Electricity and district heating in GreenREFORM

Technical documentation

Rasmus K.S. Berg* Lucas E. Shine†

August 2026

Abstract

Details the relevant economics of the electricity and district heating sector in the GreenREFORM model, as implemented on the Climate Outlook 2025 (KF25) data. We will refer to the model as the "energy module/system" along the way (shorter). The model represents every Danish generating plant individually, together with the twelve foreign price areas Denmark trades with, and solves for hourly electricity and district heat prices in each year from 2020 to 2050 as the equilibrium of a square system of first-order conditions. Sections 2–6 set out the economics plant type by plant type and then close the model; Section 8 describes how the raw data are aggregated into the version that is solved; Section 9 the solution procedure; Section 10 what the solved baseline looks like and how it compares with the projection it is built from. Appendix A maps every symbol used here to its name in the GAMS implementation. Section 1.3 describes what is distributed with the model and what a user can and cannot reproduce from it.

Contents

*University of Copenhagen, Øster Farimagsgade 5, 1353 Copenhagen K, Denmark. E-mail: rasmus.kehlet.berg@econ.ku.dk.

†DREAM Group, GreenREFORM. E-mail: lws@dreamgruppen.dk

1 Introductory remarks

The following describes the electricity and heating sector in the GreenREFORM model. The wider modelling framework it belongs to, and the approach to integrating a bottom-up energy system with a top-down general equilibrium model, are described in Berg and Eskildsen [2019] and Berg and Eskildsen [2021a]; the storage formulation draws on Berg and Eskildsen [2021b].

Figure 1 is the model on one page. Each electricity area has a price that clears supply against demand and net exports; each Danish price zone additionally has a district heat market, and the two are coupled because the same plants sell into both. Everything in Sections 2–6 is an elaboration of one of the boxes in that figure.

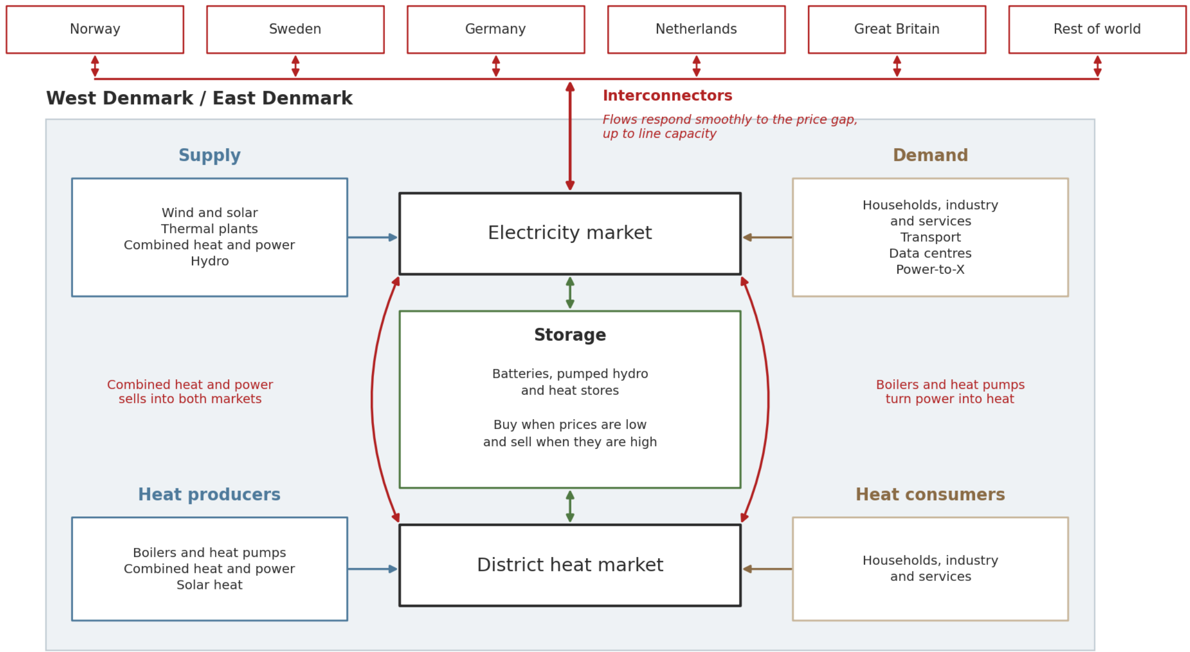


Figure 1: Market structure of one electricity area and the trade around it. The Danish price zones each carry a district heat market alongside the electricity market; the foreign areas are modelled on the electricity side only.

1.1 Notation

The model documentation uses regular math/LaTeX syntax to ease readability; appendix A includes a dictionary for LaTeX-to-GAMS notation. Here is a brief introduction to important indices and notation:

- **Time:** The energy system model covers years 2020–2050 (index t). Within each year, the model either solves for 8760 hours (index h) or aggregate states (index k); see section on aggregation for more. Where the distinction does not matter, we write h and let n_h denote the number of full hours that the state represents ($n_h = 1$ in the full hourly model).
- **Energy markets/goods:** While a range of fuels can be used as inputs (index b), the model currently models demand for (D) and supply of (S) electricity (E), district heat (DH), and hydrogen (H).
- **Geography:** Is generally indexed g, but the granularity of the index depends on the context; for electricity areas, the relevant geographic resolution is price zones (e.g. Denmark

East and West), whereas district heat areas are usually much smaller.[1] If the context is not clear, we will use g_E or g_{DH} to indicate the relevant index. If we need to reference a geographic location of e.g. a specific plant i, we use notation $g(i)$.

- **Producers:** Any energy producer in the model is defined by its producer id i. The subset of producers in a specific market is defined by i_E for electricity producers, i_{DH} for district heat producers etc.
- **Fuels:** Base fuels b include things like coal, oil, straw, wood chips etc. Fuel types f define linear combinations of base fuels. We use $f(i)$ to denote the unique fuel type that a producer i maps to.
- **Consumers:** We use c to index consumers in general. If it is not clear from the context what energy good they consume, we explicitly reference e.g. c_E for electricity. If an energy-producing plant i is also a consumer, we use $c(i)$ to denote the corresponding consumer index.
- **Technology:** Energy producing plants belong to a unique technology type (we have roughly 15–20). They are fundamentally different; we deal with each of them in turn.
- **Emissions:** The emission index m includes SO_2, NO_x, CO_2, CH_4, and N_2O. In the main version, we aggregate CH_4, N_2O, and CO_2 emissions to CO_2 equivalents.
- **Variation types** v identify a set of variations within a year; $v(i)$ identifies the variation type that producer i is mapped to (similar with $v(c)$ for consumers).
- **Resource zones** r identify domestic resources that are available in constrained quantities and are therefore priced within the model rather than at an exogenous world market price. Currently, this covers waste, hazardous waste, and biogas; see section 5. We write $r(i)$ for the resource zone that plant i belongs to, where relevant.
- **Bidding types:** Plants are split into two bidding types. Type-0 plants ($i \notin \mathcal{B}_1$) bid against the spot price in the hour. Type-1 plants ($i \in \mathcal{B}_1$) are on fixed feed-in contracts and do not respond to the spot price; see section 2.1.

1.2 Model versions and states

The GAMS model is assembled from a set of modules (`*.module.gms` files), each of which can be written in one of several *states*. A state selects which groups of variables are endogenous, which are exogenous, and which blocks of equations are imposed. The Python layer (`pyGRE.models`) assembles a named GAMS model definition from a chosen combination of module states. In practice, the two dimensions of variation that matter for reading this document are:

- **Time representation:** the modules `exoInputs`, `HY`, `HYS`, `ES` and `HS` have a `chronological` and a `clusterAvg` state. The former is used when hours are ordered in time, the latter when hours are clustered non-chronologically and transitions are described by a Markov matrix (section 8.5).
- **Closure:** the `equilibrium` and `demand_DCIIA` modules have states that switch between exogenous and endogenous prices, and between exogenous and endogenous yearly demand levels. These are used to walk the model into equilibrium; see section 9.

The model is solved as a square system of equations (GAMS `CNS`); there is no optimisation at the level of the full model. All optimisation is done analytically at the level of the individual plant, consumer, or transmission line, and only the resulting first-order conditions enter the GAMS model.

