric	Quantile Regression and Quantile Regression Forest (QRF)	Lasso penalization	Wealth of features based on all input variables, considering lagging, smoothing, and combination	Yes
UT_Argonne	Nonparametric	Ensemble of Random Forest (RF), Gradient Boosting Machines (GBM) and Support Vector Machines (SVM)	Training data selection	Wealth of features based on all input variables, considering time shifting, integration, etc.	Yes

The Recency Effect

International Journal of Forecasting
Volume 32, Issue 3, July–September 2016, Pages 585-597

Electric load forecasting with recency effect: A big data approach

Pu Wang [a], Bidong Liu [b], Tao Hong [b]

Show more

Add to Mendeley Share Cite

https://doi.org/10.1016/j.ijforecast.2015.09.006 Get rights and content

Abstract

Temperature plays a key role in driving the electricity demand. We adopt the "recency effect", a term drawn from psychology, to represent the fact that the electricity demand is affected by the temperatures of the preceding hours. In the load forecasting literature, the temperature variables are often constructed in the form of lagged hourly temperatures and moving average temperatures. In the past, computing power has limited the amount of temperature variables that can be used in a load forecasting model. In this paper, we present a comprehensive study to model the recency effect using a big data approach. We take advantage of modern computing power to answer a fundamental question: *how many lagged hourly temperatures and/or moving average temperatures are needed in a regression model in order to capture the recency effect fully without compromising the forecasting accuracy?* Using a case study based on data from the load forecasting track of the Global Energy Forecasting Competition 2012, we first demonstrate that a model with the recency effect outperforms its counterpart (a.k.a. Tao's Vanilla Benchmark Model) by 18% to 21% for forecasting the load series at the top (aggregated) level. We then model the recency effect in order to customize load forecasting models at the bottom level of a geographic hierarchy, again showing a superiority over the benchmark model of 12% to 15% on average. Finally, we discuss four different implementations of the

European Journal of Operational Research
Volume 251, Issue 2, 1 June 2016, Pages 522-530

Production, Manufacturing and Logistics

Forecasting day-ahead electricity load using a multiple equation time series approach

A.E. Clements, A.S. Hurn, Z. Li

Show more

Add to Mendeley Share Cite

https://doi.org/10.1016/j.ejor.2015.12.030 Get rights and content

Highlights

- A multiple equation time series model is built to forecast electricity load.
- Interactions in seasonal patterns are given special prominence.
- The model is easily estimated by repeated application of ordinary least squares.
- The model achieves a mean absolute percentage error of 1.36% in day-ahead forecasting over 11 years.
- Forecasts outperform the industry standard by about a third.

KEY STEPS - FORECAST

- The most important step! Reproducible code
- Find the approach that gives the lowest MASE for each time series on phase 1
- R script change PHASE value to 2 and rerun

```r
rm(list=ls())

PHASE <- 1

FLIST <- c("phase_1_data.tsf","phase_2_data.tsf")

PDAY <- c(31,30)
PMONTH <- c(10,11)

DAYS <- PDAY[PHASE]
PERIODS <- DAYS * 24 * 4
HOURS <- DAYS * 24
HOUR1 <- HOURS - 1
FIRSTPERIOD <- paste("2020-",PMONTH[PHASE],"-01 00:00:00",sep="
```

Data Replication & Reproducibility

PERSPECTIVE

Reproducible Research in Computational Science

Roger D. Peng

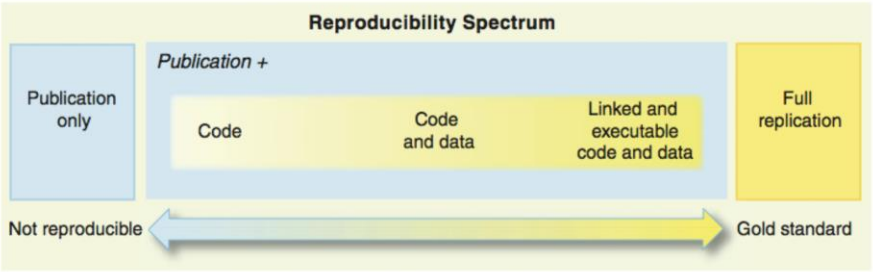


Fig. 1. The spectrum of reproducibility.

1226 2 DECEMBER 2011 VOL 334 SCIENCE www.sciencemag.org

PARADOX?

If you think the competition is just pure skill you won't enter Phase 2 but if you think luck is involved you'll definitely just run your Phase 1 model on Phase 2. i.e. it's better for the competitors and competition organizers if they believe luck is involved.

