

EXPERT MEETING IBPSA Project 1 WP1.2 MPC – OCTOBER 2020

BREAKOUT SESSIONS

Subject: WP1.2 Breakout sessions

Date: 13&14-10-2020

Location: Zoom

Minutes taken by: Javier Arroyo

PARTICIPANTS

| Name | Affiliation | BS 1 13/10 | BS 2 13/10 | BS 3 14/10 | BS 4 14/10 |
|----------------------|--------------|------------|------------|------------|------------|
| Lieve Helsen | KU Leuven | x | x | x | x |
| Filip Jorissen | | | | x | |
| Javier Arroyo | | x | x | x | x |
| Iago Cupeiro | | x | x | x | x |
| David Blum | LBNL | x | x | x | x |
| Michael Wetter | | | | | |
| Zhe Wang | | | | | |
| Valentin Gavan | ENGIE Lab | | | x | |
| Tao Yang | SDU | x | x | x | x |
| Konstantin Filonenko | | x | x | x | x |
| Christian Veje | | | | | |
| Huang Sen | PNNL | x | | | |
| Yan Chen | | | x | | |
| Jan Drgona | | x | x | x | x |
| Draguna Vrabie | | | | x | x |
| Laura Zabala Urrutia | IK4 Tekniker | x | x | x | x |
| Jesus Febres | | | | | |

| | | | | | |
|--------------------|--------------------------------|---|---|---|---|
| Aritz Bengoetxea | | | | | |
| Harald Taxt Walnum | SINTEF | x | x | x | x |
| Ettore Zanetti | Politecnico de Milano | x | x | x | x |
| Piljae Im | ORNL | | | | |
| Yeonjin Bae | | x | x | x | x |
| Gaurav Chaudhary | NTNU Norway | | | | |
| Felix Bünning | ETH Zurich/EMPA | x | x | x | x |
| Kyle Benne | NREL | x | x | x | |
| Nicholas Long | | x | x | x | x |
| Hagar Elarga | | x | | | |
| Thibault Marzullo | University of Colorado-Boulder | x | x | x | x |
| Sourav Dey | | x | | x | x |
| Gregor Henze | | | | | |
| Gerald Schweiger | IST /TU Graz | | | x | |
| Flavia Andrade | CARTIF/UCD | | | x | x |

AGENDA and REPORT

https://docs.google.com/spreadsheets/d/1J2MbPfFaNVjtRh8yi_esw9fa9wXo7jbSejMWni-g_bY/edit#gid=2018107631

| | Content - title | Presenter/Leader | time |
|------------------------------|--|------------------|---------------|
| Session 1 (Day 1) | BOPTEST: benchmarking and reporting | | 50 min |

| | | | |
|------------------------------|--|-----------------------------|---------------|
| | Presentation of results of first tests | Dave/Javier/Filip/PNNL/.. | 20 min |
| | Spreadsheet for internal reporting of BOPTTEST test cases | Javier | 15 min |
| | Public reporting (limited information, end-user agreement) | NREL/Dave | 15 min |
| Session 2 (Day 1) | Data and predictions | | 55 min |
| | Generating weather data for training grey/black box models | David Blum/Yan Chen | 20 min |
| | Dealing with uncertainties - weather prediction | Laura | 15 min |
| | Dealing with uncertainties - loads & occupancies | All | 10 min |
| | Issues related to emulators | All | 10 min |
| Session 3 (Day 2) | New developments | | 45 min |
| | Multi-objective Deep Reinforcement Learning Control (MODRLC) project | Nicholas, Thibault, Sourav, | 10 min |
| | Data-driven approaches – opportunities for MPC & BOPTTEST? | Jan Drgona | 15 min |

| | | | |
|------------------------------|--|----------------|---------------|
| | New developments in BOPTTEST | David Blum | 15 min |
| | MPC Library | Filip Jorissen | 5 min |
| Session 4 (Day 2) | Outreaching | | 45 min |
| | MPC competition: how to set up? | All | 20 min |
| | Workshop introducing BOP-TEST at BS2021? | All | 20 min |
| | Initiatives for joint papers | All | 5 min |

SESSION BS 1 BOPTTEST: BENCHMARKING & REPORTING

Presentation of results of first tests

Lieve brings up the question: who has tested BOPTTESTS?

Dave and Javier confirm they've successfully tested it.

Felix Bunning has a thesis student working on this who started a week ago and provides positive feedback on the easiness of getting it running.

