

Supported by:



on the basis of a decision
by the German Bundestag

Model-based methods to improve computing times in linear energy system optimization models

Karl-Kiên Cao, Kai von Krbek, Manuel Wetzel, Felix Cebulla

May 22nd 2019

DLR – German Aerospace Center, Department of Energy Systems Analysis

A PROJECT BY



Deutsches Zentrum
für Luft- und Raumfahrt
German Aerospace Center



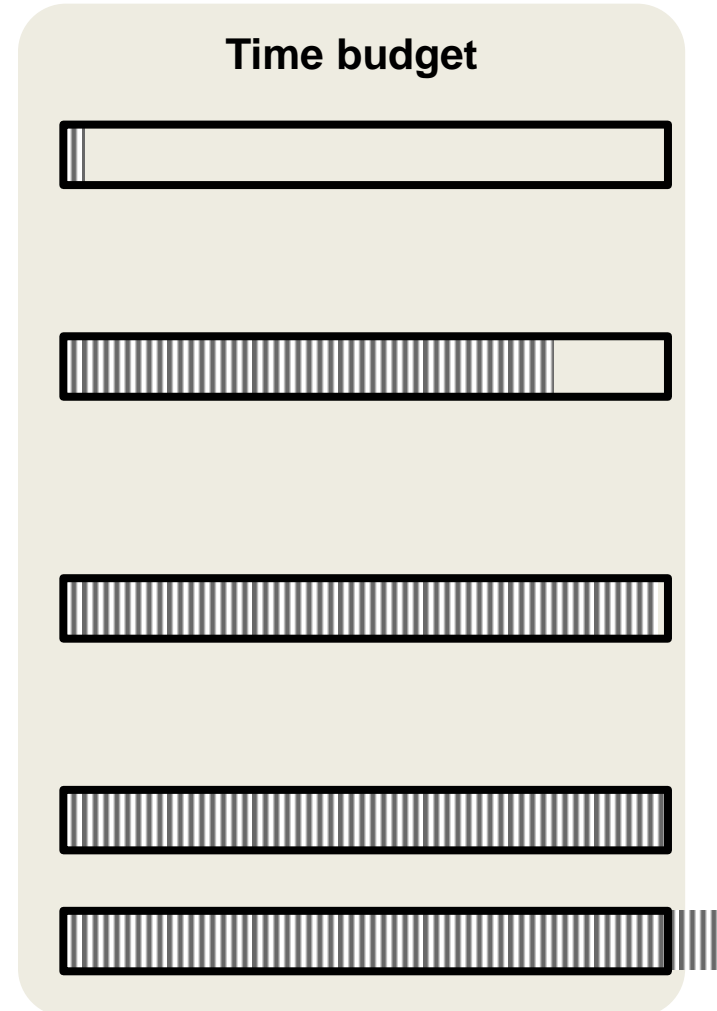
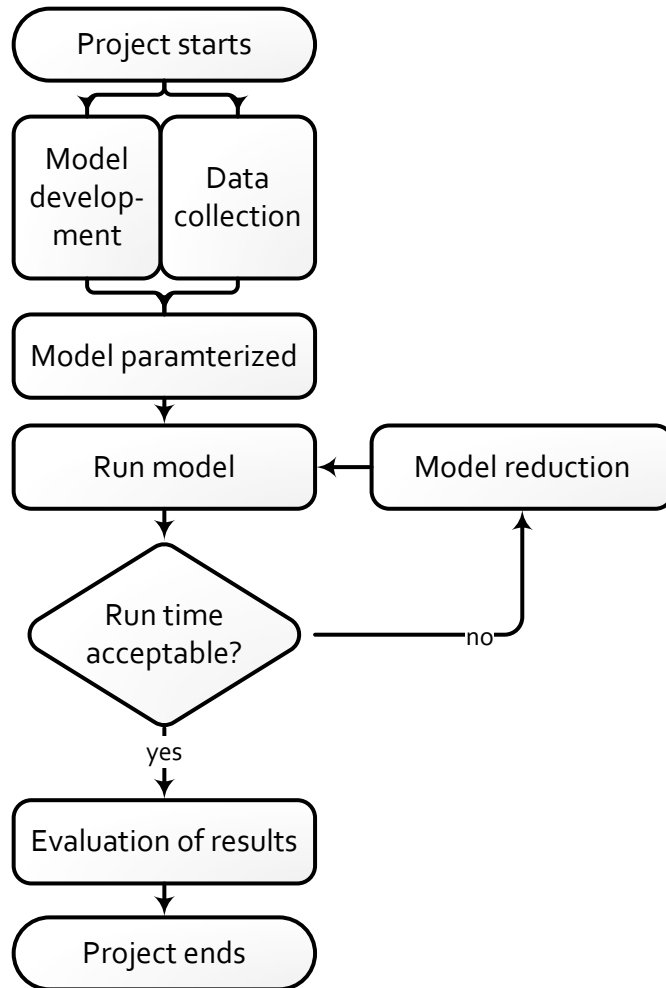
- Motivation
- Repetition: *Characteristics and dimensions of Energy system optimization models*
- Theory: *Classification of existing approaches*
- Evaluation methodology
- Results and major findings
- Conclusions

Motivation

What can modelers do
by themselves?

What have modelers done
by themselves?

Everyday's Energy Systems Analysis



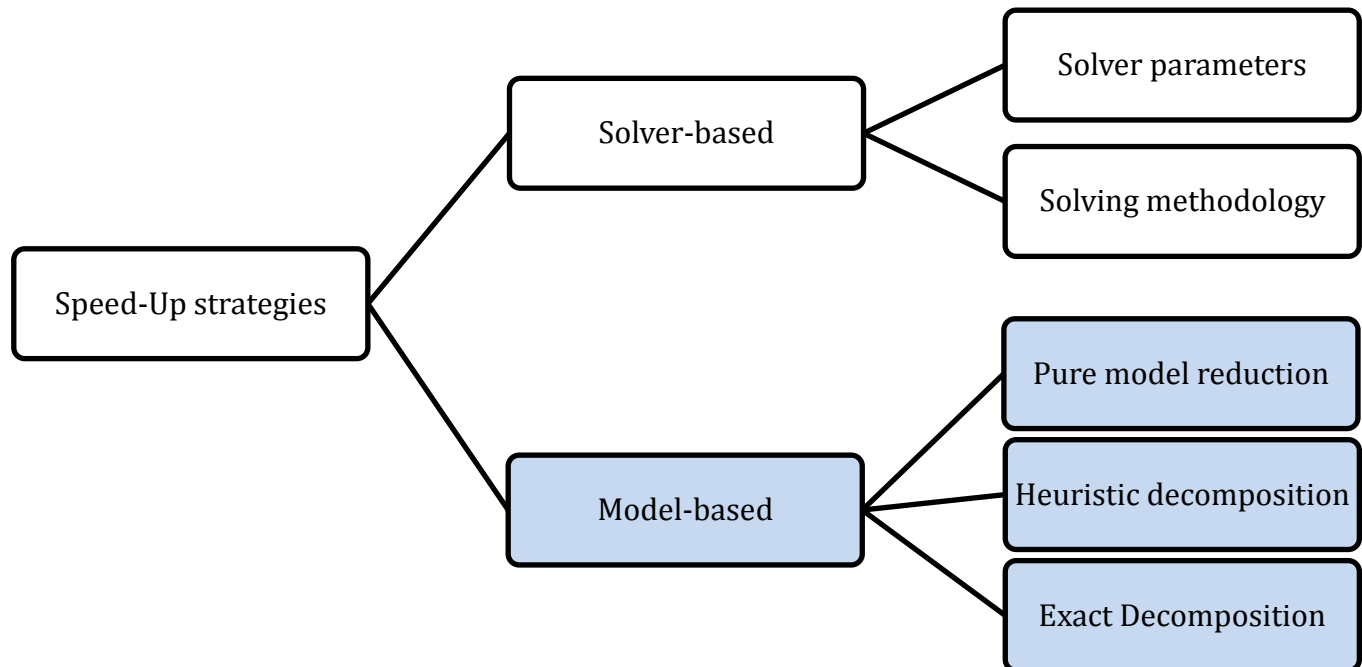
Which **speed-up** is possible
using measures that can be influenced
by „normal“ model developers?

- Large applied Energy System Optimization Models
 - LPs
 - Computing time: >12h (dominated by solver)
 - Storage and transmission
- Shared memory hardware
- Use of standard solvers

Approach I (the probably most popular one)



Approach II: Model-based speed-up strategies



Characteristics and dimensions of Energy system optimization models

Typical model dimensions

Time

Planning horizon

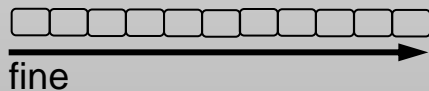
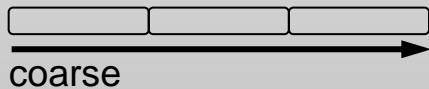
Operation

Investment

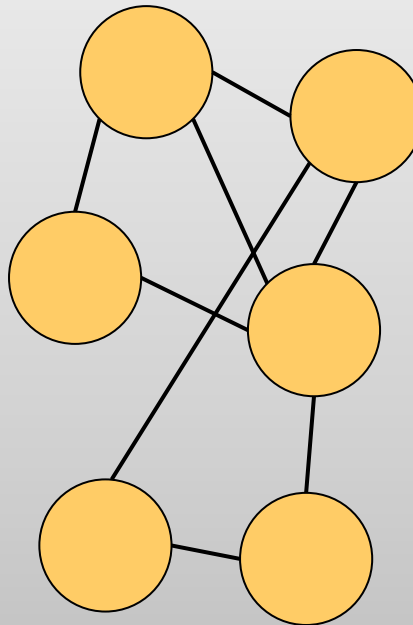
Short term

Long term

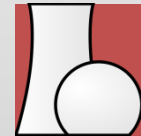
Discretisation



Regions



Technology



Linking variables & constraints

Storage energy
balance:

$$p_{s+}(t, n, u_s) - p_{s-}(t, n, u_s) - p_{ls}(t, n, u_s) = \frac{E_s(t, n, u_s) - E_s(t-1, n, u_s)}{\Delta t}$$

$$\forall t \in T; n \in N; \forall u \in U_s; U_s \subset U$$

p_{s+}/p_{s-} : storage charge/discharge power
 p_{ls} : storage self-discharge (losses)
 E_s : stored energy
 U_s : set of storage facilities