1.3 What is distributed, and what it is built from

The model is built from two sources. The plant population, their technical characteristics, their capacity paths and the hourly variation profiles come from the Ramses input data maintained by the Danish Energy Agency (see `ramsesr.pdf` in the documentation folder for the input

[1] Disaggregated data uses roughly 80–90 for Denmark only.

format). Demand levels, fuel and quota prices, tax rates and the capacity projections come from the published Climate Outlook 2025 workbooks. A pipeline of notebooks reads those sources, harmonises them, aggregates them along the five dimensions of Section 8, and writes a database that the model can be solved on.

That pipeline is not distributed. It reads Ramses input data that the Danish Energy Agency does not release, so shipping it would not help a user who cannot supply its inputs, and shipping its inputs is not ours to do. What is distributed is its output, which is enough to solve the model, run shocks against it, and reproduce everything in Section 10:

- `data/KF25_clusterAvg/` and `data/KF25_chronological/` hold the aggregated databases the model is solved on, named `db_<geography>_<hours>` for the geographical and hourly aggregations of Section 8: `db_aggEA_aggHours` is the twelve-area, 120-state version that this note reports on. These are the output of the data pipeline and the input to `1_Baseline.ipynb`.
- `data/KF25_clusterAvg/` additionally holds, for each configuration, the solved baseline (`db_baseline_*.pkl`), the GAMS checkpoint it was solved to (`cp_baseline_*.g00`), the assembled model object (`m_baseline_*.pkl`) and the starting price vector (`price_dict_*.pkl`) described in Section 9.3. A user who only wants to run shocks can start from these and never solve a baseline.
- `data/KF25_clusterAvg/modeldb_<geography>_<vintage>/` holds the working database each model object was built on. A pickled model does not carry its database; it records the path and reads it back on load, so these folders have to travel with the model objects. They are written by `1_Baseline.ipynb` and are not otherwise read.
- `makeData/RawData25/` holds the published Climate Outlook workbooks. These are not used to build the model — that has already happened — but the notebooks read them to compare the model against the projection, as in Section 10.2.
- `pyGRE/` holds the model source: the GAMS modules in `gmsModels/`, the Python layer that assembles them into a model definition, the aggregation routines, and the diagnostic and plotting tools used for the figures in this note.

The practical consequence is that the plant population, the fuel types, the tax rules, the resource-zone assignments and the mapping of hours to states are fixed in the distributed databases. Anything that changes those has to be done upstream of what is distributed. Everything downstream — prices, capacity paths, tax rates, transmission capacities, demand levels, the smoothing parameters — is a parameter in the database and can be changed in a shock (Section 9.4).

1.4 Running the model

Two notebooks in `makeModels/` do the work, using the Python environment in `virtualEnv.yml`.

- `1_Baseline.ipynb` reads an aggregated database, assembles the model, and walks it into equilibrium through the stages of Section 9. It writes the solved database, the checkpoint, and a refreshed starting price vector. Which configuration it solves is set by three variables at the top: `hour_state` (`clusterAvg` or `chronological`), `aggEA` (`aggEA` or `disaggEA`) and `kf` (`KF25` or `KF26`, the overlay of Appendix B).
- `2_Analysis.ipynb` does everything that reads a solved model. With `run_shock = True` it re-solves from the baseline checkpoint with parameters perturbed and compares the two solutions, as described in Section 9.4. It then draws the figures — energy balances, prices, trade, supply curves, market clearing, storage behaviour, validation against the projection — from the pair of solved databases. Most of the figures in this note come from it.

Only the baseline and the shock need GAMS with a CONOPT licence. Left at `run_shock = False`, `2_Analysis.ipynb` compares the two solved baselines that are distributed with the model and runs without a licence. Setting `save_figures = True` in the same cell writes every figure it draws to `results/`, one folder per plotting routine, in addition to showing it in the notebook.

Nothing in the distribution records where it was unpacked, so the folder may be placed anywhere and copied between machines. Paths inside the pickled model objects are resolved relative to the folder itself when the location they were written at no longer exists.

A full baseline solve takes on the order of an hour for the `aggEA` configuration and considerably longer for `disaggEA`; a shock re-solved from a checkpoint takes minutes.

2 Energy producing plants

In this section, we go through relevant data, the optimization problem, and the final equations characterizing the different technology types. We note that all energy producers i are mapped to a unique technology type, geographic area, fuelmix, and technology (among other things).

For a subset of plants, the short run optimization problem (within a year) is static, in the sense that each hour can be solved separately. For those where this is not the case, we will present one version for the non-chronological "clustered" model (defined over states k), and one for the chronological "hourly" model (defined over hours h). All plants are modelled with rational expectations; a version with risk (probabilistic uncertainty) is on the to-do list.

2.1 Bidding types

Plants that receive a fixed settlement price (*fast afregningspris*) are not exposed to the hourly spot price: their revenue per MWh is fixed by contract. In the data, these plants are identified by their subsidy type, and collected in the subset $\mathcal{B}_1$ (`bidT1`). For these plants, the contract price enters the marginal cost function as a negative tax (see below), and the spot price is removed from the dispatch decision. Concretely, wherever a type-0 plant compares $p^E_{t,h,g(i)} - c_{t,i}$, a type-1 plant compares $-c_{t,i}$; since $c_{t,i}$ contains the fixed settlement price with a negative sign, the plant effectively dispatches against its contract price rather than the market price. The same substitution applies to the price indices of CHP plants, where only the district heating price is retained. In what follows, we present the type-0 equations and note the type-1 modification where it matters.

2.2 Smoothed dispatch

Nearly every dispatch decision below takes the same form: the plant runs if the price covers its marginal cost, and the model writes that as $\Phi((p-c)/\sigma)$ rather than as an indicator function. The interpretation is that marginal cost is not known exactly — it varies within the plant, within the state, and across the hours the state stands for — and $\sigma_{t,i}$ is the standard deviation of that variation. What the model computes is then the expected share of capacity that runs.

This is worth being explicit about early, because σ is not a numerical convenience that washes out in aggregate. Figure 2 draws the rule for a plant with a marginal cost of 400 DKK/MWh at four values. The baseline sets $\sigma_{t,i} = 50$ DKK/MWh for every plant, and $\sigma_{t,l} = 50$ on transmission lines except those to the rest of the world, where it is 200. At 50, the dispatch share is one half when price equals marginal cost, and the plant is neither off below its cost nor fully on above it: it reaches nine tenths of capacity 64 DKK/MWh above its own marginal cost, and still offers a tenth 64 DKK/MWh below. Summed over a fleet, this is what turns a merit-order staircase into the smooth supply curve of Figure 3.

The smoothing is also what makes the model solvable: a square system of indicator functions has no derivative to hand to a Newton solver. The price of that is a modelling choice that has to be defended on its own terms rather than treated as an approximation error, and the same parameters are the ones the solution procedure is allowed to relax when a stage fails (Section 9.2).

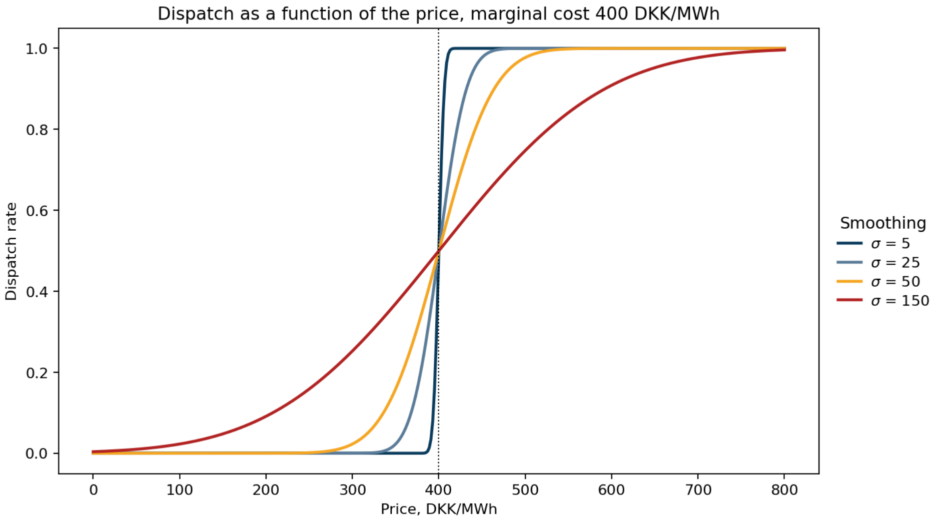


Figure 2: The dispatch rule $\Phi((p-c)/\sigma)$ for a plant with marginal cost 400 DKK/MWh, at four values of the smoothing parameter.