Nicholas Long tried one year ago and is now catching up with new developments.

Iago's thesis student is going to start using BOPTTEST as well.

Dave presents his first experiments with the BESTEST air based case. Using MPCPy with an R3C3 grey box model in an MPC that controls fan speed and supply air temperature setpoints. The model is trained with one week of data and uses the following week for testing. The MPC uses a moving horizon estimator, a prediction horizon of six hours and a sampling time of 10 minutes. His MPC works well with BOPTTEST and success on testing. Issues raised from experience:

- “Weather station” block is needed within BOPTEST to model a weather station nearby the building that is capable of providing weather data
- Dynamic and highly cost options need to be implemented in the BOPTEST interface when computing KPIs.
- Data input time checking also to be implemented.
- Feature needed to reset time ratio.

The presentation can be found [here](#).

Javier presents his experience with BOPTEST and illustrates through results in one publication and through the results of an MPC in the residential single-zone hydronic case with heat pump. The MPC also uses a reduced order grey box model trained with one week of data trained in January and tested in February. The MPC uses an unscented Kalman filter as state estimator, one day ahead prediction horizon and 15 minutes sampling time. The main lesson learnt is the need to characterize controllers and test cases in order to enable assessment and comparison across different controllers in the same test cases. The presentation can be found [here](#).

Felix is implementing different data-driven MPC methods and using BOPTEST for comparison of different identification methods. The tutorial from Rome is really helpful to install and start hands on with BOPTEST. We'll update the tutorial and move it to the repository.

Dave brings up other issues regarding software development like the memory issues raised by Harald. **Harald** has done some testing with their own emulators. Very user friendly! Most difficult part is getting Docker running on Windows. He also explains an issue with the forecaster at the end of the year. A potential solution could be the repetition of the year: weather data at the end of the year is continued with weather data from the beginning of the same year → may lead to discontinuities.

Ettore is running month by month which means that the memory problem raised by Harald. He is also using his own emulator model. Ettore raises an issue when gathering data and setting parameters for his own emulator. Lieve recalls that BOPTEST is not meant as a simulation framework but as a benchmark to assess different controllers and therefore we can't allow users to parametrize the test cases too much.

Spreadsheet for internal reporting of BOPTEST test cases

Javier elaborates on the topic with a presentation. Main objective is to characterize deployed test cases to enable comparison of different controllers. A deployed test case is the combination of a controller and a test case. A controller is defined as an aggregation of features and can be really complex. Which features should remain in the definition of a controller? A trade-off between accuracy and simplicity should be accomplished. Regarding test cases, these can be defined as a conjunction of one of the previously defined building models and a set of parameters that define boundary conditions: weather forecast uncertainty, price scenario...

A spreadsheet has been started up in order to start sharing results and to learn about how to report results from experience.

Public reporting (limited information, end-user agreement)

Discussion on the BOPTEST web application: tags for controllers comparison (post-processing), what information needed about the controllers and tests.

Kyle takes over to show the current status of the graphical user interface being developed by Devetry for public reporting. He interactively shows the status of the interface. Main idea is to gather and display results in a web service. Around 75% already developed. Lieve points out that we should constrain the definition of test cases as much as possible. This means that we shouldn't mix results from different seasons and that we should filter clearly such that we can do benchmarking.

Discussion. What are the keys, categories, choices,... in order to sort and characterize controllers?

Kyle: I've the feeling It's imperative that test cases are accurately defined whereas controller definitions are more flexible.

Lieve: still, controllers need to be accurately identifiable to find insights from comparisons.

A disclaimer is needed for public sharing such that users assume responsibility of the veracity and integrity of their results.

Kyle: we're considering launching a hosted instance of BOPTEST. That would mean that users would not have to deploy containers in their local machines if using this hosted version. In those cases users would have the possibility to decide whether they want to share metadata of their controllers.

SESSION BS 2 DATA & PREDICTIONS

Generating weather data for training grey/black box models

Dave introduces the problem of year-long test period runs when data is required for system identification. Potential solutions would be to use TMY3 for training and TMYx for testing or other combinations. Dave has gone through different TMY formats and found out that they indeed follow slightly different distributions. The presentation can be found [here](#).

Felix: would it be possible to hide the data to not use it for testing? → Really challenging as the other data needs to be used for forecasting for instance. It's not a matter of just data but also of morale.