DC power flow:

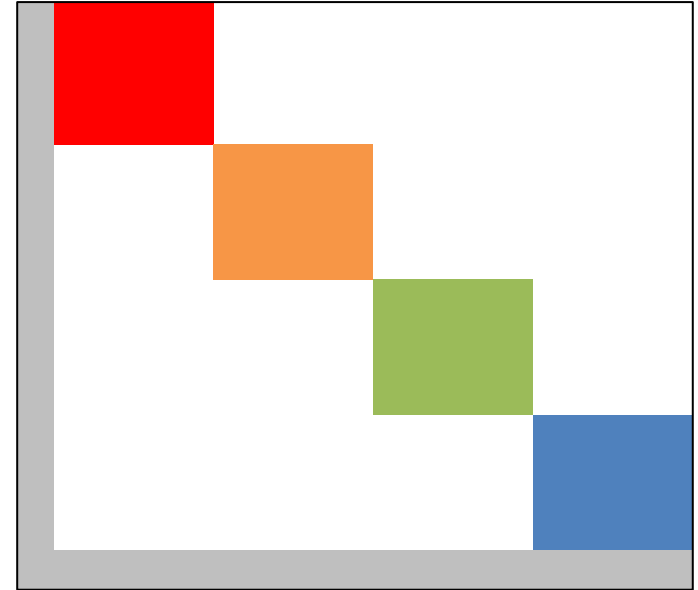
$$p_{im}(t, n) - p_{ex}(t, n) - p_{lt}(t, n) = \sum_n B(n, n') \cdot \theta(n', t)$$

$$\forall t \in T; \forall n \in N$$

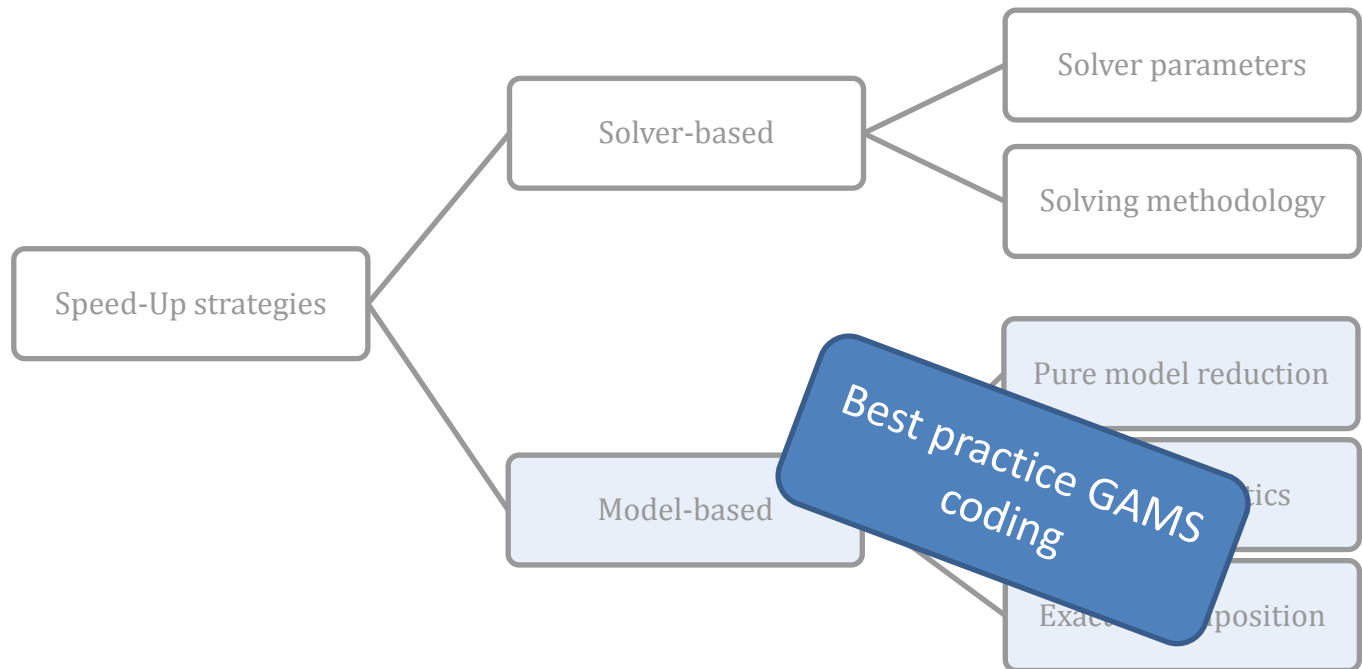
$$p_{f+}(t, l) - p_{f-}(t, l) = \sum_l \sum_n B_{diag}(l, l') \cdot K^T(l, n) \cdot \theta(n, t)$$

$$\forall t \in T; \forall l \in L$$

p_{im}/p_{ex} : power import/export
 p_{lt} : transmission losses
 p_{f+} : active power flow along/against line direction
 $/p_{f-}$:
 θ : voltage angle
 B : susceptances between regions
 B_{diag} : diagonal matrix of branch susceptances
 K : incidence matrix

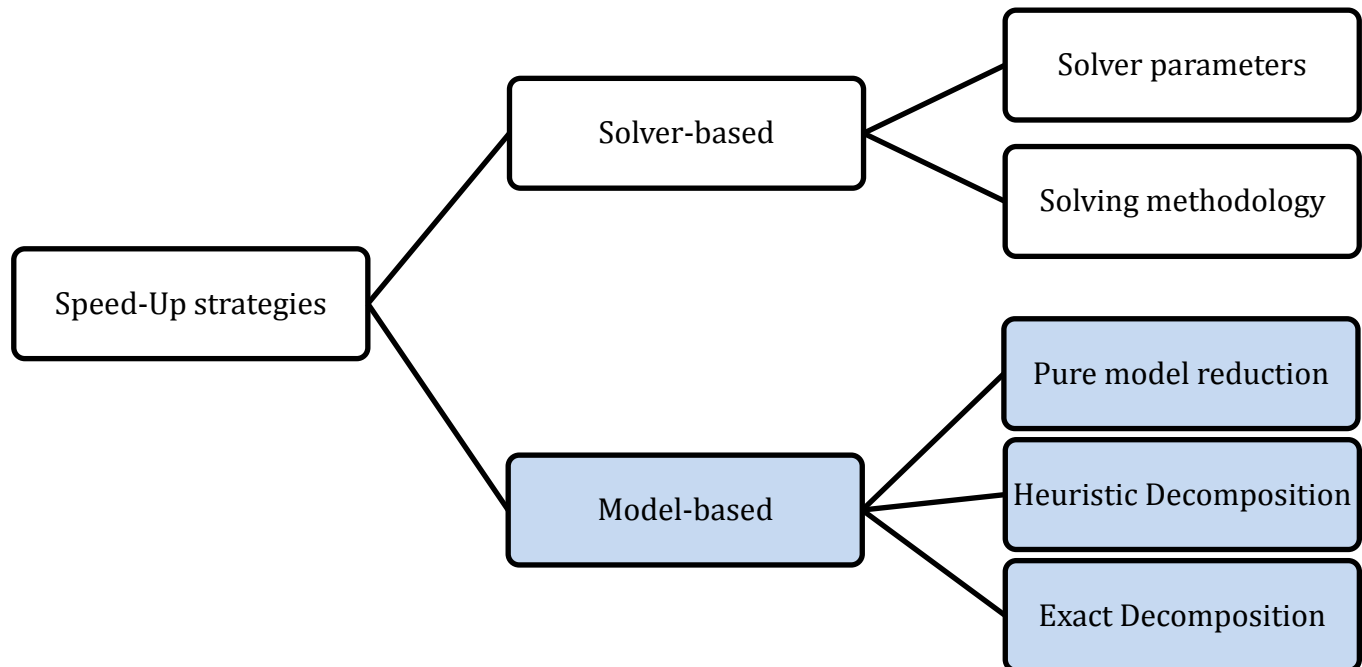


„Low Hanging Fruits“



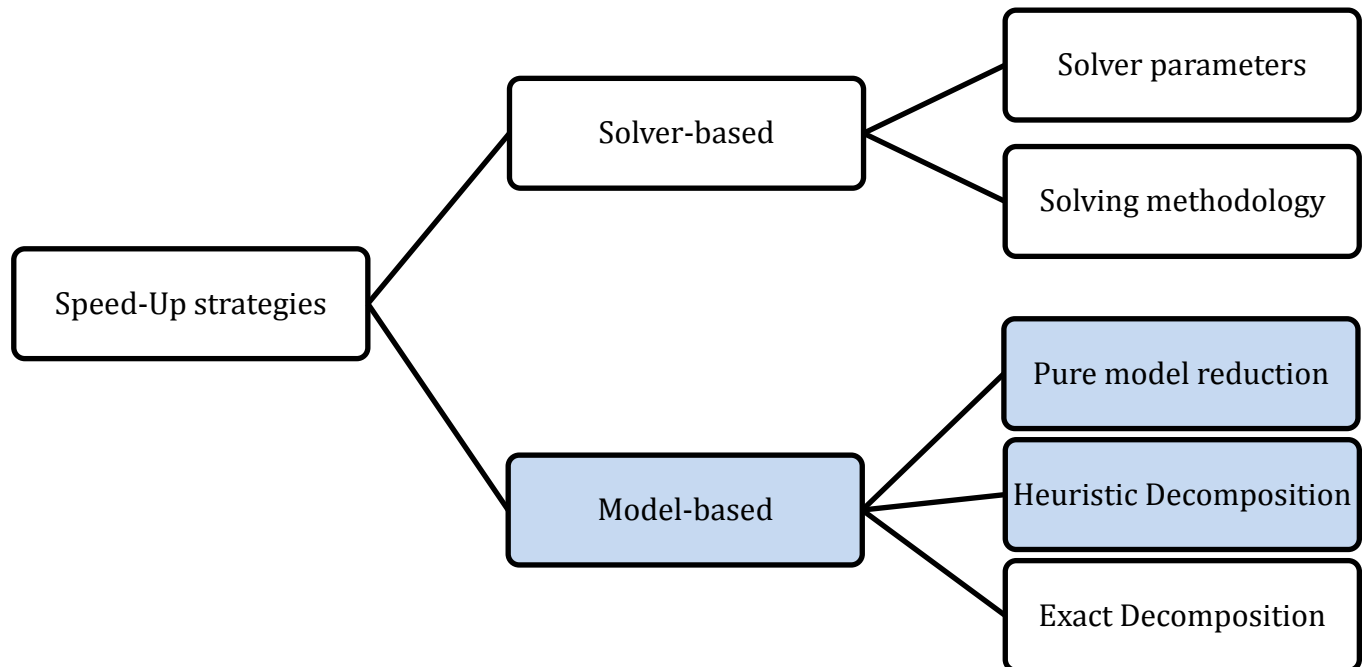
- Selection of measures (also useful for decrease memory need):
 - Input data should not differ much in its order of magnitude
 - Index order influences computing time
 - Useful, but not necessarily faster
 - Assignment statements with a different set order can be faster
 - It can be better to place large index sets at the beginning
 - Use of “option kill” , e.g. for long time-series input parameters saves memory
 - Abundant use of “Dollar Control over the Domain of Definition”
 - Consistent (and limited) use of defined variables
 - Avoidance of the consideration of technologies providing the same service at the same costs
 - Consideration of alternative formulation of model constraints (dense vs. sparse)
- Helpful references: “Speeding up GAMS Execution Time”
by Bruce A. McCarl <https://www.gams.com/mccarl/speed.pdf>