RANDOM FORESTS

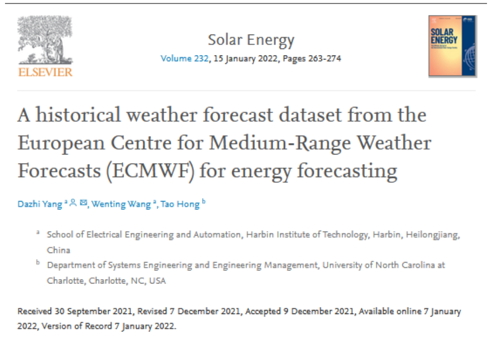
Led Zeppelin

The Cure

Thom Pace

PHASE 1 FORECASTING

- Weather data provided was originally daily, but 15 minute forecasts required – daily max/min temp, rain, solar at 3 sites
- Initially used **GAMs (generalized additive models)** which is great for visualizations and explainability, especially for bike-sharing demand
- Switched to **random forests** as the competition was only about performance
- Inspired by GEFCOM – 12 ECMWF variables used, here 8; same as Redback work
- https://apps.ecmwf.int/codes/grib/param-db 5,656 parameters


Solar Energy
Volume 232, 15 January 2022, Pages 263-274

A historical weather forecast dataset from the European Centre for Medium-Range Weather Forecasts (ECMWF) for energy forecasting

Dazhi Yang [a], Wenting Wang [a], Tao Hong [b]

[a] School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, Heilongjiang, China
[b] Department of Systems Engineering and Engineering Management, University of North Carolina at Charlotte, Charlotte, NC, USA

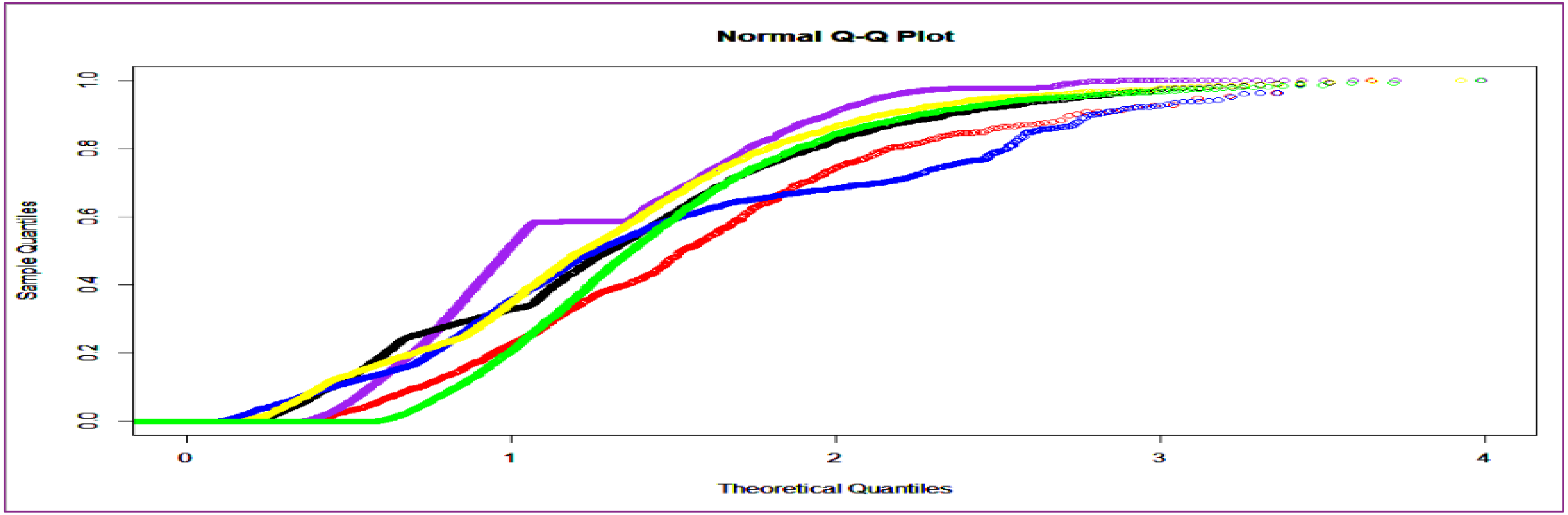
Received 30 September 2021, Revised 7 December 2021, Accepted 9 December 2021, Available online 7 January 2022, Version of Record 7 January 2022.