2.3 Condensation plants (CD)

Condensation plants produce electricity relying on a combination of fuels defined by their fuel type $f(i)$. We define approximate marginal costs, $c_{t,i}$, as the sum of fuel costs, emission taxes, other taxes, and other variable operating and maintenance costs ($v_{t,i}$):[2]

$$c_{t,i} = v_{t,i} + \frac{c^{fuel}_{t,f(i)} + c^{r}_{t,i}}{\eta_{t,i}} + T^{M}_{t,i} + T^{E}_{t,i} \tag{1}$$

$$c^{fuel}_{t,f(i)} = \sum_b \alpha_{f(i),b} \times p_{t,b} \tag{2}$$

$$c^{r}_{t,i} = \sum_b \alpha_{f(i),b} \times p^{r}_{t,r(i),b} \tag{3}$$

$$T^{M}_{t,i} = \frac{1}{\eta_{t,i}} \sum_m \left(\mu^{m}_{t,i}\tau^{m}_{t,i}\right), \tag{4}$$

where $\mu^{m}_{t,i}$ is the emission intensity of emission type m, $\tau^{m}_{t,i}$ is the effective emission tax rate on the plant, $\eta_{t,i}$ is the conversion efficiency, $\alpha_{f,b}$ are fuel coefficients such that $\sum_b \alpha_{f,b} = 1$, $p_{t,b}$ are base fuel costs, and $T^{E}_{t,i}$ are other electricity-related taxes/subsidies for the specific plant. The term $c^{r}_{t,i}$ is new relative to the previous version: it is the shadow price of the constrained domestic resources described in section 5, and is zero for plants that are not mapped to a resource zone.

Cost variables $c_{t,i}$, $v_{t,i}$, $T^{M}_{t,i}$, and $T^{E}_{t,i}$ all measure costs in DKK/MWh output electricity, and $c^{fuel}_{t,f}$, $c^{r}_{t,i}$ measure price indices for fuel type f in DKK/MWh of the fuel f.

Emission intensities and regulation. The CO_2e intensity is defined by a plant-specific term $\tilde{\mu}^{CO_2}_{t,i}$ that captures CH_4 and N_2O emissions in CO_2-equivalents, and fuelmix-related CO_2 emissions.

CO_2e emissions are regulated at several different rates depending on the output of a plant (electricity or heat) and to what extent the plant is covered by EU-ETS (or EU-ETS2).[3] Currently, we have six different rates:

[2]We refer to this as "approximate" marginal costs, as we assume that short run costs are normally distributed around this mean, cf. below.

[3]This structure of CO_2 intensities and taxes is adopted due to the structure of data inputs.

i. The cost of EU-ETS quotas ($\tau_t^{\text{ETS-1}}$),
ii. the price of EU-ETS2 quotas ($\tau_t^{\text{ETS-2}}$),
iii. the additional tax on Danish emissions from electricity generation that is covered by EU-ETS ($\tau_t^{CO_2,E}$),
iv. a tax on Danish emissions from electricity generation that is not covered by EU-ETS ($\tau_t^{\text{Non-ETS},E}$),
v. the additional tax on Danish emissions from district heat production that is covered by EU-ETS ($\tau_t^{CO_2,H}$),
vi. and a tax on Danish emissions from district heat production not covered by the EU-ETS, but (in the future) the EU-ETS2 ($\tau_t^{\text{Non-ETS},H}$).

For condensation plants, we define:

$$\mu_{t,i}^{CO_2} \equiv \tilde{\mu}_{t,i}^{CO_2} + \sum_b \left(\alpha_{f(i),b} \times \mu_{t,b}^{CO_2}\right) \tag{5}$$

$$\tau_{t,i}^{CO_2} \equiv CO2_{t,i} \times \left(\tau_t^{\text{ETS-1}} + \tau_t^{CO_2,E}\right) + (1 - CO2_{t,i})\,\tau_t^{\text{Non-ETS},E} \equiv \tau_{t,i}^{CD}, \tag{6}$$

where $CO2_{t,i}$ indicates the share of emissions from the specific plant that is covered by the EU-ETS.[4] CO_2 intensities measure ton CO_2e per MWh of energy input and CO_2 taxes measure DKK/ton CO_2e.

The NO_x intensity is defined by a plant-specific term $\tilde{\mu}_{t,i}^{NO_x}$; there are no additional terms related to the fuelmix. The tax is identical across i as well, i.e. we simply have $\mu_{t,i}^{NO_x} = \tilde{\mu}_{t,i}^{NO_x}$ and $\tau_{t,i}^{NO_x} = \tau_t^{NO_x}$. For NO_x emissions, intensities measure kg/MWh energy input and taxes measure DKK/kg emissions.

The SO_2 intensity is defined purely from fuel use (no $\tilde{\mu}_{t,i}^{SO_2}$ parameter), but the plants have individual degrees of desulphurization (i.e. removal of SO_2), defined by the parameter $DS_{t,i}$. Taxes are the same across i plants. For SO_2, intensities measure kg/MWh energy input and taxes DKK/kg. Formally, we thus have

$$\mu_{t,i}^{SO_2} \equiv (1 - DS_{t,i}) \sum_b \left(\alpha_{f(i),b} \times \mu_{t,b}^{SO_2}\right), \qquad \tau_{t,i}^{SO_2} \equiv \tau_t^{SO_2}.$$

Foreign plants. Plants outside Denmark are only subject to the EU-ETS price, i.e. for $i \notin DK$ we replace $T_{t,i}^M$ by $\mu_{t,i}^{CO_2}\tau_t^{\text{ETS-1}}/\eta_{t,i}$, and the resource-zone term $c_{t,i}^r$ is zero.

Short run optimization. In the short run (within a given year t), the producer takes installed generating capacity $Q_{t,h,i}^E$ and equilibrium prices $p_{t,h,g(i)}^E$ as given. The producer then maximizes profits given by

$$Q_{t,i}^E \Pi_{t,i} = \sum_h n_h\, Q_{t,h,i}^E \left[p_{t,h,g(i)}^E s_{t,h,i}^E - F_{t,i}\left(s_{t,h,i}^E\right)\right] - Q_{t,i}^E \times FOM_{t,i},$$

where $\Pi_{t,i}$ is profits per installed capacity, $s_{t,h,i}^E \equiv S_{t,h,i}^E / Q_{t,h,i}^E$ measures the utilization rate, $S_{t,h,i}^E$ is electricity generation, $FOM_{t,i}$ measures fixed operating and maintenance costs (proportional to installed capacity), and F is the short-run average cost function, which is convex in $s_{t,h,i}^E$. It can be verified that if costs are normally distributed around $c_{t,i}$ with standard deviation $\sigma_{t,i}$, the

[4]In the input data, $CO2_{t,i}$ is simply a dummy.

optimal dispatch decision yields

$$s^E_{t,h,i} = \Phi\left(\frac{p^E_{t,h,g(i)} - c_{t,i}}{\sigma_{t,i}}\right) \tag{7}$$

$$\pi_{t,h,i} = s^E_{t,h,i}\left(p^E_{t,h,g(i)} - c_{t,i}\right) + \sigma_{t,i}\phi\left(\frac{p^E_{t,h,g(i)} - c_{t,i}}{\sigma_{t,i}}\right) \tag{8}$$

$$Q^E_{t,i}\Pi_{t,i} = \sum_h n_h\, Q^E_{t,h,i}\pi_{t,h,i} - Q^E_{t,i} \times FOM_{t,i}, \tag{9}$$

where ϕ, Φ denote the density and cumulative density of the standard normal distribution, respectively, and $\pi_{t,h,i}$ measures hourly profits per installed capacity $Q^E_{t,h,i}$.