Harald: We already discussed this in the past and concluded that it is not possible to avoid cheating entirely. Lieve: it's still possible to state a list of good practices. There are other ways of directing towards good practices that pop up during discussion like a disclaimer where it's needed to check a box stating that e.g. testing data has not been used for training.

Jan brings to the discussion whether we need to include a KPI quantifying the memory requirements of a controller. Dave points out that that is related to how the controller is characterized. Dave points to the "[Development requirements and guide](#)" where there are some guidelines on how to deal to this topic.

Dealing with uncertainties - weather prediction

Laura elaborates on the topic with a presentation. The presentation can be found [here](#). The objective is to provide the user/tester with uncertain data for forecast. We've gathered data (historical forecasts) from three different locations: Norway, Berkeley and Leuven, in order to test and assess methodologies to emulate forecast uncertainty.

First approach implemented is the autoregressive model developed by Krzystof. It was calibrated with data on temperature forecasts from Berkeley and Norway. Working well in Berkeley but not in Norway. Not all locations follow a normal distribution. It also has technical limitations related to initialization. A robust solution for this approach is expected by February 2021.

Second approach proposed by Spanish company working with predictions and named ECDF-QQ mapping. When calibrating this model for the different locations it was found to have great errors in extremes of the ECDF. To be continued after February 2021.

Next steps for the short term are to keep working with the autoregressive model as there is an initial working version and improve from there. A deadline has been set for February to have a working version in BOPTTEST. After that the ECDF-QQ mapping approach may be considered.

Dealing with uncertainties - loads & occupancies

Topic of discussion: how should uncertainty be defined in order to allow users to retrieve stochastic forecasts from BOPTTEST?

Lieve: probably the best way is to define it in scenarios → deterministic, lower uncertainty level, medium uncertainty level, high uncertainty levels.

Dave: agree, then the next question is how to define uncertainty level?

Lieve: it can be defined in terms of the number of times that a "sigma" distribution is added on top of the deterministic forecast. That is assuming normal distribution, what to do with the other distributions? Like for instance loads and occupancies? Occupancies do not follow normal distributions. For loads we should contact Ina and Anke as they've work to be published on this.

The main focus on uncertainties from now and until the next meeting will be in weather predictions, not on loads and occupancies. Lieve mentions a PhD on propagation of uncertainties in the construction parameters and disturbances. She will share with the team once first results are obtained.

Issues related to emulators

Dave raises the issue of the single-zone commercial hydronic case that simulates fine in Dymola but is crashing at a very initial time when simulating using pyfmi in the BOPTEST framework. He receives some inputs from colleagues on potential solutions and follow up on this is expected in the github repository.

Dave asks whether the multi-zone commercial hydronic building simple (Infrax model) had the same issue. Javier: it's not exactly the same as that one was related to event indicators, details can be found in the pull request. Konstantin and Tao will take up on this.

Lieve closes the meeting.

Emulators overview:

| Emulator | Developer |
|--|-------------------|
| Single-zone BESTEST hydronic | Filip & Javier |
| Single-zone BESTEST hydronic (modulating HP) | Javier & Filip |
| Single-zone BESTEST air-based (gas boiler) | Dave |
| Multi-zone (8z) residential hydronic heating (gas boiler) | Valentin & Javier |
| Multi-zone (8z) residential hydronic heating + air cooling | Valentin |
| Single-zone commercial air-based | Dave |
| Single-zone commercial hydronic | Krzysztof / Tao |
| Multi-zone (5z) commercial air-based | Dave |

| | |
|--|---------------|
| Multi-zone office hybrid (simple) | Iago & Javier |
| Multi-zone office hybrid (complex) | Filip |
| Multi-zone commercial air-based | Yeonjin |
| Multi-zone prototype air-based (complex) | Sen Huang |

§ Merged to master. Ready to be used!

§ In pull request. Under peer-review (working on latest changes and documentation)

§ In pull request. BOPTEST Additions needed (like IO blocks, test case data or unit-tests)

§ In pull request. Model changes needed

§ Emulator model under development

SESSION BS 3 NEW DEVELOPMENTS

Multi-objective Deep Reinforcement Learning Control (MODRLC) project

Intro project of University of Colorado Boulder & NREL

Thibault, Sourav and Nicholas from NREL and University of Colorado present MODRLC. Their presentation can be found [here](#). Aim is to develop multi-objective deep RL control for GEB as well as test transfer learning, and perform comparisons of performance among different controllers. Additionally, a comparison between RBC, MPC and RL will be performed in the project. The OpenAI Gym interface is to be used for the latter. The roadmap is: 1. ACTB

development, 2. Evaluation of control strategies (using BOPTEST), and 3. Field implementation. The software architecture integrates a spawn model, the use of the BOPTEST interface and a python controller agent that is using RL. The spawn model provides several advantages like reliability of EnergyPlus and the flexibility of Modelica.