Model-based speed-up strategies

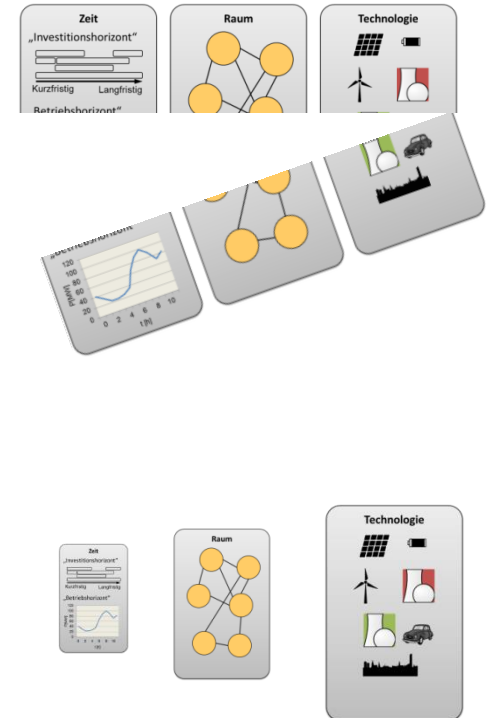
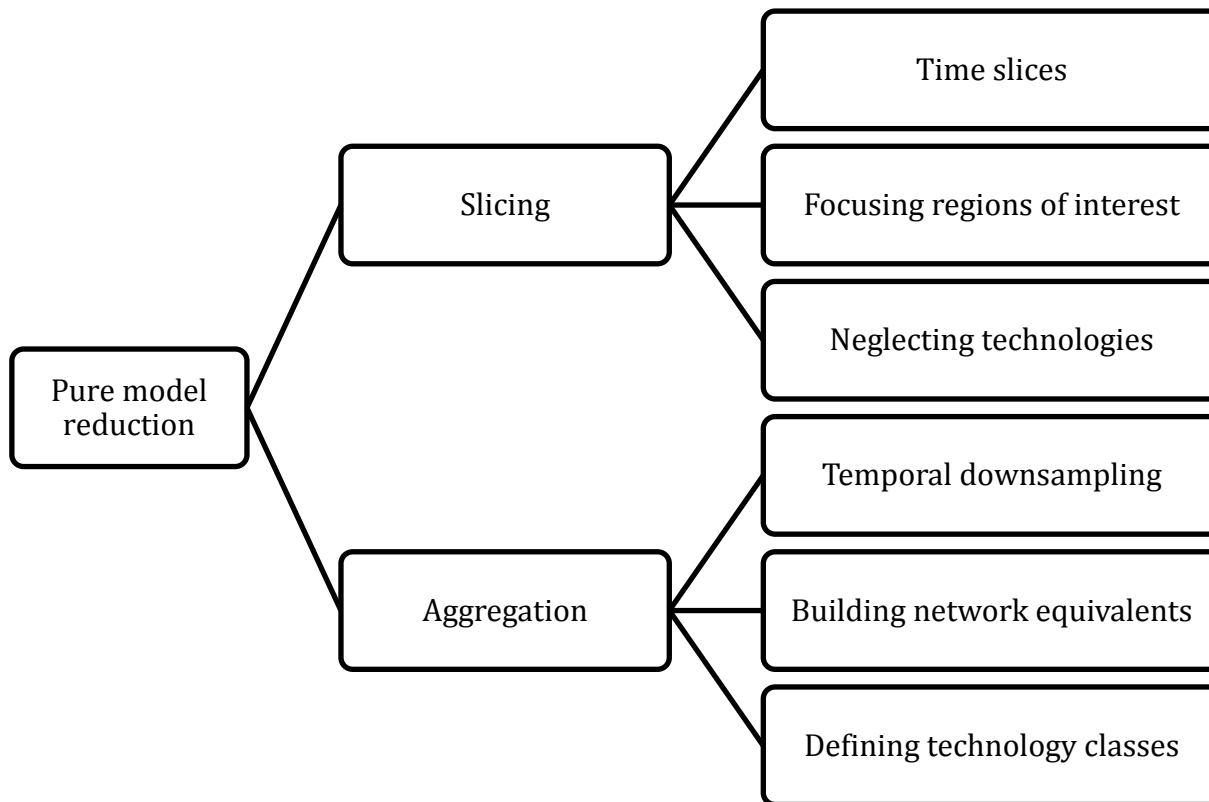


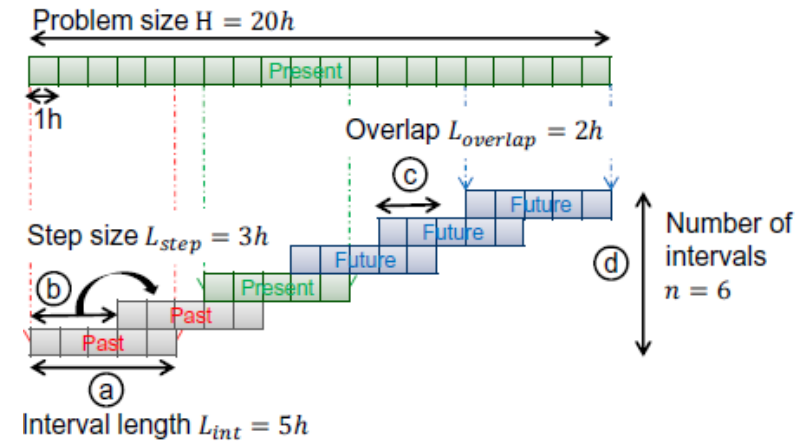
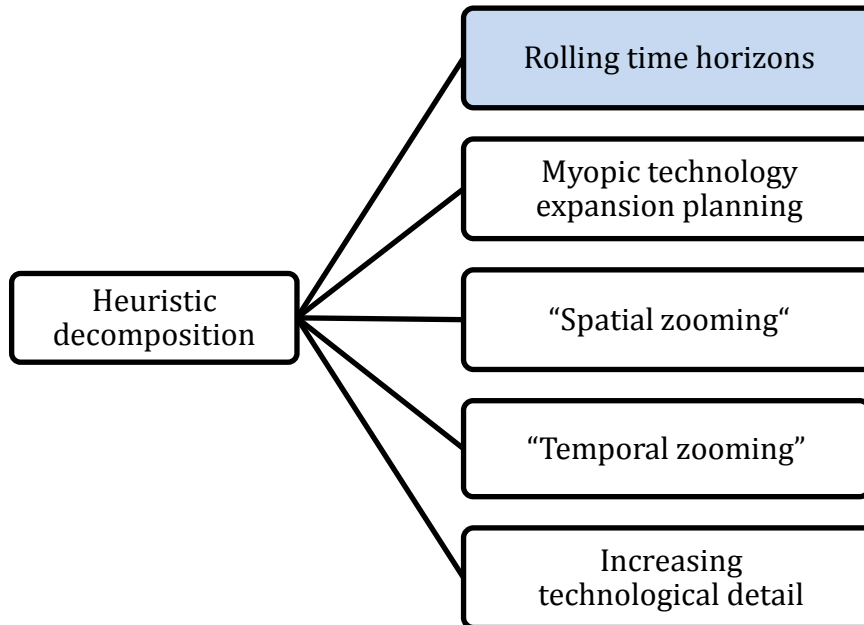
Authors	Math. problem type	Descriptive problem type	Decomposed model scale	Decomposition technique
Alguacil and Conejo [56]	MIP/NLP	Plant and grid operation	Time, single sub-problem	Benders decomposition
Amjady and Ansari [57]	MIP/NLP	Plant operation		Benders decomposition
Binato et. al [58]	MIP/LP	TEP		Benders decomposition
Esmaili et. al [59]	NLP/LP	Grid operation		Benders decomposition
Flores-Quiroz et. al [60]	MIP/LP	GEP	Time, 1-31 sub-problems, sequentially solved	Dantzig-Wolfe decomposition
Habibollahzadeh et. al [61]	MIP/LP	Plant operation		Benders decomposition
Khodaei et. al [62]	MIP/LP	GEP-TEP	Time, 2 sub-problem types, sequentially solved	Benders decomposition
Martinez-Crespo et. al [63]	MIP/NLP	Plant and grid operation	Time, 24 sub-problems, sequentially solved	Benders decomposition
Roh and Shahidehpour [64]	MIP/LP	GEP-TEP	Time, up to 10 · 4 sub-problems, sequentially solved	Benders decomposition and Lagrangian Relaxation
Virmani et. al [65]	LP/MIP	Plant operation	Technology (generation units), up to 20 sub-problems, sequentially solved	Lagrangian Relaxation
Wang et. al [66]	LP/MIP	Plant and grid operation	Space, 26 sub-problems, sequentially solved	Lagrangian Relaxation
Wang et. al [67]	MIP/NLP	Plant and grid operation	Scenarios and time, 10 · 4 sub-problems, sequentially solved	Benders decomposition
Wang et. al [68]	LP	Plant and grid operation	Technology (circuits) and time (contingencies), 2 sub-problem types, sequentially solved	Lagrangian Relaxation and Benders decomposition