The generating capacity in hour h is tied to a yearly capacity $Q^E_{t,i}$ following a simple decomposition

$$Q^E_{t,h,i} = Q^E_{t,i} \times CapF_{t,i} \times \gamma_{h,v(i)}, \qquad \gamma_{h,v(i)} \equiv \sum_v \mu^{cap}_{t,i,v}\,\gamma_{h,v},$$

where $CapF_{t,i}$ measures a capacity factor $\in [0,1]$, and $\gamma_{h,v(i)}$ measures the hourly-to-average capacity for each hour for the variation profile of plant i.[5] The share parameters $\mu^{cap}_{t,i,v}$ satisfy $\sum_v \mu^{cap}_{t,i,v} = 1$ for all (t,i) and allow a plant to be described as a linear combination of base variation profiles, e.g. when data are only available for a few years and intermediate years are interpolated.

Using this, we write unit profits as

$$\Pi_{t,i} = \sum_h n_h\left(CapF_{t,i} \times \gamma_{h,v(i)} \times \pi_{t,h,i}\right) - FOM_{t,i}. \tag{10}$$

For type-1 (fixed-contract) plants, (7)–(10) apply with $p^E_{t,h,g(i)}$ set to zero, so that dispatch is governed by $\Phi(-c_{t,i}/\sigma_{t,i})$ and the profit margin by $-c_{t,i}$.

The default units applied in this module: prices measure DKK/MWh, generation and capacities in GWh and GW, fixed costs in million DKK per GW, and short run profits $\Pi_{t,i}$ in million DKK/GW.[6]

Interactions with other parts of the model. The supply of electricity $S^E_{t,h,i} = s^E_{t,h,i}Q^E_{t,h,i}$ enters the equilibrium condition in the relevant geographic market (g_E). Fuel consumption and emissions are computed simply from multiplying the total production throughout a year onto the emission intensities $\mu^m_{t,i}$. Consumption of base fuels is given by total production times the fuelmix coefficient $\alpha_{f,b}$ divided by the conversion efficiency of the plant $\eta_{t,i}$; see section 7.

The measure $\Pi_{t,i}$ is a key decision variable for investments in further capacity; this is work-in-progress, however.

2.4 Boiler heaters (BH)

Notationally, boiler heaters can be presented as condensation plants, but operating in the district heat markets instead. There are two notable differences, though: first, as mentioned in the section on CD plants, CO_2 emissions are taxed differently when the purpose is district heating. Second, fuels used for district heating purposes are taxed further at a rate specific to the base fuel $\tau^{DH}_{t,b}$.

[5] For condensation plants, the convention is to have $CapF_{t,i} = \gamma_{h,v(i)} = 1$.

[6] This means that the equation for $\Pi_{t,i}$ technically involves an appropriate adjustment of short run revenues; this is left out for readability.

Specifically, we have

$$\tau_{t,i}^{CO_2} = CO2_{t,i} \times \left(\tau_t^{\text{ETS-1}} + \tau_t^{CO_2,DH}\right) + (1 - CO2_{t,i}) \times \left(\tau_t^{\text{ETS-2}} + \tau_t^{\text{Non-ETS},DH}\right) \equiv \tau_{t,i}^{BH} \quad (11)$$

$$T_{t,i}^{DH} = \frac{\sum_b \alpha_{f(i),b} \times \tau_{t,b}^{DH}}{\eta_{t,i}}. \quad (12)$$

Note the asymmetry with (6): the EU-ETS2 price enters the non-ETS term for heat, but not for electricity.[7] Marginal costs then read $c_{t,i} = v_{t,i} + \left(c_{t,f(i)}^{fuel} + c_{t,i}^{r}\right)/\eta_{t,i} + T_{t,i}^{M} + T_{t,i}^{DH}$, with $T_{t,i}^{M}$ evaluated using $\tau_{t,i}^{BH}$, and dispatch follows (7) with $p_{t,h,g(i)}^{DH}$ replacing $p_{t,h,g(i)}^{E}$.

2.5 Wind (WS, WL), solar (PV), and run-of-river hydro (RO)

The four plant types are essentially modelled as condensation plants, but with a few simplifications: no fuel consumption nor emissions/emission taxes, so that $c_{t,i} = v_{t,i} + T_{t,i}^{E}$. The investment decisions should be handled differently than condensation plants, however. A large share of these plants are on fixed settlement contracts and are therefore of bidding type 1.

2.6 Solar heating (SH)

Similar to wind and solar PV, but active on the district heat "market" instead, with $c_{t,i} = v_{t,i}$.

2.7 Exogenous production (EPE, EPH)

Plants whose production is largely determined outside the energy sector are modelled as condensation plants if they produce electricity (`EPE`) and as boiler heaters if they produce heat (`EPH`). In the previous version these were treated as a single `EP` type; the split is now explicit in the code, which matters for the fuel and emission accounting in section 7, where `EPE` is grouped with the electricity producers and `EPH` with the heat producers. In the data, they cover a range of different things. Danish plants cover surplus production of electricity or heat from other industries; some production occurs more or less exogenously and is not tied to any specific fuel use, other production is more flexible and can be adjusted up/down depending on market prices in the given hour.

Notes – follow up on this: It is not clear where the Danish EP plants' fuel consumption should be accounted for; are they included in the energy sector of the industry that the producers primarily engage in?

2.8 CHP plants: allocating fuel and emissions between electricity and heat

Combined heat and power plants produce two outputs from one fuel input, and Danish tax law taxes fuel used for heat production differently from fuel used for electricity production. Some rule is therefore needed to split the fuel input of a CHP plant between the two purposes. This is a substantive change relative to the previous version of the documentation, which derived the split from the *output shares* of the plant. The current implementation instead follows the statutory rules, which the plant may choose between.

Consider a CHP plant operating in a given mode, producing E units of electricity and H units of district heat from $F = 1/\eta_{t,i}$ units of fuel (all measured per unit of the composite output of that mode; see the plant-type sections below for the exact endpoints). Danish law offers two rules for splitting F into a part attributed to electricity, F^E, and a residual attributed to heat, $F^H = F - F^E$:

[7] In the current data preparation, the EU-ETS2 price is *also* added to both $\tau_t^{\text{Non-ETS},E}$ and $\tau_t^{\text{Non-ETS},DH}$ in the KF25 patch. Combined with (11), non-ETS heat production is therefore charged the EU-ETS2 price twice, while non-ETS electricity production is charged it once. This should be resolved by either removing the explicit $\tau_t^{\text{ETS-2}}$ term from (11) or by not adding it in the data preparation.

- **The V-rule** (heat-efficiency rule): the fuel attributed to heat is $F^H = H/\overline{\nu}$ with $\overline{\nu} = 1.25$, and the remainder is attributed to electricity, $F^E = F - H/\overline{\nu}$.
- **The E-rule** (electricity-efficiency rule): the fuel attributed to electricity is $F^E = E/\overline{\varepsilon}$ with $\overline{\varepsilon} = 0.65$, and the remainder is attributed to heat.

In addition, the fuel attributed to electricity under the V-rule is capped, so that the implied electricity efficiency cannot fall below $\overline{\gamma} = 0.35$:

$$F^E \leq E/\overline{\gamma}.$$

We refer to a plant whose V-rule allocation is bound by this cap as operating under a *guarded V-rule.*

Assigning a rule to each plant. A plant is assumed to choose the rule that attributes as much fuel as possible to electricity, since fuel attributed to heat carries both the heat CO_2 tax $\tau_{t,i}^{BH}$ and the fuel-specific heat tax $\tau_{t,b}^{DH}$, whereas fuel attributed to electricity only carries $\tau_{t,i}^{CD}$. For each of the (at most two) modes that a plant can operate in, we compute

$$D = \min\left(F^E_{V\text{-rule}},\, E/\overline{\gamma}\right) - F^E_{E\text{-rule}},$$

so that $D > 0$ means that the V-family attributes more fuel to electricity than the E-rule. The plant is assigned to the V-rule if $D > 0$ in both modes, to the E-rule if $D < 0$ in both modes, and, when the sign of D differs across the two modes, to whichever rule dominates by the larger margin. If the V-rule is selected and the cap binds in either mode, the plant is assigned to the guarded V-rule.