Parameter ID	078	079	134	137	151	157	164	165	166	167	168	169	175	178	186	206	212	228	243	228021	228022	228129	#
Short Name	TCLW	TCIW	SP	TCWV	MSL	R	TCC	10U	10V	2T	2D	SSRD	STRD	TSR	LCC	TCO3	TISR	TP	FAL	FDIR	CDIR	SSRDC	
GEFCOM2014-S	x	x	x			x	x	x	x	x		x	x	x				x					12
IEEE Tech Challenge 2021					x	x	x	x		x	x	x	x										8
Espejo et al 2019										x		x								x	x	x	5
Yang et al 2022 (N. America/Europe)			x	x				x	x	x	x	x	x		x	x	x	x	x	x			14
OEMOF Feedinlib / PVLIB								x	x	x		x								x			5
Greco-Project PVCompare				x				x	x	x		x								x			6

Phase 1 Forecasting

- Random forest libraries have the useful "variable importance" ranking
- Lagging and leading data (up to 3 hours)
- Day of week, time of day, Julian date (Fourier values)
- "Days before" temperature effect
- Data cleaning – missing data

"Everyone wants to do the model work, not the data work": Data Cascades in High-Stakes AI



Phase 1 & 2 Challenges – data issues

- 1 October – participant found bug in optimisation evaluation code – adding solar power onto net load, not subtracting it. Competition phase 1 extended by a week
- 4 October – seeing this, I wrote to organizers. Time zone issues: optimisation was happening 9am-5pm UTC i.e. 8pm-4am Melbourne time. Like bug bounties – perhaps there should be rewards for finding these.
- 21 October – several bugs in optimisation evaluation; “recurring activities”
- 25-29 October – leaderboard recalculated
- Repeated leaderboard outages, spammers
- Missing data points treated as “zero” values – random effects

KEY STEPS

- Quantile regression forest – forecast **median** to minimise MAE
- (i.e. sum of deviations from actual value)
- Most important parameter to tune – *"mtry"*
- Training against individual phase 1 time series (without overfitting)
- Each hour gets 4 random forests (each quarter hour)
- Choosing building start months of 2020 (Building 0,1,3,6)
- Removing building outliers
- Choosing solar start months (Solar1 has some cumulative data)
- Predictor variables: ECMWF vars lead/lag 3h, day of week, day of year etc
- Public holiday – 23 October Grand Final holiday excluded from training
- Building4 and Building5 set to median values of Oct 2020 (1 and 19 kW)

- Forecast groups of buildings and solar together with normalization (critical, but mentioned by organizers *"cross-learning across time series"*)
- Using BOM daily and ECMWF 1 hour data together (critical … is this surprising?)
- Solar0 and Solar5 thresholding hugely improves MASE (critical)

Predictor Variables for Building

wh	temperature	dewpoint	wind	MSLP	R	SSRD	STRD	TCC	b8	b9	b10	t1	t2	t3
6	9.7	7.3	5.0	1005	0.85	59.4	338.4	1.00	4.5	2.8	2.6	9.7	9.8	10.1
tf	tf2	tf3	t24	t48	t72	s1	s2	s3	sf	sf2	sf3	st1	st2	st3
10.2	10.7	10.6	13.8	13.4	12.3	15.4	1.1	0	182	244	152	344	347	345
stf	stf2	stf3	w1	w2	w3	wf1	wf2	wf3	d1	d2	d3	df1	df2	df3
319	328	334	4.6	4.4	6.3	5.7	5.9	5.8	7.6	7.7	7.9	7.3	7.3	7.8
rh1	rh2	rh3	rhf1	rhf2	rhf3	cc1	cc2	cc3	ccf1	ccf2	ccf3	mslp1	mslpf	hr1
0.87	0.87	0.86	0.82	0.8	0.83	1.00	1.00	1.00	0.97	0.99	0.99	1005	1005	0
sin hr	cos hr	wd	wd1	wd2	wdx0	wdx1	wdx2	wdx3	wdx4	wdx5	wdx6	sin day	cos day	x1
0	1	0	1	0	0	1	0	0	0	0	0	0.50	-0.87	0.63

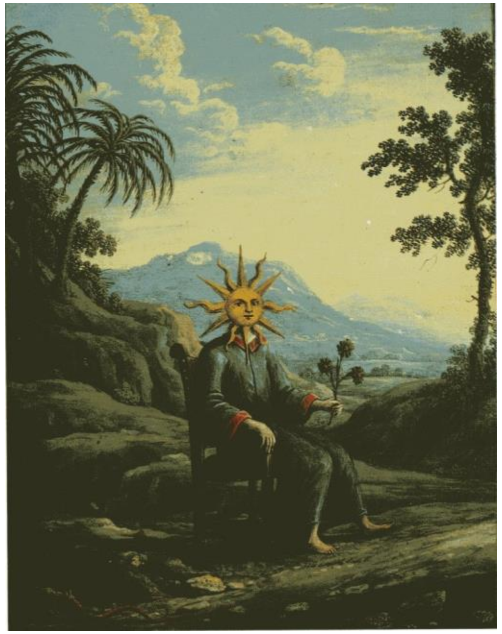
PREDICTOR VARIABLES FOR SOLAR

wh	temperature	MSLP	SSRD	STRD	TCC	b8	b9	b10	t1
0	12	1018	203	321	0.37	8.8	8	6.6	12
t2	t3	tf	tf2	tf3	s1	s2	s3	sf	sf2
12	12	12	11	10	221	172	173	145	54
sf3	st1	st2	st3	stf	stf2	stf3	cc1	cc2	cc3
0.66	318	332	347	311	308	303	0.56	0.81	0.94
ccf1	ccf2	ccf3	mslp1	mslpf1	sin hr	cos hr	sin day	cos day	x1
0.48	0.58	0.74	1017	1018	0.97	0.26	0.66	-0.76	0