Model-based speed-up strategies

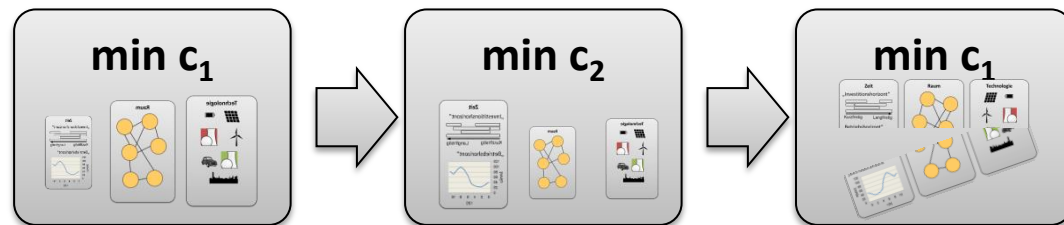


Types of model reductions in ESM

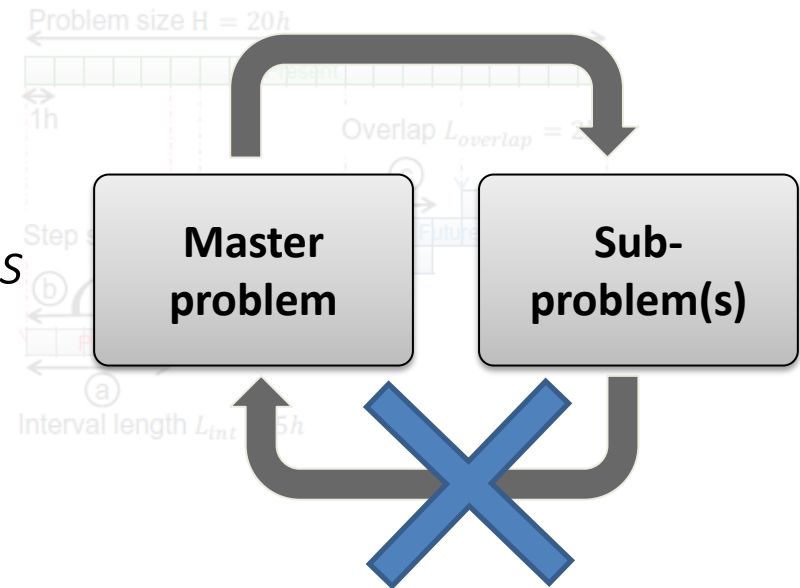




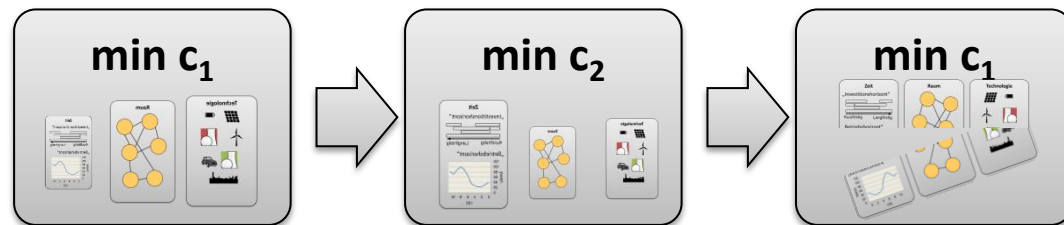
Stepwise solving
reduced models

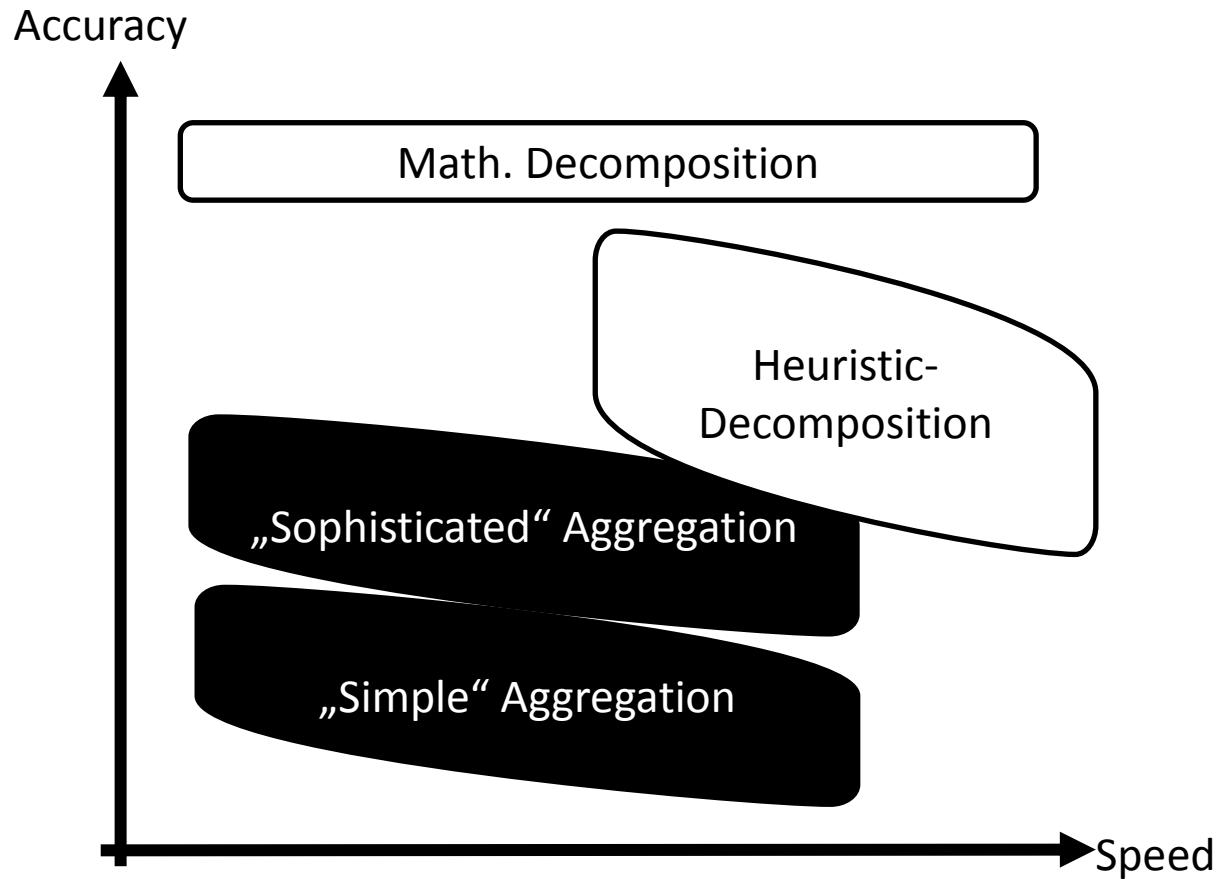


„Decomposition which is **similar** to exact decomposition approaches that are stopped within the first iteration“