This classification is carried out once, on the raw plant data, and stored as a mapping $\mathcal{R}(i) \in \{V\text{-rule},\, E\text{-rule},\, \text{guarded } V\text{-rule}\}$ (`id2TaxRule`). The mapping is treated as a categorical attribute of the plant and is therefore respected by the aggregation of producer ids (section 8.4): plants with different tax rules are never aggregated. The assignment is thus fixed in a given database and does not respond to shocks; making it endogenous would introduce a discrete choice with a jump in marginal costs, which is why it is resolved in the data step.

Fuel allocation factors. Given the assigned rule, we define per-output allocation factors $a_{t,i}^E$ (fuel per unit of electricity output) and $a_{t,i}^H$ (fuel per unit of heat output) for each mode. For back-pressure mode ($E = 1$, $H = 1/\kappa_{t,i}$, $F = 1/\eta_{t,i}$):

$$\text{V-rule:} \quad a_{t,i}^{E,BP} = \frac{1}{\eta_{t,i}} - \frac{1}{\overline{\nu}\kappa_{t,i}}, \qquad a_{t,i}^{H,BP} = \frac{1}{\overline{\nu}} \tag{13}$$

$$\text{E-rule:} \quad a_{t,i}^{E,BP} = \frac{1}{\overline{\varepsilon}}, \qquad a_{t,i}^{H,BP} = \kappa_{t,i}\left(\frac{1}{\eta_{t,i}} - \frac{1}{\overline{\varepsilon}}\right) \tag{14}$$

$$\text{guarded:} \quad a_{t,i}^{E,BP} = \frac{1}{\overline{\gamma}}, \qquad a_{t,i}^{H,BP} = \kappa_{t,i}\left(\frac{1}{\eta_{t,i}} - \frac{1}{\overline{\gamma}}\right). \tag{15}$$

For bypass mode ($E = 0$, $H = 1/\kappa_{t,i} + 1/\xi_{t,i}$) the E-rule and the guard both attribute all fuel to heat, so that $a_{t,i}^{H,BB} = (1/\eta_{t,i})/(1/\kappa_{t,i} + 1/\xi_{t,i})$, while the V-rule gives $a_{t,i}^{H,BB} = 1/\overline{\nu}$. For extraction (condensation) mode ($E = 1 + \xi_{t,i}/\kappa_{t,i}$, $H = 0$), the V-rule attributes all fuel to electricity, $a_{t,i}^{E,EX} = (1/\eta_{t,i})/(1 + \xi_{t,i}/\kappa_{t,i})$ and $a_{t,i}^{H,EX} = 0$, whereas the E-rule and the guard attribute $a_{t,i}^{E,EX} = 1/\overline{\varepsilon}$ (resp. $1/\overline{\gamma}$) with the residual $a_{t,i}^{H,EX} = (1/\eta_{t,i})/(1 + \xi_{t,i}/\kappa_{t,i}) - a_{t,i}^{E,EX}$ taxed at heat rates but levied per unit of *electricity* output, since the mode produces no heat.

Output-oriented taxes. With these factors, the tax terms used in the CHP marginal cost functions are

$$
\begin{aligned}
T_{t,i}^{E,BP} &= a_{t,i}^{E,BP}\,\mu_{t,i}^{CO_2}\tau_{t,i}^{CD} + \tilde{T}_{t,i}^{E} && (16)\\
T_{t,i}^{DH,BP} &= a_{t,i}^{H,BP}\left[\mu_{t,i}^{CO_2}\tau_{t,i}^{BH} + \sum_b \alpha_{f(i),b}\tau_{t,b}^{DH}\right] && (17)\\
T_{t,i}^{DH,BB} &= a_{t,i}^{H,BB}\left[\mu_{t,i}^{CO_2}\tau_{t,i}^{BH} + \sum_b \alpha_{f(i),b}\tau_{t,b}^{DH}\right] && (18)\\
T_{t,i}^{E,EX} &= a_{t,i}^{E,EX}\,\mu_{t,i}^{CO_2}\tau_{t,i}^{CD} + a_{t,i}^{H,EX}\left[\mu_{t,i}^{CO_2}\tau_{t,i}^{BH} + \sum_b \alpha_{f(i),b}\tau_{t,b}^{DH}\right] + \tilde{T}_{t,i}^{E}, && (19)
\end{aligned}
$$

where $\tilde{T}_{t,i}^{E}$ collects other taxes and subsidies levied per MWh of electricity output (in the current data, essentially the production subsidy with a negative sign). Note that no electricity-related subsidy enters $T_{t,i}^{DH,BB}$, since bypass mode produces no electricity.

Finally, since CO_2 is now handled entirely through the output-oriented tax terms, the emission tax function common to both modes only covers SO_2 and NO_x:

$$
T_{t,i}^{M,CHP} = \frac{1}{\eta_{t,i}} \sum_{m \neq CO_2} \mu_{t,i}^{m}\tau_{t,i}^{m}. \tag{20}
$$

2.9 Back-pressure plants (BP)

One can think of back-pressure plants as condensation plants that are able to utilize surplus heat from production as district heating. For the BP plants, the parameter $\kappa_{t,i}$ defines the fixed ratio of electricity-to-heat generation of the plant.

Let $c_{t,i}$ be the approximate marginal cost from increasing electricity generation marginally with $1/\kappa_{t,i}$ units of co-produced district heating; we then have

$$
c_{t,i} = v_{t,i} + \frac{c_{t,f(i)}^{fuel} + c_{t,i}^{r}}{\eta_{t,i}} + T_{t,i}^{M,CHP} + T_{t,i}^{E,BP} + \frac{T_{t,i}^{DH,BP}}{\kappa_{t,i}}. \tag{21}
$$

These variables should be interpreted carefully due to the co-production of electricity and district heat:

- $c_{t,i}$ and $v_{t,i}$ are defined as DKK per unit of the composite good that consists of 1 unit of electricity and $1/\kappa_{t,i}$ units of district heat;
- producing one unit of this composite good requires $1/\eta_{t,i}$ units of fuelmix $f(i)$. Thus, all variables that are measured per fuel input are scaled by $1/\eta_{t,i}$ – this includes the emission taxes $T_{t,i}^{M,CHP}$;
- $T_{t,i}^{E,BP}$ is measured in DKK per electricity output and $T_{t,i}^{DH,BP}$ in DKK per district heat output. In the definition of $c_{t,i}$, these are thus weighted by 1 and $1/\kappa_{t,i}$, respectively.

Short run optimization. In the short run, the producer takes installed generating capacity $Q_{t,h,i}^{E}$ and equilibrium prices $p_{t,h,g(i)}^{E}, p_{t,h,g(i)}^{DH}$ as given.[8] The producer then maximizes

$$
\begin{aligned}
Q_{t,i}^{E}\Pi_{t,i} &= \sum_h n_h\, Q_{t,h,i}^{E}\left[s_{t,h,i}^{E} p_{t,h,i} - F_{t,i}^{c}\left(s_{t,h,i}^{E}\right)\right] - Q_{t,i}^{E} \times FOM_{t,i},\\
p_{t,h,i} &\equiv p_{t,h,g(i)}^{E} + \frac{1}{\kappa_{t,i}} p_{t,h,g(i)}^{DH}.
\end{aligned}
$$

[8] We use electricity generating capacity as the main variable here, but the constant electricity-to-heat ratio means that we could alternatively write this in terms of district heat capacity.

If costs are normally distributed around $c_{t,i}$ with standard deviation $\sigma_{t,i}$, optimization yields a solution identical to the CD plant, only with prices and costs redefined:

$$\Delta^q_{t,h,i} \equiv p_{t,h,i} - c_{t,i} \tag{22}$$

$$s^E_{t,h,i} = \Phi\left(\frac{\Delta^q_{t,h,i}}{\sigma_{t,i}}\right) \tag{23}$$

$$\pi_{t,h,i} = s^E_{t,h,i}\Delta^q_{t,h,i} + \sigma_{t,i}\phi\left(\frac{\Delta^q_{t,h,i}}{\sigma_{t,i}}\right) \tag{24}$$

$$\Pi_{t,i} = \sum_h n_h\left(CapF_{t,i} \times \gamma_{h,v(i)} \times \pi_{t,h,i}\right) - FOM_{t,i}. \tag{25}$$

For type-1 plants, $p_{t,h,i} = p^{DH}_{t,h,g(i)}/\kappa_{t,i}$, i.e. the spot electricity price drops out of the price index.