MASE Expectation

Almost all my forecast MASE improvement came after Phase 1 data was released.
Obviously lots of room to improve Solar0/5 still

"Progress usually comes from many small improvements; a change of 1% can be a reason to break out the champagne."



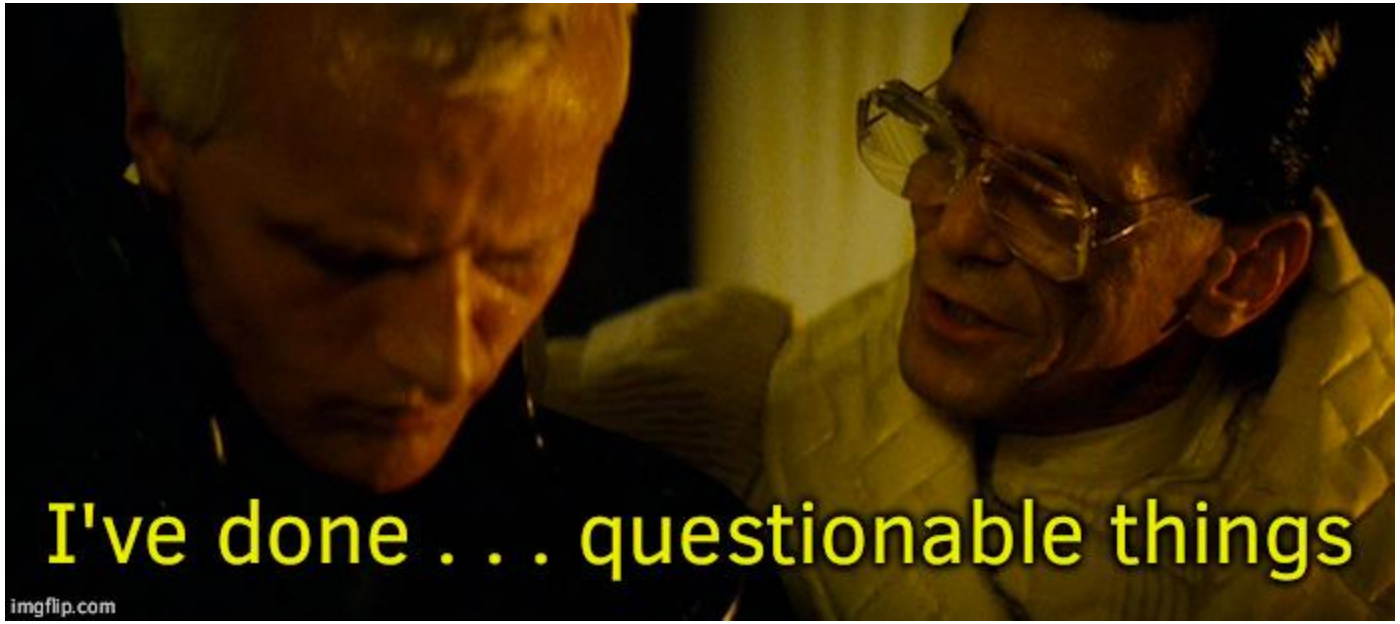
Case	MASE Phase 1	MASE after tuning
Building0	0.4301	0.3859
Building1	0.6115	0.4251
Building3	0.3310	0.2913
Building4	0.5637	0.5637
Building5	1.0370	0.8383
Building6	0.7676	0.7336
Solar0	0.8479	0.6558
Solar1	0.4619	0.3619
Solar2	0.5251	0.4139
Solar3	0.5910	0.4990
Solar4	0.5624	0.4219
Solar5	0.8559	0.6092
Mean	**0.6320**	**0.5166**

ERA-5 DATA PROVIDED

Questioning how provided ERA-5 data was derived.

Inverse distance weighting (exponent 2) of four ERA-5 points (0.25 degrees).

Lots of subtleties e.g. exponent choice in IDW, losing wind speed dir/quant nuances.