Stepwise solving
reduced models

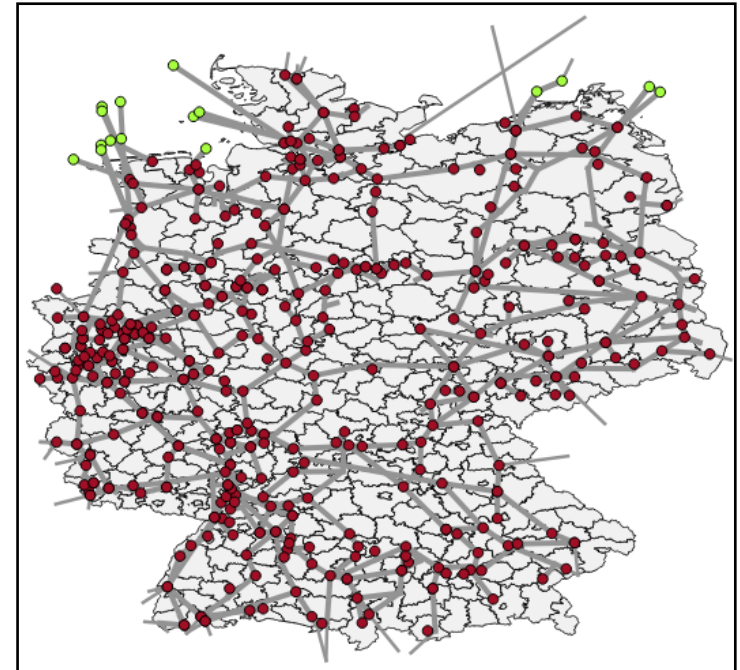




PIPS

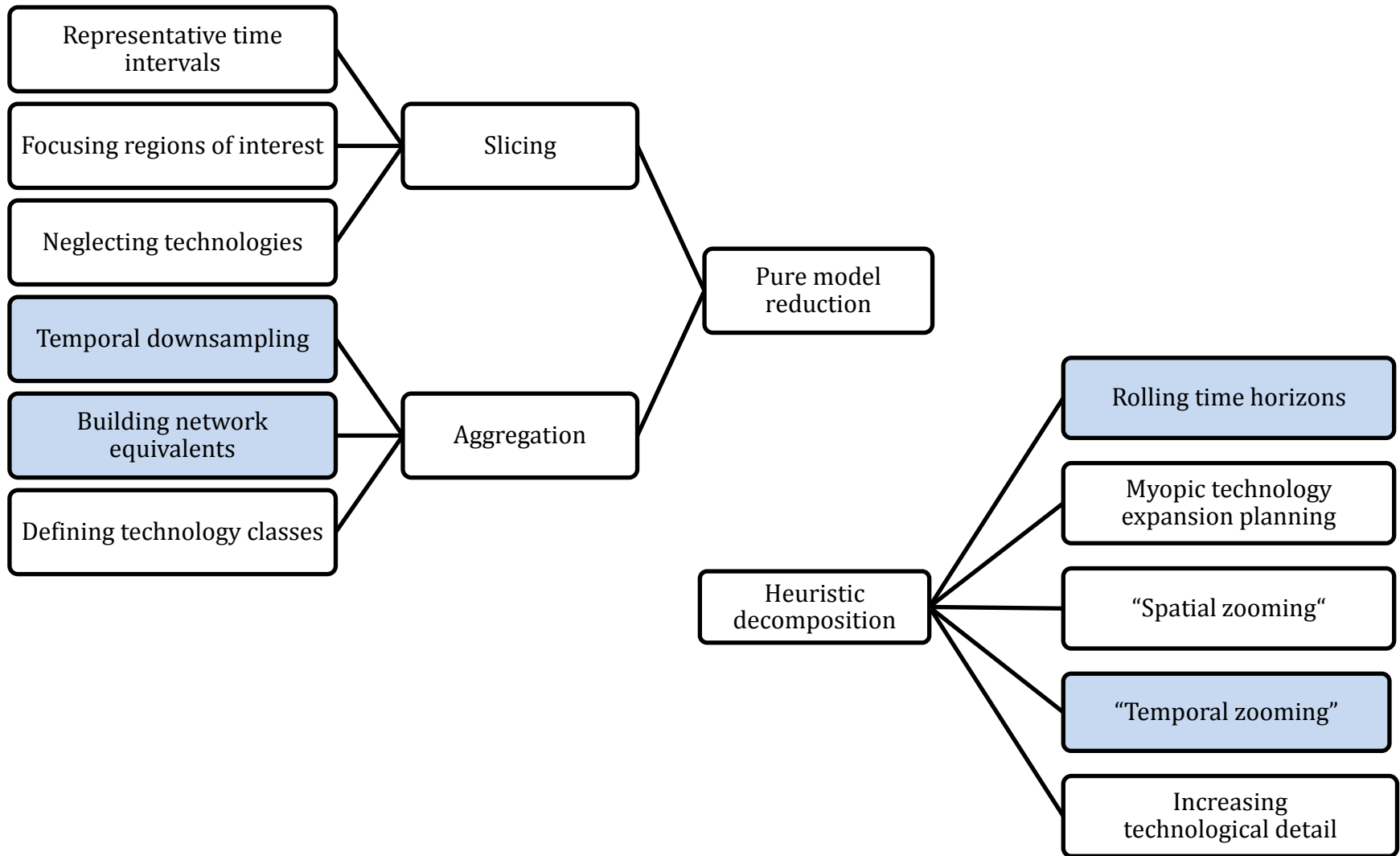
Evaluation methodology

Model name	REMix
Author (Institution)	German Aerospace Center (DLR)
Model type	Linear programming minimization of total costs for system operation economic dispatch / optimal dc power flow with expansion of storage and transmission capacities
Sectoral focus	Electricity
Geographical focus	Germany
Spatial resolution	> 450 nodes (reference model)
Analyzed year (scenario)	2030
Temporal resolution	8760 time steps (hourly)



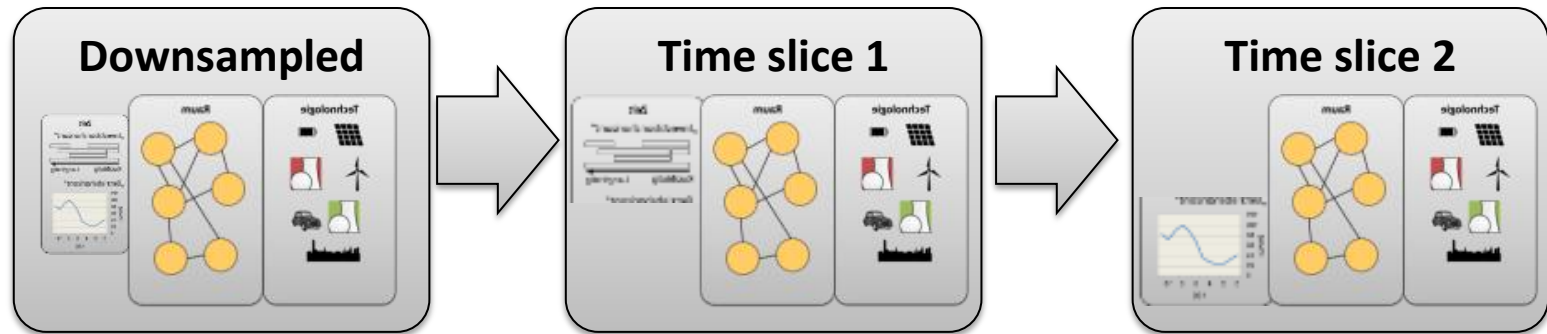
Solver	Commercial
Algorithm	Barrier
Cross-over	Disabled
Max. parallel barrier threads	16
Scaling	Aggressive

Evaluated speed-up approaches

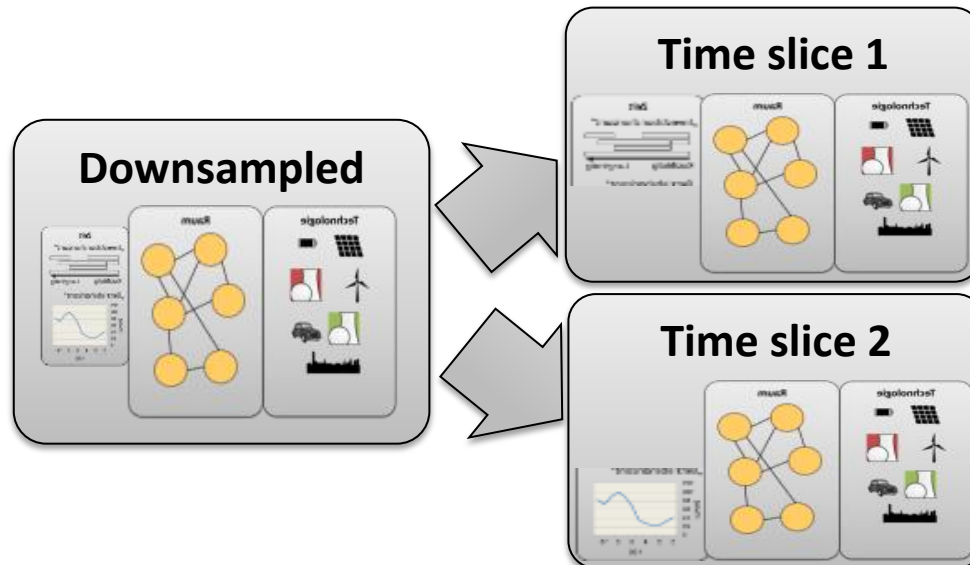


„Temporal zooming“ implementations

- Sequential



- Parallel (using GAMS's grid computing facility)



„Temporal zooming“ implementations

- Sequential



Parallelization limited

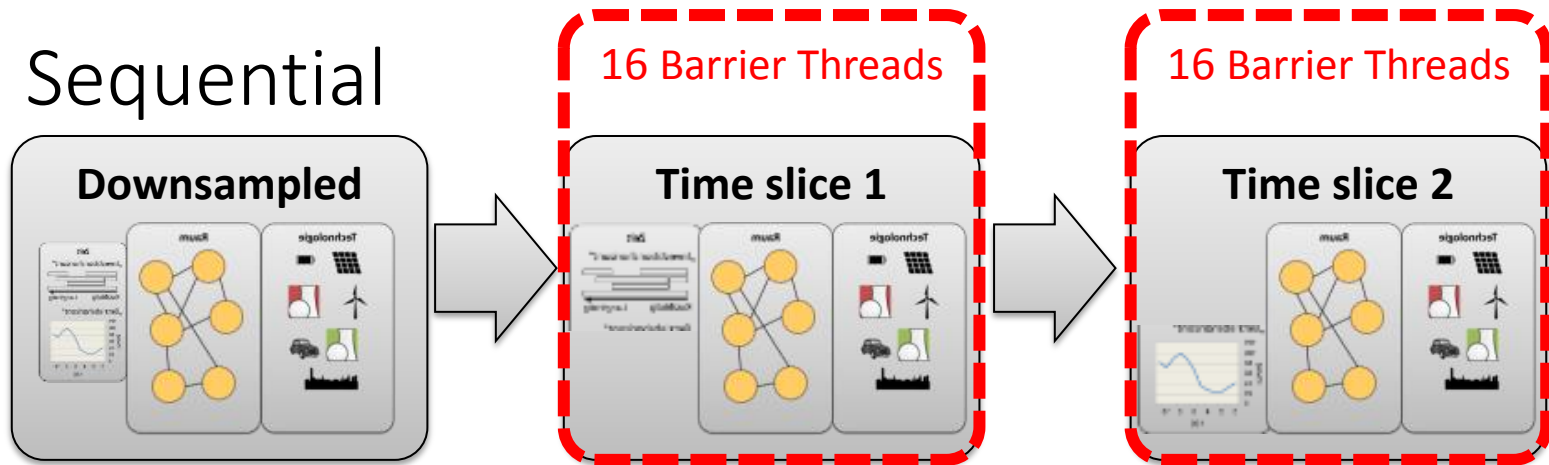
- Parallel (using GALVIS's grid computing facility)

due to shared memory!