The control strategies to be evaluated are a simple rule based, then an MPC and then three RL agents like DQN, PPO and SAC.

Work to date: EnergyPlus model verified. Developed a spawn model of a DOE prototype (only works on Dymola right now, still some development to work on pyfmi). Sensors and setpoints are already exposed through the BOPTEST interface.

Data-driven approaches – opportunities for MPC & BOPTEST?

Further discussion.

Lieve: are you going to develop your own testbed or is BOPTEST the framework where you want to embed your testbed. Nicholas: ideally we would set up a synergy with the BOPTEST team to extend BOPTEST with our test case.

Lieve: why not to use Modelica directly? Dave: it's only the building envelope that is developed in EnergyPlus.

Konstantin mentions that the modelicagym package may be worth checking: <https://github.com/ucuapps/modelicagym> as it utilizes modelica models through an OpenAIGym interface.

Dave: there has already been some discussion on OpenAI Gym interface for BOPTEST.

Javier: Here there is an example of the development of an OpenAIGym interface around BOPTEST: <https://git.isis.vanderbilt.edu/SmartBuildings/bop-test-gym>. And here my own development: <https://github.com/JavierArroyoBastida/project1-bop-test/tree/gymEnv>

Dave: concentration is needed to guarantee quality under the hood.

Lieve: Is it envisaged in the project to study some kind of hybrid models?

Sourav: No, not at this project. Maybe in the future.

Thibault clarifies that the focus of the project is on transfer learning in real implementation.

Ettore: does it make sense to develop a guide for BOPTEST developers? Not only test case developers but for actual backend developers. Dave: certainly... that's something we intend with the design requirements and guide. Ettore will start by having a look at it.

New developments in BOPTEST

Discussion on developments and maintenance

Dave presents the latest developments in BOPTEST since the Rome meeting. Presentation can be found [here](#). Summary:

- Created test case repository
- Specify start time with warmup period
- Quick-start documentation with README.md updated
- Added CO2 to KPI calculator
- Read block now reads Zone operative temperature
- Specific zone designation in signal read block
- Unused columns in csv data omitted
- Header lines in csv data omitted (only used columns read)
- Header lines in csv data for documentation
- Demonstrated customized KPI calculation
- BOPTEST-Sercive development branch created

Outstanding issues:

- Reported large memory usage
- Shift towards Python3 for simulation while still using JModelica for compilation of Models
- Price scenario API needs parameter for choosing price scenario
- Weather station to retrieve weather data
- Forecasts after one year
- Normalize energy KPI by floor area
- Signal exchange blocks as arrays and combine
- Error handling, logging, and user messaging
- Users guide and tutorial
- KPI measuring actuator travel
- Enhanced documentation for test cases

Future thinking:

- Provide controllers with input/output tagging and metadata schemes like Haystack
- Prototype the needed implementation in Modelica models and pass-through to the API
- Architecture for proposed web-hosted instance of BOPTEST
- Leveraging development for previous Alfalfa project

Jan presents further discussion topics. Presentation can be found [here](#). Summary:

- How to prevent cheating? Checkbox with accepting rules statement?
- How to promote fair comparison of MPC with data-driven controllers? A solution could be peer-reviewed KPIs, like offline cpu time, memory requirements, data requirements...
Lieve: this is difficult to peer review → should be included in the rules agreement.

- Differentiation in data usage? We could use BOPTEST for generation of static data to attract the ML community. Dave: this one is potentially for the future as the focus of BOPTEST is control evaluation which relates to dynamic data rather than static data. Related to this there was an ASHRAE project to predict the energy consumption: <https://www.kaggle.com/c/ashrae-energy-prediction>.

MPC Library

Future plans on the MPC library, this is a repository of simplified models that are usable for optimization.

Filip: there is not much to discuss here as there are only two people using it.

Dave: someone interested in using or collaborating in such a library? Iago would be happy to contribute with his borefield models once I finished the writing of my PhD.

Dave: There is still value in having simplified models for MPC.

Filip: is it a priority and worth investing time? It can not be done overnight...