BOM three points – 8.3, 3.2, 16.1 km away
ERA5 four pts – 21.1, 15.2, 20.7, 14.5 km
ERA5-Land three points – 11.5, 2.9, 10.3 km
MERRA-2 four pts - 15.2, 46.9, 44.6, 63.1 km
JRA-55 one pt – 440 metres
37.75 s 145 e
37.75 s 145.25 e
Seven Seeds Coffee Roasters
BOM Olympic Park
37.9 s 145.1 e
37.9 s 145.2 e
BOM Oakleigh
JRA-55 grid point
Monash OikoLab point
BOM Moorabbin
Image © 2021 TerraMetrics
38 s 145 e
38 s 145.1 e
38 s 145.2 e
Imagery Date: 12/1/2018

Other variables

- Could Phase 1 forecast have been improved with extra data (NWP, AEMO etc) or a different approach? (using AEMO data might be a bit circular)

Yes, but not by large amounts

- AEMO price and demand data (had to download 3 files for competition Phase 1 & 2) is *half hourly* – is microgrid subject to wholesale price? Price/Demand improves B0/B6 forecast!
- AEMO Rooftop PV Actual data from NemWeb is *half hourly*
- ERA5 precipitation data – e.g. ILSPF “Instantaneous large-scale surface precipitation fraction”
- ERA5-Land data is 0.1 degrees – but only 3 points to interpolate from
- Other solar vars for PVLib: FDIR ~ GHI, SSRDC, CDIR to derive DNI, DHI etc. Diffuse radiation.
- Wind direction
- JRA-55 has 3-hourly data grid point 400 m from Monash
- NASA MERRA-2 1h data - SWGNT ~ SSRD
- GFS reanalysis data (3-hourly) is painful to process
- PvOutput.org has many nearby points (5 min data, $15 donation for 1 year access) or Solar Analytics
- WeatherMan/Solcast approach – derive solar installation parameters from data, resimulate

PERFORMANCE IN PHASE 2

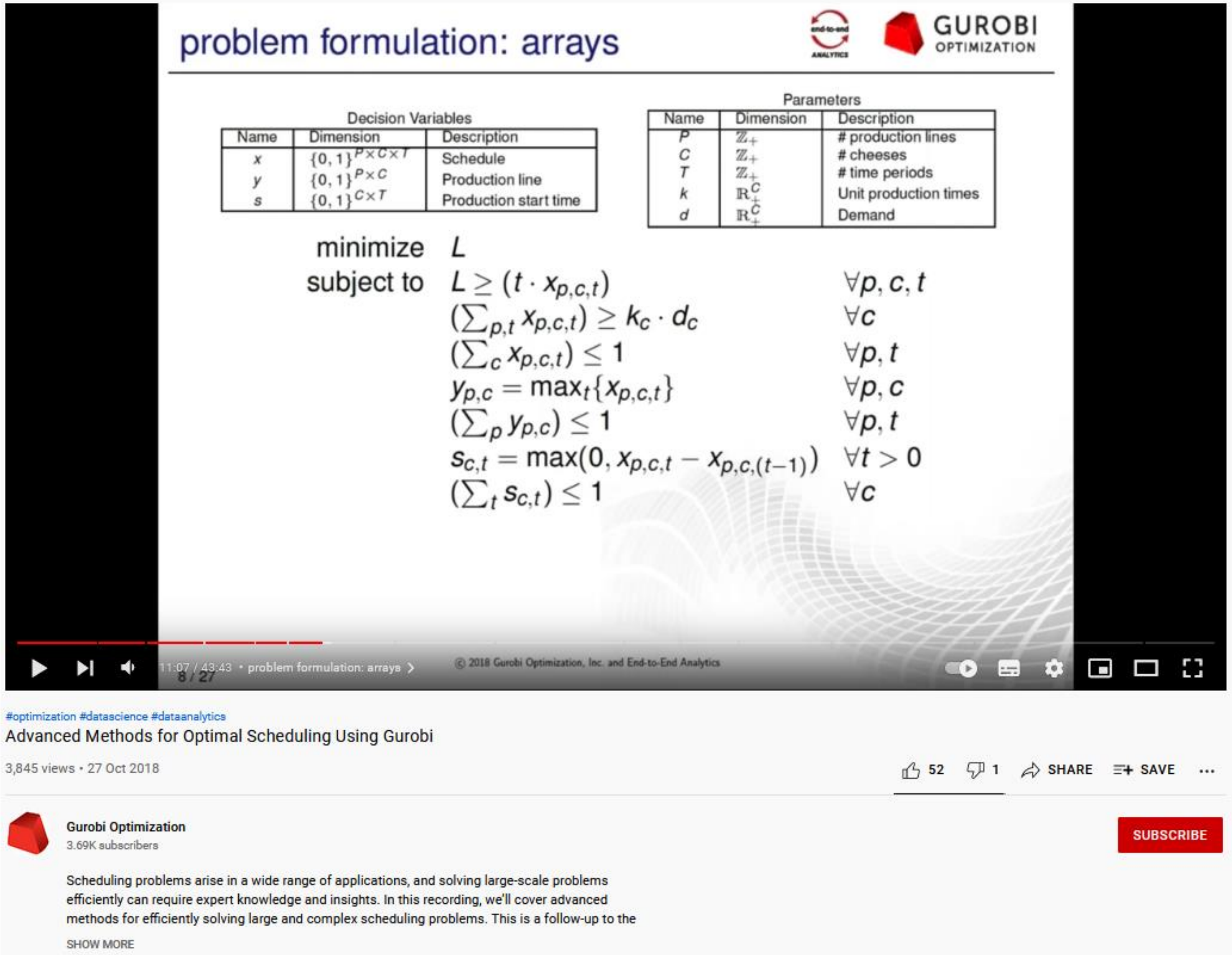
- Best forecast of 9 of 12 time series (3/6 Buildings, 6/6 Solar)
- Combining all entries only improves to 0.6276
- Using AEMO 5-minute Vic demand allows more improvement