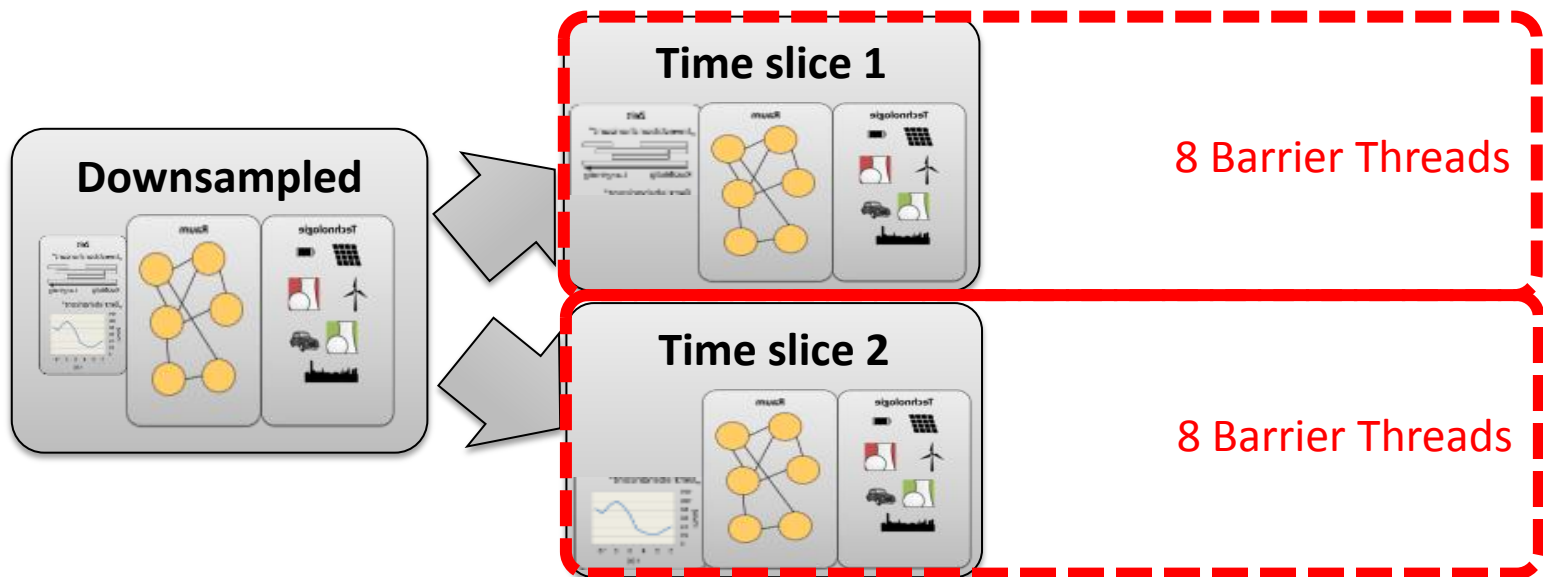


„Temporal zooming“ implementations

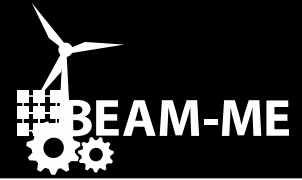
- Sequential



- Parallel (using GAMS's grid computing facility)



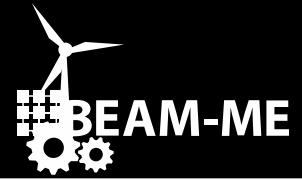
Speed-up approach parameters



Speed-up approach	* w	* wo	Parameter	
			Name	Evaluated range
Spatial aggregation	✓	✓	number of regions (clusters)	{1, 5, 18, 50, 100, 150, 200, 250, 300, 350, 400, 450, 488}
Temporal Downsampling	✓	✓	temporal resolution	{1, 2, 3, 4, 6, 8, 12, 24, 48, 168, 1095, 4380}
Rolling horizon dispatch	✗	✓	number of intervals overlap size	{4, 16, 52} {1%, 2%, 4%, 10%}
Temporal zooming (sequential)	✓	✗	number of intervals resolution of down-sampled run	{4, 16, 52} {4, 8, 24}
Temporal zooming (grid computing)	✓	✗	number of intervals resolution of down-sampled run number barrier threads number of parallel runs	{4, 16, 52} {4, 8, 24} {2, 4, 8, 16} {2, 4, 8, 16}

*w/wo: expansion of storage and transmission capacities

Speed-up approach parameters

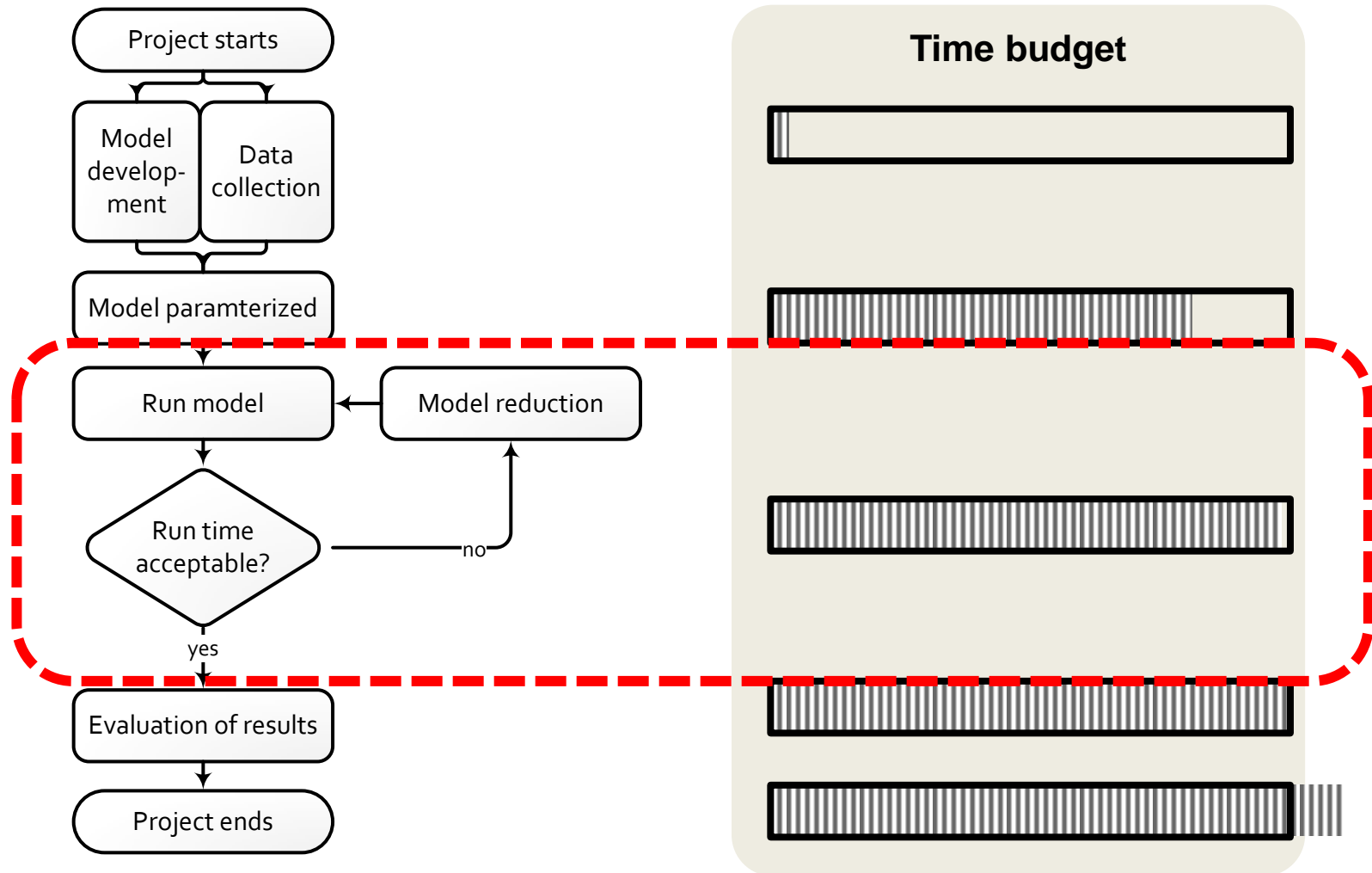


Speed-up approach	*w	*wo	Parameter	
			Name	Evaluated range
Spatial aggregation	✓	✓	number of regions (clusters)	{1, 5, 18, 50, 100, 150, 200, 250, 300, 350, 400, 450, 488}
Temporal Downsampling	✓	✓	temporal resolution	{1, 2, 3, 4, 6, 8, 12, 24, 48, 168, 1095, 4380}
Rolling horizon dispatch	✗	✓	number of intervals overlap size	{4, 16, 52} {1%, 2%, 4%, 10%}
Temporal zooming (sequential)	✓	✗	number of intervals resolution of down-sampled run	{4, 16, 52} {4, 8, 24}
Temporal zooming (grid computing)	✓	✗	number of intervals resolution of down-sampled run number barrier threads number of parallel runs	{4, 16, 52} {4, 8, 24} {2, 4, 8, 16} {2, 4, 8, 16}

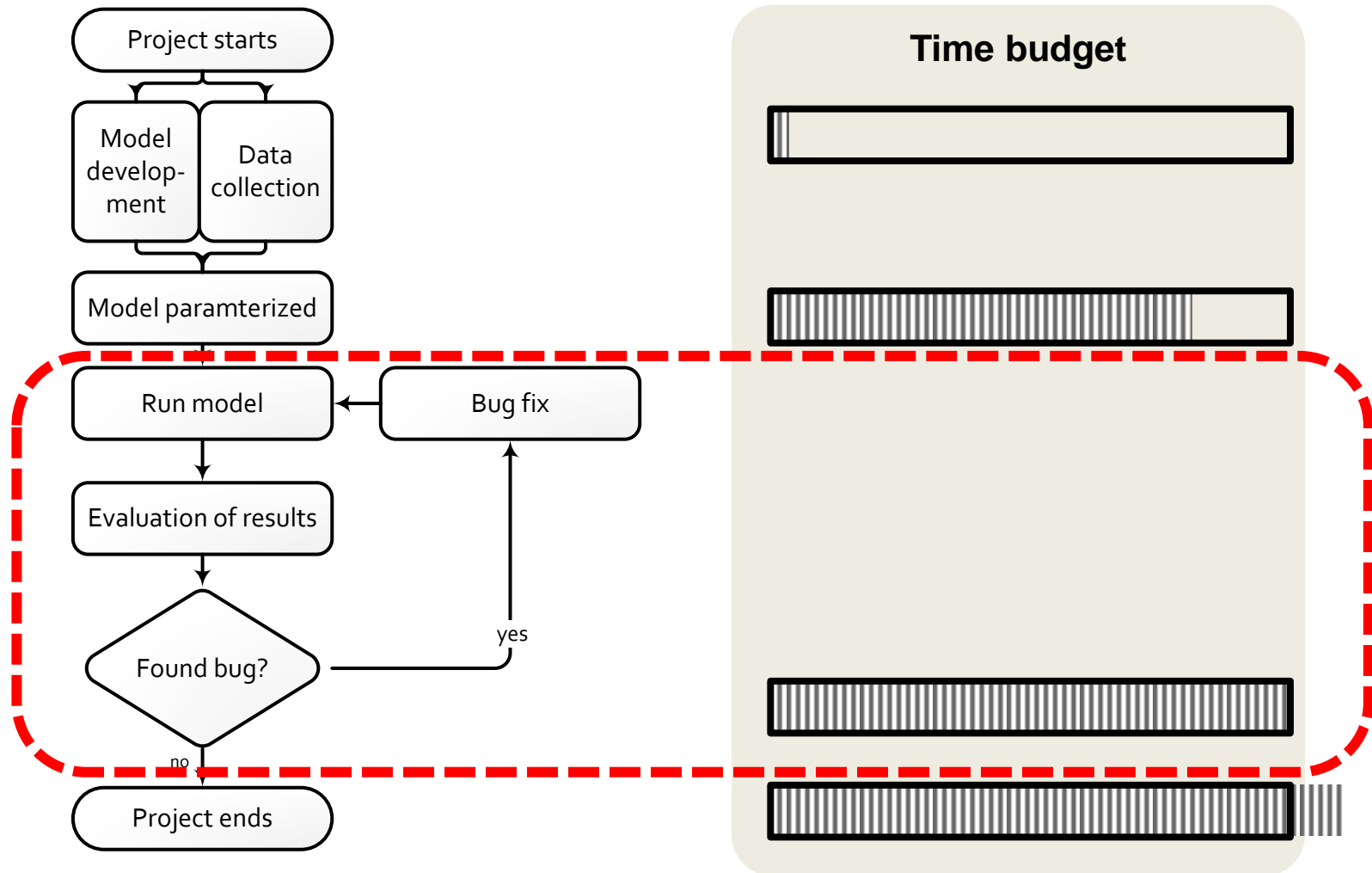
*w/wo: expansion of storage and transmission capacities