Interactions with other parts of the model. The supply of electricity and district heating enters the relevant equilibrium condition with $S^E_{t,h,i} = n_h\, s^E_{t,h,i}Q^E_{t,i}$ and $S^{DH}_{t,h,i} = S^E_{t,h,i}/\kappa_{t,i}$. Fuel consumption is given by $S^E_{t,h,i}\alpha_{f(i),b}/\eta_{t,i}$ (as $\eta_{t,i}$ measures the efficiency of electricity generation).

2.10 Back-pressure with bypass (BB)

Back-pressure plants with bypass can either operate in back-pressure mode or transform electricity output to heat output at a constant rate $\xi_{t,i}$; specifically, $\xi_{t,i} > 0$ measures the loss in electricity from increasing heat output marginally.[9] We use a smooth approximation for a discrete choice between operating in back-pressure or bypass mode.[10]

Let $c^{BP}_{t,i}$ denote the marginal cost of operating as a back-pressure plant, cf. (21). In bypass mode, each unit of electricity is converted to $1/\xi_{t,i}$ units of district heating, i.e. the output shares are 0 and $1/\xi_{t,i} + 1/\kappa_{t,i}$ for electricity and district heat, respectively. The cost index in bypass mode is

$$c^{BB}_{t,i} = v_{t,i} + \frac{c^{fuel}_{t,f(i)} + c^r_{t,i}}{\eta_{t,i}} + T^{M,CHP}_{t,i} + \left(\frac{1}{\kappa_{t,i}} + \frac{1}{\xi_{t,i}}\right)T^{DH,BB}_{t,i}. \tag{26}$$

The difference in profit margins between back-pressure and bypass mode is

$$\Delta_{t,h,i} = p^E_{t,h,g(i)} - \frac{p^{DH}_{t,h,g(i)}}{\xi_{t,i}} + c^{BB}_{t,i} - c^{BP}_{t,i},$$

which will tend to be positive if electricity prices are high relative to district heat prices, or if heat is taxed at higher rates than electricity. We define the share of the plant operating in back-pressure mode as

$$u^{BP}_{t,h,i} = \Phi\left(\frac{\Delta_{t,h,i}}{\sigma^{EH}_{t,i}}\right), \tag{27}$$

where $\sigma^{EH}_{t,i} > 0$ is a smoothing parameter. For type-1 plants, the p^E term drops out of $\Delta_{t,h,i}$.

[9] In the current version, this parameter is simply 1 for all BB plants.

[10] As an alternative to this approximation approach, we can assume that electricity can be transformed to heating in a nonlinear fashion, e.g. following a normalized constant elasticity of transformation (see Berg and Eskildsen, 2021a). A third option is to base the smoothing on extreme value distributions.

Short run optimization. The producer problem is as for the BP plant, with revenues and costs being a linear combination of the two modes:

$$\Delta^q_{t,h,i} = u^{BP}_{t,h,i}\left[p^E_{t,h,g(i)} + \frac{p^{DH}_{t,h,g(i)}}{\kappa_{t,i}} - c^{BP}_{t,i}\right] + (1-u^{BP}_{t,h,i})\left[\left(\frac{1}{\kappa_{t,i}} + \frac{1}{\xi_{t,i}}\right)p^{DH}_{t,h,g(i)} - c^{BB}_{t,i}\right] \tag{28}$$

$$s^E_{t,h,i} = \Phi\left(\frac{\Delta^q_{t,h,i}}{\sigma_{t,i}}\right) \times CapF_{t,i} \times \gamma_{h,v(i)}. \tag{29}$$

Note that in the current implementation, the capacity factor and the hourly variation multiply the dispatch rate directly (rather than entering as a separate capacity decomposition); we write it this way in the remainder of the document for the CHP types.

Interactions with other parts of the model. The supply of electricity and district heat is

$$S^E_{t,h,i} = n_h\, Q^E_{t,i}\, s^E_{t,h,i}\, u^{BP}_{t,h,i} \tag{30}$$

$$S^{DH}_{t,h,i} = n_h\, Q^E_{t,i}\, s^E_{t,h,i}\left[\frac{1}{\kappa_{t,i}} + \frac{1-u^{BP}_{t,h,i}}{\xi_{t,i}}\right]. \tag{31}$$

Fuel consumption is proportional to $n_h Q^E_{t,i} s^E_{t,h,i}/\eta_{t,i}$, i.e. to the composite output rather than to electricity output alone, which is why the fuel accounting in section 7 treats BB and EX plants separately from BP plants.

2.11 Extraction plants (EX)

Extraction plants can either operate in back-pressure or condensation mode. $\xi_{t,i}$ once again denotes the loss in electricity from increasing heat output marginally. To avoid notational confusion with the condensation plant type, we refer to the second mode as extraction (EX). In extraction mode, the resulting output shares are $1 + \xi_{t,i}/\kappa_{t,i}$ and 0 for electricity and district heat, respectively.[11] The cost index in extraction mode is

$$c^{EX}_{t,i} = v_{t,i} + \frac{c^{fuel}_{t,f(i)} + c^r_{t,i}}{\eta_{t,i}} + T^{M,CHP}_{t,i} + \left(1 + \frac{\xi_{t,i}}{\kappa_{t,i}}\right)T^{E,EX}_{t,i}. \tag{32}$$

The difference in profit margins between back-pressure and extraction mode is

$$\Delta_{t,h,i} = \frac{p^{DH}_{t,h,g(i)}}{\kappa_{t,i}} - \frac{\xi_{t,i}p^E_{t,h,g(i)}}{\kappa_{t,i}} + c^{EX}_{t,i} - c^{BP}_{t,i},$$

and $u^{BP}_{t,h,i} = \Phi(\Delta_{t,h,i}/\sigma^{EH}_{t,i})$ as before. The composite profit margin is

$$\Delta^q_{t,h,i} = u^{BP}_{t,h,i}\left[p^E_{t,h,g(i)} + \frac{p^{DH}_{t,h,g(i)}}{\kappa_{t,i}} - c^{BP}_{t,i}\right] + (1-u^{BP}_{t,h,i})\left[\left(1 + \frac{\xi_{t,i}}{\kappa_{t,i}}\right)p^E_{t,h,g(i)} - c^{EX}_{t,i}\right], \tag{33}$$

with dispatch $s^E_{t,h,i} = \Phi(\Delta^q_{t,h,i}/\sigma_{t,i})CapF_{t,i}\gamma_{h,v(i)}$ and supplies

$$S^E_{t,h,i} = n_h\, Q^E_{t,i}\, s^E_{t,h,i}\left(1 + \left[1-u^{BP}_{t,h,i}\right]\frac{\xi_{t,i}}{\kappa_{t,i}}\right) \tag{34}$$

$$S^{DH}_{t,h,i} = n_h\, Q^E_{t,i}\, s^E_{t,h,i}\,\frac{u^{BP}_{t,h,i}}{\kappa_{t,i}}. \tag{35}$$

[11] We get this from $\xi_{t,i} \equiv -\Delta E/\Delta H$; setting $\Delta H = -1/\kappa_{t,i}$ (the back-pressure production), the additional potential production is $\Delta E = \xi_{t,i}/\kappa_{t,i}$.

2.12 Storage technologies

The model contains four storage technologies: hydro with reservoir (HY), hydro with pumped storage (HYS), electricity storage (ES), and district heat storage (HS). They share a common structure, and the code shares a common set of macros. Relative to the previous version, there are two changes worth noting:

1. all four types now exist in both a `chronological` and a `clusterAvg` state, so that storage can be solved in the clustered model as well;
2. charging and discharging are netted, so that a plant cannot charge and discharge simultaneously.

2.12.1 Hydro with storage (HY)

Hydro power plants with storage rely on an exogenous inflow of potential energy as the standard variable renewable energy (VRE) plant types. They can delay dispatch by storing the energy in a reservoir with a given storage capacity. The marginal cost is identical to the case of simple VRE plants ($c_{t,i} = v_{t,i} + T^E_{t,i}$), but the decision to dispatch or postpone is dynamic.