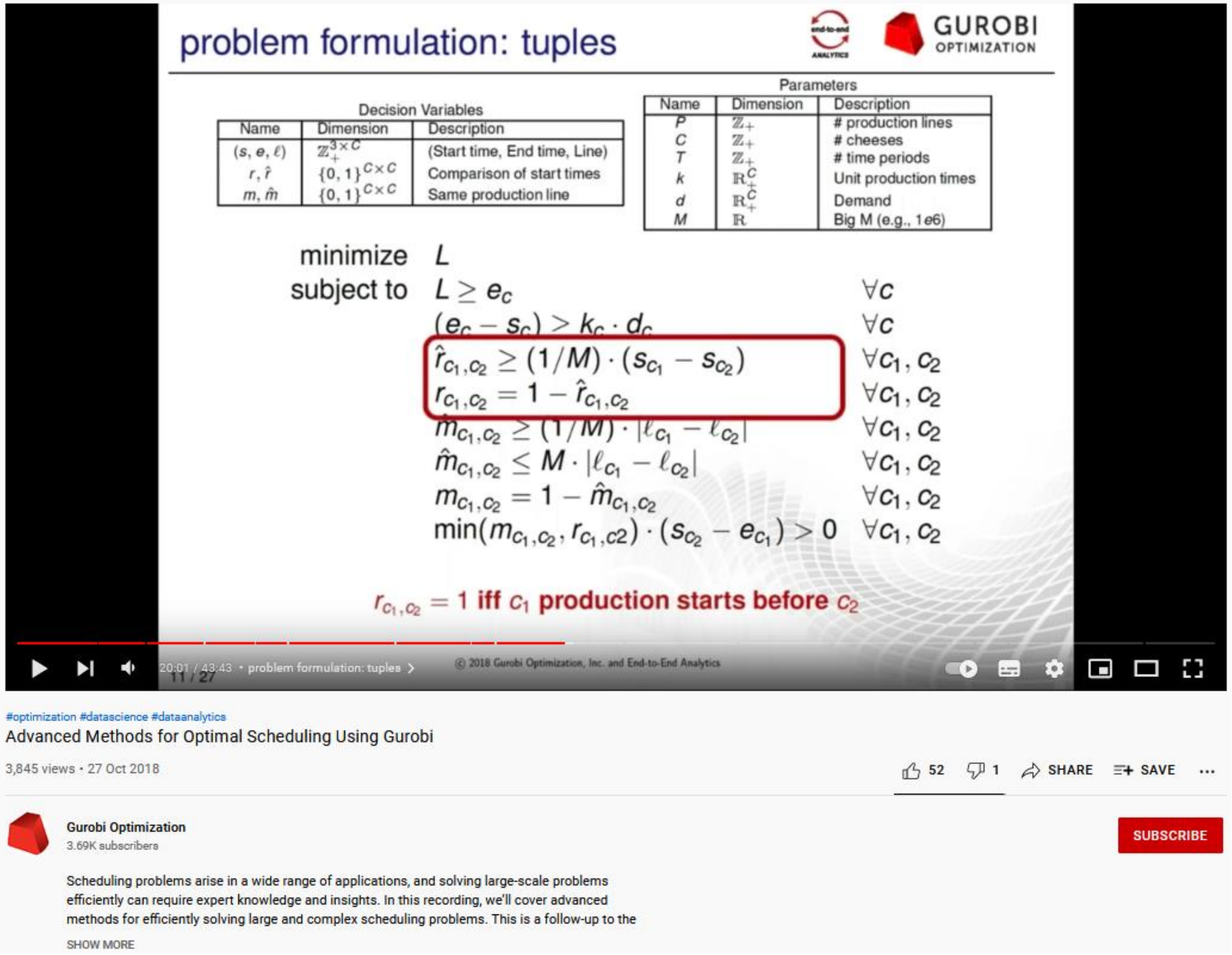
MASE	Bean	Abolghasemi	SZU	EVERGi	Stratigakos	FRESNO	Best	with AEMO demand
mean	0.6460	0.7441	0.7750	0.8073	0.8474	1.0026	0.6276	0.6395
Building 0	1.0438	0.9081	0.9413	1.2008	1.3227	1.2376	0.9081	0.9876
Building 1	0.8840	0.9610	1.0171	1.1341	1.0362	1.1077	0.8840	0.8724
Building 3	0.6494	0.7524	0.6099	0.6398	0.7785	0.7711	0.6099	0.6486
Building 4	**0.7236**	0.6775	0.7236	0.8096	0.8269	0.7236	0.6775	0.7236
Building 5	**0.7922**	0.9654	0.8563	0.9493	0.8463	0.9157	0.7922	0.7922
Building 6	0.7476	0.7822	0.8611	1.0182	0.8577	0.7694	0.7476	0.7457
Solar 0	0.6019	0.9305	0.9159	1.0439	1.0170	1.4421	0.6019	0.6016
Solar 1	0.3860	0.4187	0.5222	0.3988	0.5416	0.8155	0.3860	0.3831
Solar 2	0.4148	0.5314	0.6186	0.5248	0.6656	0.9462	0.4148	0.4166
Solar 3	0.5475	0.7032	0.6678	0.7221	0.7118	1.0440	0.5475	0.5420
Solar 4	0.4179	0.5616	0.6516	0.5173	0.7002	0.9947	0.4179	0.4183
Solar 5	0.5435	0.7366	0.9145	0.7289	0.8639	1.2640	0.5435	0.5417