Lieve: since it's in the main interest of only two institutes, namely LBNL and KU Leuven, we can conclude that it's not a priority and therefore we'll park this task instead of bringing it up every meeting so that we can focus on other priorities first.

SESSION BS 4 OUTREACHING

MPC competition: how to set up?

Targeted for 2022. Not really near term but would be nice to have some thoughts about it.

To be organized as a follow up after the workshops are organized either in Spring or Summer.

Workshop introducing BOP-TEST at BS2021?

Incl. inspiration towards use cases such as: training building operators, assisting teaching ...

Lieve: is it something that we are going to organize?

Dave and Javier agree to take the lead on it.

The organization of this is not paid by the conference. The sponsorship of Engie could be used for this. Lieve will pass the information of what is needed to Dave and Javier.

The workshops are organized the day before the conference starts. The length of the workshop is flexible.

May be a good change to announce that there will be a contest on advanced controllers for BOPTEST in Spring or Summer 2022. For the workshop itself Dave proposes an exercise where the controller has not to be developed from scratch as that could be too much for the workshop.

Jan proposes to do this kind of workshop periodically in the building simulation conferences and ML communities. Javier takes the lead for the first workshop on the BS community (September) and Jan for the workshops on the ML community (December, see: <https://www.climatechange.ai/>).

Initiatives for joint papers

Dave has been thinking for a while for a journal paper on BOPTEST in order to have a more complete description of the features and initial tests. Lieve agrees and thinks it is really needed. Dave and Lieve agree that the Journal of Building Performance Simulations may be the best target. Dave will develop and outline and will share that in an upcoming meeting. Lieve puts it on the agenda for the meeting of November.

Priority setting for monthly progress meetings

Changed format of these meetings: first a short update (like one quarter of an hour) and follow with some more time for discussion (like three quarters of an hour). Issues to be tackled in coming meeting:

- **November meeting:** Controller characterization, BOPTEST workshops, Journal paper (the latter only if time allows otherwise in december).
- **December meeting:** Journal paper, weather uncertainty emulation.
- **January meeting:** Reporting results for testing.

Emulator issues can be better treated through the repository.

We agree that meetings will last for 1,5 hours instead of only 1 hour.

Javier will raise an issue for the OpenAI Gym interface and will organize a meeting with Dave, Jan, Sourav and Konstantin to go through the development of this wrapper interface for ML algorithms.

Issues on the service branch for BOPTEST are: the deployment of multiple workers to run several simulations at the same time and the ability to upload custom test cases when running in sort of a privileged way. Dave: the latter is outside of the scope of BOPTEST as test cases are very carefully predefined.

WRAP UP BREAKOUT SESSION

1. BS1 - BENCHMARKING & REPORTING

Positive feedback from initial BOPTEST trials. At least 5 people already tested in simple cases and shared their experience. Some issues and lessons learnt have been discussed. The tutorial used in Rome has proven to be really helpful. There is a call to test and report issues.

The reporting of results takes place initially through a spreadsheet while the official public reporting interface is being developed as a web service. Crucially important is the definition of test cases and characterization of controllers to compare different controllers and find insights.

2. BS2 - DATA & PREDICTIONS

Different TMY formats may be used to generate data for training grey-box and black-box models in data periods out of the testing period. A disclaimer will be included to avoid cheating. Users will have to check a box to declare that they comply with ethics rules and that they are responsible for the integrity of the results they are sharing.

A clear plan for defining uncertainties for the weather forecast is drawn. Two approaches are envisaged: autoregressive models and QQ mapping. The level of uncertainty will be picked by the user from a set of qualitative ranges, namely: deterministic, low uncertainty, high uncertainty.

The emulators under development are running fine in Dymola, although few issues arise when simulating in JModelica which need to be addressed.

3. BS3 - NEW DEVELOPMENTS

University of Colorado Boulder presents a project where RBC, MPC and RL are to be compared. Their emulator could replace one of our emulators that still needs to be developed.

The MPC library task is parked as is not a priority. Only two partners interested and contributing to it right now: LBNL and KU Leuven. Iago's borefield model can still be added.

The latest developments in BOPTEST have been presented along with current issues and figure steps.

4. BS4 - OUTREACHING

A BOPTEST paper for the Journal of Building Performance Simulation is envisaged.

A BS2021 workshop will be organized that is focused on BOPTEST. In this workshop we will teach how to use BOPTEST and announce a future contest in Spring and Summer where advanced controllers will be tested in BOPTEST. Additional workshops are also conceived in the ML community starting from December 2021.