Let $Q^E_{t,i}$ denote the installed dispatch capacity and $\zeta_{t,i}$ the duration, defined as the storage capacity divided by the dispatch capacity, i.e. the number of hours the generator can produce at maximum capacity before emptying a full reservoir. The hourly generation capacity is constant throughout a year, but the inflow varies. We let $CapF^I_{t,i}$ denote the average utilization rate throughout a year and $\gamma^I_{h,v(i)}$ the hourly-to-average inflow.[12]

Writing the problem in per-capacity terms ($r_{t,h,i} \equiv R_{t,h,i}/Q^E_{t,i}$), the plant solves

$$\begin{aligned} \max\ \Pi_{t,i} &= \sum_h n_h \left[p^E_{t,h,g(i)} s^E_{t,h,i} - F^c_{t,h,i}\left(s^E_{t,h,i}, r_{t,h,i}\right)\right] - FOM_{t,i} \\ \text{s.t.}\ r_{t,h,i} &= r_{t,h-1,i} + n_h\, \gamma^I_{h,v(i)} \times CapF^I_{t,i} - n_h\, s^E_{t,h,i}. \end{aligned}$$

Assuming that costs F^c are convex and follow a specific functional form (see Berg and Eskildsen, 2021b), the conditions identifying the optimal dispatch path are

$$s^E_{t,h,i} = \Phi\left(\frac{p^E_{t,h,g(i)} - c_{t,i} - \lambda_{t,h,i}}{\sigma_{t,i}}\right) \tag{36}$$

$$\lambda_{t,h,i} = \lambda_{t,h+1,i} + \Lambda(r_{t,h,i}) \tag{37}$$

$$r_{t,h,i} = r_{t,h-1,i} + n_h\, \gamma^I_{h,v(i)} \times CapF^I_{t,i} - n_h\, s^E_{t,h,i}, \tag{38}$$

where the marginal storage cost term is

$$\Lambda(r) \equiv x^{out}_{t,i}\left(\frac{\zeta_{t,i}}{2} - r\right) + \left(\frac{x^{out}_{t,i}}{2} - \frac{x^{in}_{t,i}}{\zeta_{t,i}}\right)\left(\sqrt{r^2 + \left(x^{\Delta}_{t,i}\right)^2} - \sqrt{(r-\zeta_{t,i})^2 + \left(x^{\Delta}_{t,i}\right)^2}\right). \tag{39}$$

The first condition states that dispatch depends on whether the hourly price exceeds the marginal cost of production plus the continuation value ($\lambda_{t,h,i}$) of storing the energy. The second states that the value of storing energy is the continuation value in $h{+}1$ plus a marginal storage cost term that depends on how full the reservoir is. This cost term is a smooth version of a kinked linear cost function with two marginal cost regimes: one outside the natural bounds of the reservoir $(0, \zeta_{t,i})$ where the marginal cost of changing $r_{t,h,i}$ is very large, and one inside the natural bounds where it is very low. The three parameters are $x^{out}_{t,i}$ (approximate costs outside bounds), $x^{in}_{t,i}$ (approximate costs inside bounds), and $x^{\Delta}_{t,i}$ (smoothing between the two regimes).[13]

[12] Specifically, we let $CapF^I_{t,i} \equiv I_{t,i}/(8760 \times Q^E_{t,i})$ where $I_{t,i}$ is the yearly inflow.

[13] In the baseline, $x^{in}_{t,i} = 1$ and $x^{\Delta}_{t,i} = 10^{-3}$ uniformly, while $x^{out}_{t,i} = \max(100,\, 3000/\zeta_{t,i})$ scales with how large the reservoir is relative to the plant's power rating; Section 2.12.3 explains why. All three are in the set of parameters that the solution procedure is allowed to relax when a solve fails, see Section 9.2.

Profits per installed capacity are[14]

$$\Pi_{t,i} = \sum_h n_h \left(s^E_{t,h,i} \left(p^E_{t,h,g(i)} - c_{t,i} \right) + \sigma_{t,i} \phi \left(\frac{p^E_{t,h,g(i)} - c_{t,i} - \lambda_{t,h,i}}{\sigma_{t,i}} \right) \right) - FOM_{t,i}. \tag{40}$$

2.12.2 Hydro with pumped storage (HYS), electricity storage (ES), district heat storage (HS)

Hydro power plants with pumped storage differ from HY plants in that they can fill up the storage capacity by pumping up water – a process that uses electricity instead; we let $\eta_{t,i} \in (0,1)$ be the round-trip efficiency of the pumped storage. The charging speed $\chi_{t,i}$ measures gross charge capacity relative to discharge capacity.[15] Electricity storage (ES) is notationally identical but without exogenous inflow; district heat storage (HS) is identical to ES with capacities, consumption and generation related to heating and with $c_{t,i} = v_{t,i}$ (no electricity subsidy term).

We extend the choice set of the plant with a charging rate. The plant forms a discharge intent $s^E_{t,h,i}$ and a charge intent $d^E_{t,h,c(i)}$:

$$s^E_{t,h,i} = \Phi \left(\frac{p^E_{t,h,g(i)} - c_{t,i} - \lambda_{t,h,i}}{\sigma_{t,i}} \right) \tag{41}$$

$$d^E_{t,h,c(i)} = \Phi \left(\frac{\lambda_{t,h,i}\eta_{t,i} - p^E_{t,h,g(i)} - c_{t,c(i)}}{\sigma_{t,i}} \right), \tag{42}$$

where $c_{t,c(i)}$ is the marginal cost of charging in addition to the equilibrium price. Currently, we set $c_{t,c(i)} \approx 0$, but this can be used to include different types of tariffs, e.g. from accessing the distribution grid.

Netting of charge and discharge. A plant that is close to indifferent between charging and discharging would, with the smooth decision rules above, do both at the same time. Since this generates spurious round-trip losses and spurious volumes in the market clearing conditions, the two intents are netted before they enter the rest of the model:

$$\widetilde{s}_{t,h,i} \equiv \max\left(s^E_{t,h,i} - \chi_{t,i} d^E_{t,h,c(i)},\, 0 \right) = \tfrac{1}{2} \left(\left| s^E_{t,h,i} - \chi_{t,i} d^E_{t,h,c(i)} \right| + s^E_{t,h,i} - \chi_{t,i} d^E_{t,h,c(i)} \right) \tag{43}$$

$$\widetilde{d}_{t,h,i} \equiv \max\left(\chi_{t,i} d^E_{t,h,c(i)} - s^E_{t,h,i},\, 0 \right). \tag{44}$$

Only the net position affects the storage level, the market, and profits:

$$r_{t,h,i} = r_{t,h-1,i} + n_h \gamma^I_{h,v(i)} CapF^I_{t,i} - n_h \widetilde{s}_{t,h,i} + \eta_{t,i} n_h \widetilde{d}_{t,h,i} \tag{45}$$

$$S^E_{t,h,i} = n_h\, Q^E_{t,i}\, \widetilde{s}_{t,h,i}, \qquad D^E_{t,h,c(i)} = n_h\, Q^E_{t,i}\, \widetilde{d}_{t,h,i}. \tag{46}$$

Profits per installed capacity are

$$\begin{aligned}\Pi_{t,i} = &\sum_h n_h \left[\widetilde{s}_{t,h,i} \left(p^E_{t,h,g(i)} - c_{t,i} \right) + \sigma_{t,i} \phi \left(\frac{p^E_{t,h,g(i)} - c_{t,i} - \lambda_{t,h,i}}{\sigma_{t,i}} \right) \right] \\ &- \sum_h n_h \left[\widetilde{d}_{t,h,i} \left(p^E_{t,h,g(i)} + c_{t,c(i)} \right) - \sigma_{t,i} \phi \left(\frac{\eta_{t,i}\lambda_{t,h,i} - p^E_{t,h,g(i)} - c_{t,c(i)}}{\sigma_{t,i}} \right) \right] - FOM_{t,i}.\end{aligned} \tag{47}$$

The inflow term is absent for ES and HS, and for HS all electricity prices and quantities are replaced by their district heating counterparts.

[14]These profits are only approximate, because we do not account for the nonlinear costs from fluctuations in $r_{t,h,i}$. With the current set of parameters used in the model, these are very close to zero anyway.

[15]This means that a plant that can consume twice the amount of electricity per hour than it can potentially generate has $\chi_{t,i} = 2$.