Optimization

- Solving the model as a MIP is much easier than solving the MIQP.
- Almost all of the submitted solution depends on first deriving the best MIP solution possible (i.e. minimizing the recurring load or minimizing the recurring + once-off load) and only then solving as an MIQP
- Gurobi 9.1.2 (laptop phase 1, UQ HPC phase 2)
- Various papers about "Predict+Optimize" problem but Phase 1 and leaderboard seem to indicate no close relationship between forecast result and cost. Complex problem, competition issues, limited time

Arrays vs tuples approach (Nicole Taheri)





Now that we have a good starting point ...

- **Conservative** is just choosing the lowest recurring load and lowest recurring + once off load and evaluating cost using a naive or flat forecast. This was probably the winning approach for cost in Phase 1, as some competitors had winning results with no forecast, or a poor forecast, but seemed pointless to me as the organizers said quality of forecast should contribute to results in phase 2.
- **Forced discharge** forbids any charging in peak hours, and forces at least one of the two batteries to be discharging in every peak period.
- **No forced discharge** forbids any charging in peak hours, but the MIQP solver decides whether to discharge or do nothing in those hours.
- **Liberal** allows charging in peak, but the maximum of recurring + once off + charge effect for each period is limited to the maximum of recurring + once off load over all periods. This is to avoid nasty surprises when the solver thinks that a period has low underlying load and schedules a charge (due to a low price in that period) but then accidentally increases the maximum load over all periods, which can be very costly.
- **Very liberal** allows charging over peak and does not attempt to control the maximum of recurring + once off + charge effect. This would be the best approach if the forecast was perfect.

Estimated Cost

Only Large2/Large4 had the once-off load in, all activities, in peak.
The estimated cost is very different from the real cost.

Winning solution had almost all once-off activities included.

Case	Estimated Cost	Actual Cost
small0	26681	34166
small1	26233	33682
small2	26251	33235
small3	26452	33977
small4	26107	33462
large0	26265	32417
large1	26666	33842
large2	25389	32841
large3	26010	33149
large4	25849	33334
Total	**261906**	**335107**

		bean	esmailbeigi	evergi	akylas	sample
small	0	34166	34509	35676	37281	57941
	1	33682	33265		36862	50096
	2	33236	32428		34342	59924
	3	33977	33136		38344	46427
	4	33463	32490		35263	99669
large	0	33417	32643		34644	46404
	1	33842	33055		34949	78291
	2	32841	31712		36050	42501
	3	33149	32219		35389	55874
	4	33334	32903		39390	52230
	total	335107	328359		362515	589357

Anonymous Peer Review stage

Final scoring not just based on cost, but on presentation + 4-page report.
https://arxiv.org/abs/2202.00894
Eight member scientific committee of academics, 3 from Monash
Ratings – 1 excellent, 2 very good, 3 good, 1 acceptable, 1 poor.
Evaluations ran from

"This was the submission I judge to be the best"