Results

Everyday's Energy Systems Analysis...

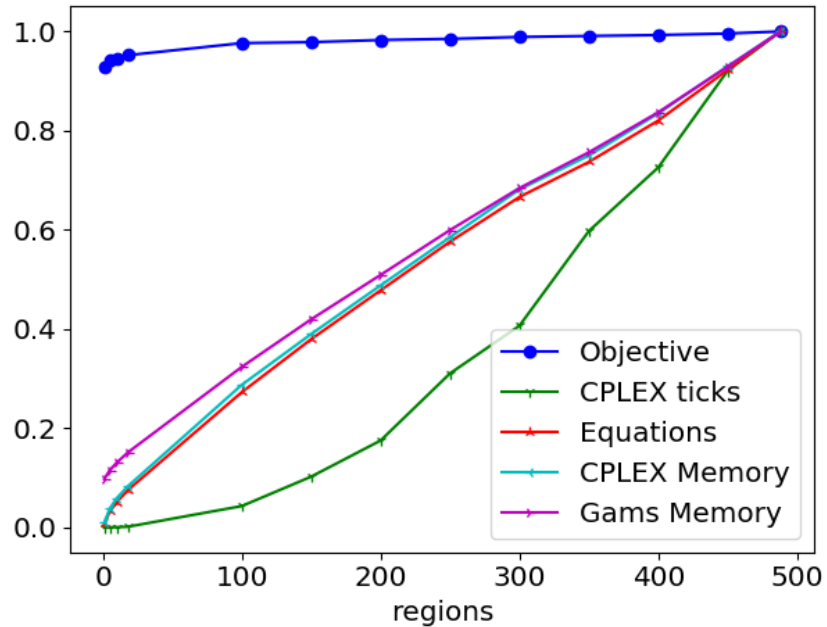


Everyday's Energy Systems Analysis...