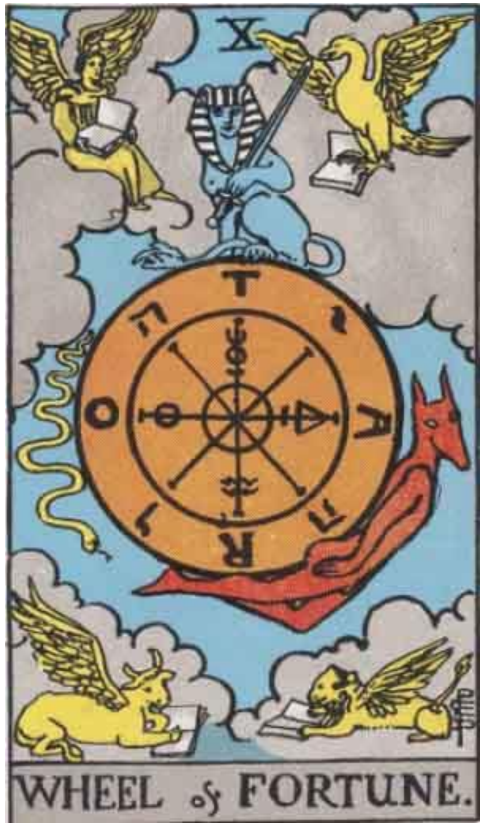
to

"There seems to be some manual tuning and heuristics, but overall the paper is well explained and the decisions justified, and I think this would be very helpful to the readers that want to implement something similar."

to

"The results are obviously good, but the methodology is very ad-hoc, to the point of forecasting manually chosen constants in some cases."

Some armchair quarterbacking going on. Objective is to win forecasting and lowest cost approach, and there was no time to assess multiple approaches. Ad hoc approaches were clearly the best on several time series. Struggled against organizer mistakes – luck involved.



Scientific Committee

- Mark Wallace (Professor, Monash University)
- Guido Tack (Associate Professor, Monash University)
- Pablo Montero Manso (Lecturer, University of Sydney)
- John Betts (Senior Lecturer, Monash University)
- Yanfei Kang (Associate Professor, Beihang University)
- Alejandro Rosales (PhD, CIMAT, Mexico)
- Isaac Triguero (Associate Professor, University of Nottingham)
- Daniel Peralta Cámara (PhD, Ghent University)

Summary

- Random forest – 4 models for each hour
- Use daily BOM solar data + ECMWF hourly data + temporal variables
- Train buildings and solar together in groups
- Thresholding two solar series

- Arrays approach with 0-1 Mixed Integer Program (MIP)
- First minimize recurring and recurring + once-off load, then solve MIQP
- “No forced discharge” approach chosen from 5 approaches

THE UNIVERSITY OF QUEENSLAND AUSTRALIA

LESSONS

- Luck is an important factor
- The most effective methods may be absolutely *ad hoc*
- “AI competitions don’t produce useful models”
- Ranger package with multithreading is very useful
- Python is much better than R for Gurobi
- If you’re a competition organizer …
  - Timely communication is essential (on errors, assessment, prizes)
  - Clarify rules early – had to submit code at end of Phase 2

How a Kaggle Grandmaster cheated in $25,000 AI contest with hidden code – and was fired from dream SV job

Pet adoption ML coder apologizes and says desire to be ranked #1 'compromised my judgement'

Katyanna Quach 101

Tue 21 Jan 2020 // 09:24 UTC

Secondly, and most importantly: **Everyone who wins money from a competition should be required to open-source their solution**. I am not the first one to say this and I have no idea why it is still not the case. I understand that competition sponsors should be able to opt-out of this for privacy reasons (but then again, why host a Kaggle competition in the first place if that's the case?), but it should definitely be the default.

I would even go as far as automatically publicizing all submitted solutions after the competition. At least for Kernels competitions this would be easy to implement and it would reduce fraud to essentially zero. Unless I am missing something this should not be a problem.

How has this helped in other work?

- Phase 2 data now public – meaningful improvement is difficult
- Preparing paper for **"International Journal of Forecasting"**
- **Project SHIELD** – data quality indicator using ECMWF data
- **Cybersecurity** of energy data – Springer chapter to appear
- What *are* the best solar/load predictor variables in an Australian context?
- **Bike sharing / e-scooter data** – strong parallels to energy data, but:
  - Data quality is much higher
  - Geographically more diverse (e.g. study with 40 cities / 16 countries)
  - User information is available (gender, age, subscriber type)
  - With energy systems, user information is highly confidential and it is difficult to obtain NMIs, addresses, occupancy, gender, age, employment info etc. Huge delays due to legal issues
  - Grants?

Thank you

Contact information

R.Bean1@uq.edu.au