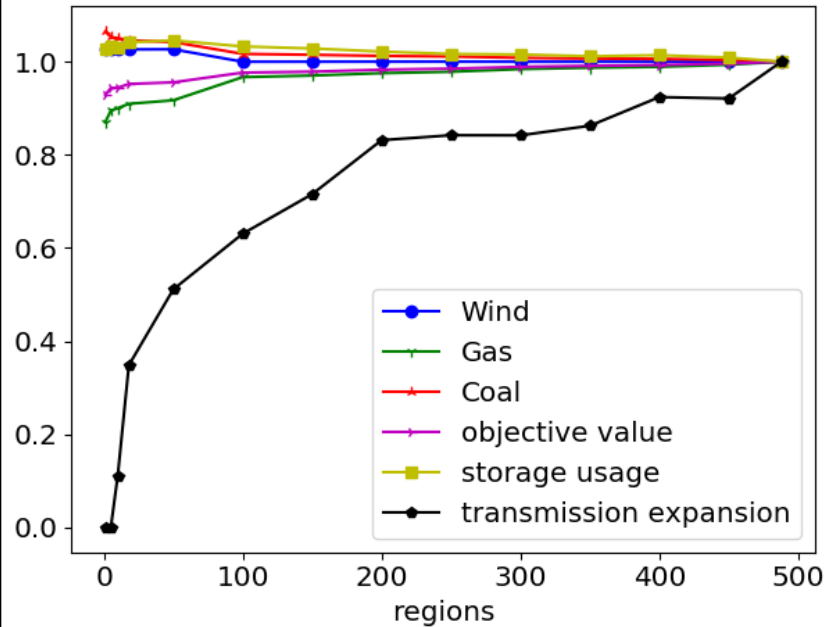


Spatial aggregation

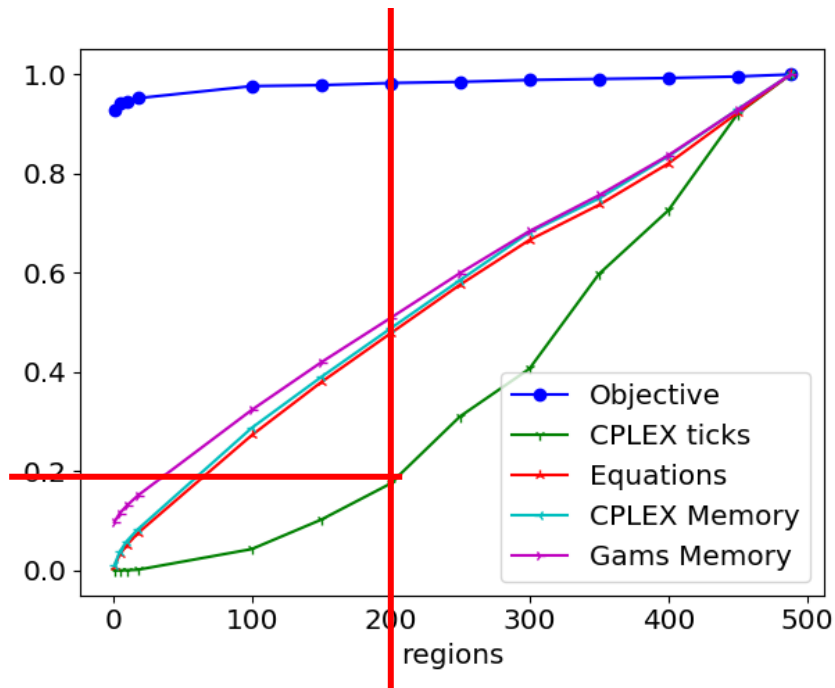
Performance



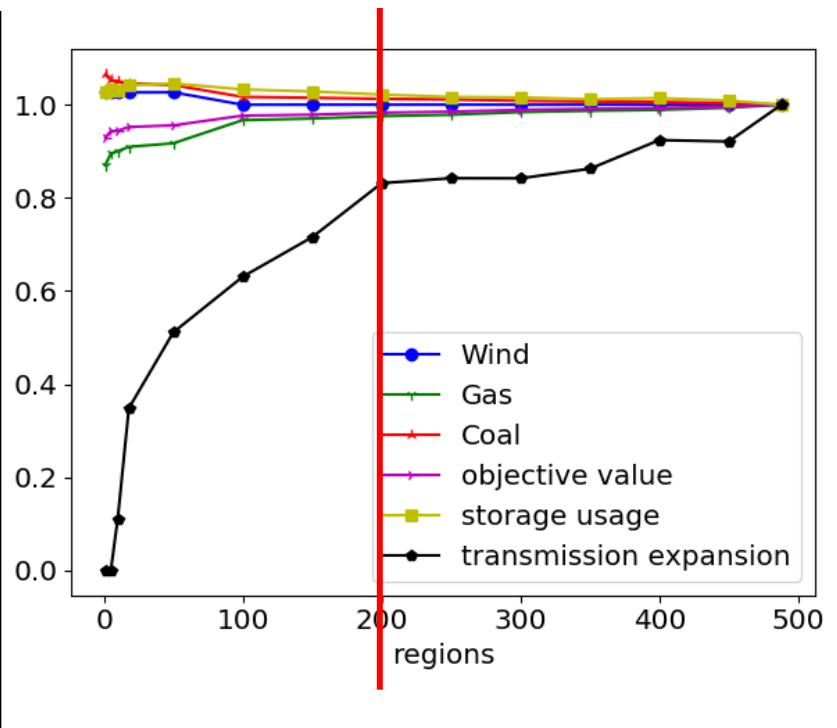
Accuracy



Performance



Accuracy

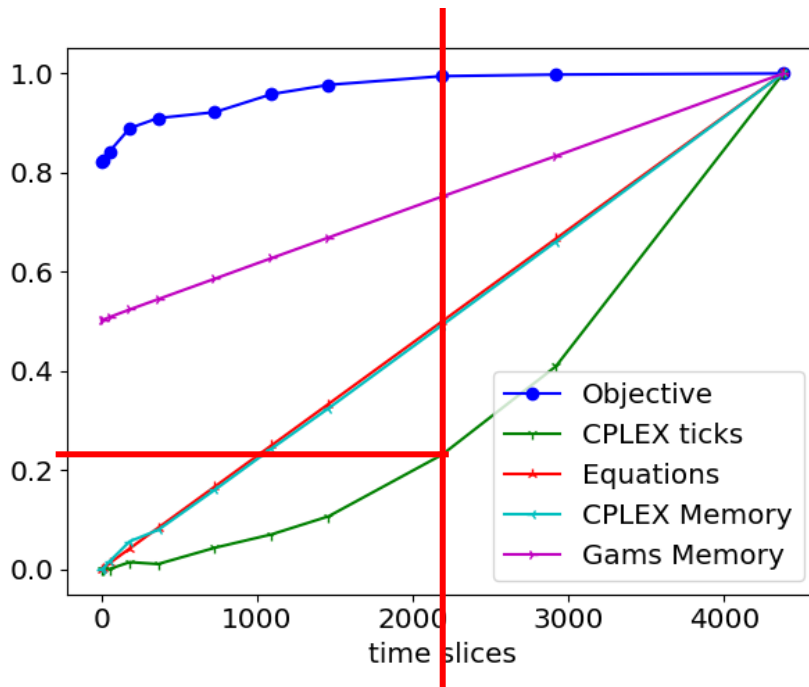


1 a) Speed-up factor: ≈ 5

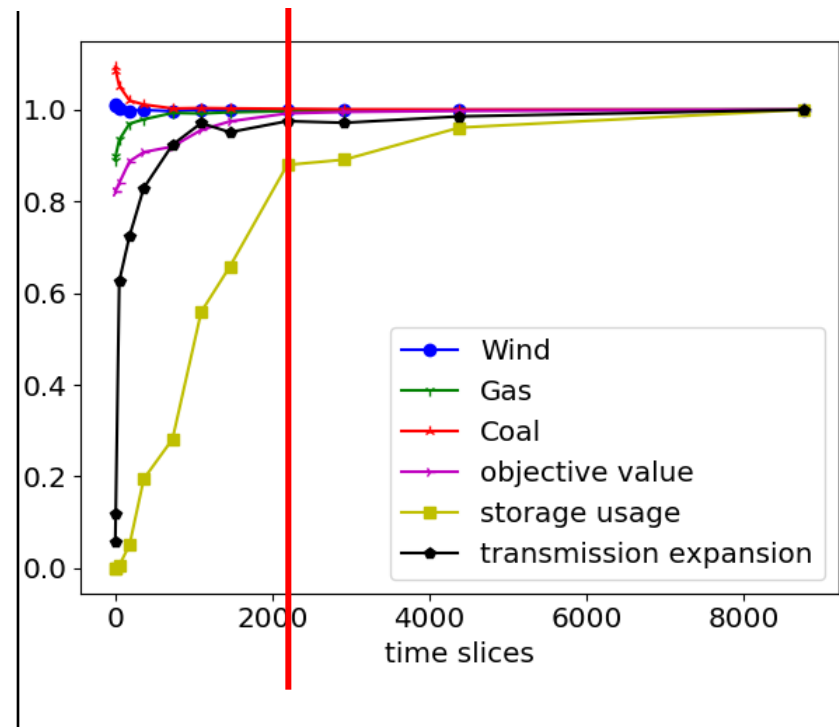
2 a) Accuracy error mainly $< 10\%$ (grids: $\approx 20\%$)

Temporal downsampling

Performance



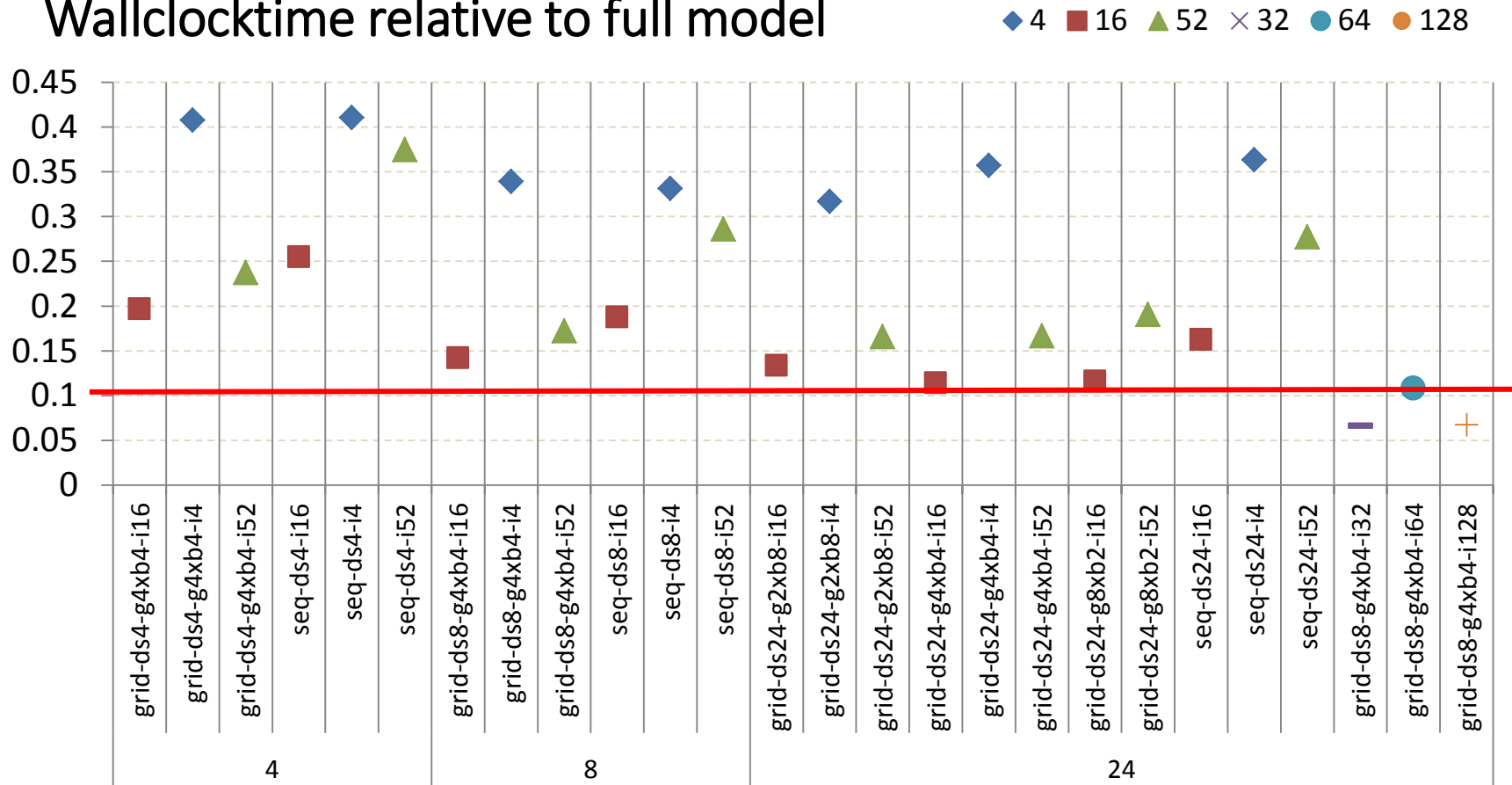
Accuracy



1 b) Speed-up factor: ≈ 5

2 b) Accuracy error mainly $< 10\%$ (storage: $\approx 20\%$)

Wallclocktime relative to full model



3) Speed-up factor: >10 reachable, at least >2

Conclusions

Which **speed-up** is possible
using measures that can be influenced
by „normal“ model developers?

10!

- 4 speed-up strategies evaluated
- 2 slightly different models
- Aggregation
 - 1) Speed up ≈ 5
 - 2) Accuracy error $< 10\%^*$
- Temporal zooming
 - 3) Speed up ≈ 10

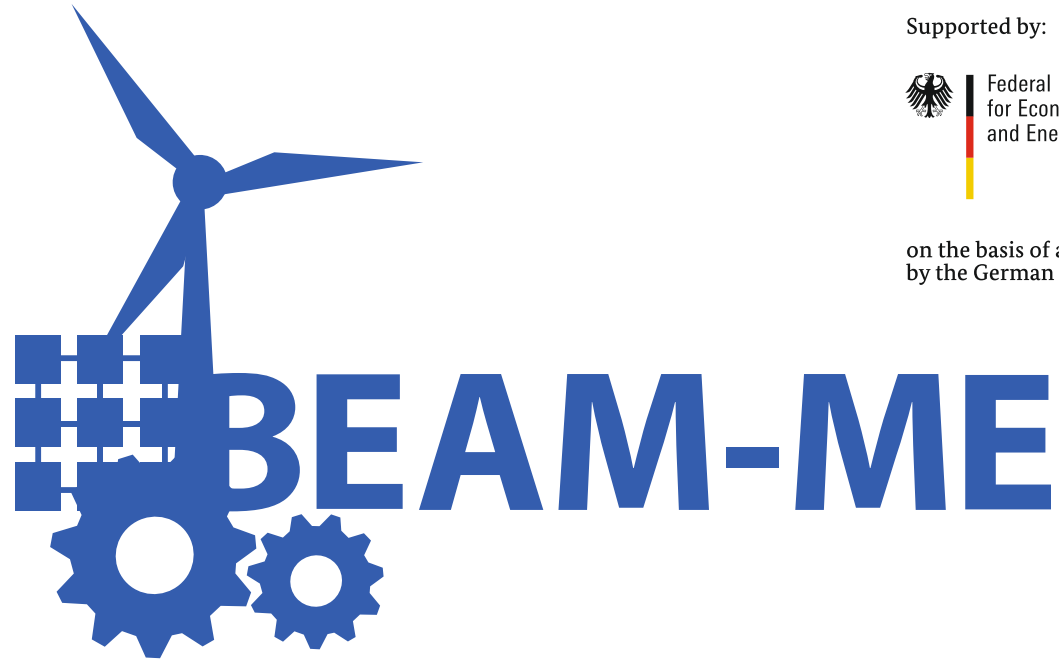
*except of indicators related to aggregated dimension

Project BEAM-ME

Thank you!

Contact

Karl-Kiên Cao
Deutsches Zentrum für Luft- und
Raumfahrt – Energiesystemanalyse
Tel. +49711 6862-459
karl-kien.cao@dlr.de



Supported by:



Federal Ministry
for Economic Affairs
and Energy

on the basis of a decision
by the German Bundestag

A PROJECT BY



Deutsches Zentrum
für Luft- und Raumfahrt
German Aerospace Center