2.12.3 Non-chronological representation

When hours are clustered non-chronologically, "$h+1$" and "$h-1$" are no longer well defined. We replace them by expectations under the Markov transition matrices estimated in the aggregation step (section 8.5). Let $\Pr^+(h \to h')$ denote the probability that state h is followed by state h', and $\Pr^-(h' \to h)$ the probability that state h was preceded by h'. The continuation value and the storage balance then read

$$\lambda_{t,h,i} = \sum_{h'} \Pr^+(h \to h')\, \lambda_{t,h',i} + \Lambda(r_{t,h,i}) \tag{48}$$

$$r_{t,h,i} = \sum_{h'} \Pr^-(h' \to h)\, r_{t,h',i} + n_h \gamma^I_{h,v(i)} CapF^I_{t,i} - n_h \widetilde{s}_{t,h,i} + \eta_{t,i} n_h \widetilde{d}_{t,h,i}, \tag{49}$$

with all remaining equations unchanged. This is an approximation: it tracks the *state average* of the reservoir level rather than its full distribution. A full dynamic programme over the joint distribution of state variables is not tractable at the scale of the model, cf. the discussion in section 8.5.

Keeping the stock inside its bounds. The approximation costs something that is worth stating plainly. In the chronological model the state equation is a chain, and $r_{t,h,i}$ is pinned down hour by hour from a known starting point. Replacing $h-1$ by an expectation over the transition matrix turns the chain into a system whose solution is determined only up to the level of the reservoir: adding a constant to $r_{t,h,i}$ in every state satisfies the same equations. What ties the level down is (39), and specifically the penalty $x^{out}_{t,i}$ that applies outside $(0, \zeta_{t,i})$. If that penalty is weak relative to the price differences the plant is arbitraging, the solved stock can drift outside its physical range without the solver objecting.

This is why $x^{out}_{t,i}$ scales as $\max(100,\ 3000/\zeta_{t,i})$ rather than being set uniformly. The stock is measured in hours of full output, so a given absolute drift is a much larger share of a three-hour battery than of a hydro reservoir with 600 hours of storage; scaling the penalty by $1/\zeta_{t,i}$ makes the cost of a given *fractional* excursion roughly comparable across plants, and the floor keeps a meaningful penalty on the long-duration reservoirs.

What remains is not zero, and Section 10 reports it: in the solved baseline the stock lies outside $(0, \zeta_{t,i})$ by 3–4 per cent of a full store on average across storage plants and states, weighted by the hours a state represents. The residual is concentrated in short-duration electricity storage, where the mean excursion is 5–9 per cent and individual states reach a third of capacity, and it is small for hydro and heat storage, which stay within about one per cent. The pattern is what the approximation would predict: a battery with three hours of storage exists to cycle within the day, and the state average of its stock is exactly the statistic that a non-chronological representation cannot pin down. Results that turn on the behaviour of short-duration storage should be read with that in mind, and the design-days project of Section 11.6 is aimed at it.

2.13 Electrical heaters (EH)

Electrical heaters and heat pumps are notationally identical, but typical parameter values and regulation differ. They rely on electricity to generate district heating with time-varying coefficients of performance (COPs).[16] We define the hourly COP as $\eta_{t,h,i} \equiv \eta_{t,i} \times CapF_{t,i} \times \gamma_{h,v(i)}$.

Marginal costs per input of electricity are

$$c_{t,h,i} = v_{t,i} + p^E_{t,h,g(i)} + T^{E,EH}_{t,i}, \tag{50}$$

[16] The COP of heat pumps varies across the year due to variation in temperatures; the COP of a heat pump is generally decreasing in the "lift", i.e. the difference between the sink (e.g. house) and source (ambient air).

where $v_{t,i}$ measures costs associated with heat production per electricity input and $T^{E,EH}_{t,i}$ measures taxes/subsidies per electricity consumed by electrical heaters and heat pumps. The producer takes installed capacity for electricity consumption $Q^E_{t,c(i)}$ and hourly prices as given, and profit maximization under normally distributed costs yields

$$\Delta_{t,h,i} \equiv \eta_{t,h,i} \times p^{DH}_{t,h,g(i)} - c_{t,h,i} \tag{51}$$

$$d^E_{t,h,c(i)} = \Phi\left(\frac{\Delta_{t,h,i}}{\sigma_{t,i}}\right) \tag{52}$$

$$\pi_{t,h,i} = d^E_{t,h,c(i)}\Delta_{t,h,i} + \sigma_{t,i}\phi\left(\frac{\Delta_{t,h,i}}{\sigma_{t,i}}\right) \tag{53}$$

$$\Pi_{t,i} = \sum_h n_h\, \pi_{t,h,i} - FOM_{t,i}. \tag{54}$$

Note that the capacity factor and the hourly variation enter through the COP rather than through a separate capacity decomposition, which is why the cost-wedge term $\sigma_{t,i}\phi(\cdot)$ is not scaled by $CapF_{t,i}\gamma_{h,v(i)}$ for this plant type.

Interactions with other parts of the model. Electricity consumption is $D^E_{t,h,c(i)} = n_h\, d^E_{t,h,c(i)}Q^E_{t,c(i)}$ and district heat generation is $S^{DH}_{t,h,i} = \eta_{t,h,i} \times D^E_{t,h,c(i)}$.

2.14 Power-to-Hydrogen producers (EF)

Currently, we simply model hydrogen using a fixed price p^H_t that is constant throughout the year. This corresponds to a small open economy assumption; the price is now an explicit time series (`pHy[t]`) rather than a scalar, so that a path for the world hydrogen price can be imposed. We assume that PtX producers can operate their facility flexibly and only produce in hours where the electricity price is sufficiently low.

Notationally, this is an electrical heater whose output is priced at the hydrogen price. Costs per electricity input are $c_{t,h,i} = v_{t,i} + p^E_{t,h,g(i)} + T^{E,EF}_{t,i}$ and the first-order conditions are those of the EH plant with

$$\Delta_{t,h,i} \equiv \eta_{t,h,i} \times p^H_t - c_{t,h,i}. \tag{55}$$

Electricity consumption is $D^E_{t,h,c(i)} = n_h\, d^E_{t,h,c(i)}Q^E_{t,c(i)}$ and hydrogen generation is $S^H_{t,h,i} = \eta_{t,h,i}D^E_{t,h,c(i)}$.

2.15 Power-to-Hydrogen with co-production of district heat (PX)

PX plants are identical to EF plants, with the exception that they co-produce $1/\kappa_{t,i}$ units of district heat per unit of hydrogen output.[17] The profit margin per unit of electricity input becomes

$$\Delta_{t,h,i} \equiv \eta_{t,h,i}\left(p^H_t + \frac{p^{DH}_{t,h,g(i)}}{\kappa_{t,i}}\right) - c_{t,h,i}, \tag{56}$$

with the remaining conditions as for the EF plant. Hydrogen generation is $S^H_{t,h,i} = \eta_{t,h,i}D^E_{t,h,c(i)}$ and district heat generation is $S^{DH}_{t,h,i} = S^H_{t,h,i}/\kappa_{t,i}$.

2.16 Other plant types

The input data includes a few additional plant types:

- Hydrogen producers relying on natural gas and CCS. There is only a single plant in Germany doing this. This is currently not included in the model, but should be added in particular if we add a model of the European hydrogen market more carefully.

[17] Note that this interpretation differs slightly from CHP plants where $\kappa_{t,i}$ is defined as the electricity-to-district heat ratio.

- Similarly, we have "plants" that are importers of hydrogen from areas that are not covered in the model. This is not included, but should be if we add the European hydrogen market model.
- Flexible demand (FD) appears to aggregate blocks of demand for each electricity area that only demand a fixed level if prices exceed either 2000 DKK/MWh or 2777 DKK/MWh. Currently, we model flexible demand in an alternative way (see section on demand).
- There are two Danish carbon capture plants (CC). This is currently not included.

3 Demand for electricity and district heating

The model includes several different types of consumers of electricity. First, a range of the energy producing plants included in the model are also consumers of electricity and district heating. This includes various storage types that can charge from the grid (e.g. pumped hydro, batteries), electrical heaters/heat pumps, and power-to-hydrogen facilities. Their hourly demand is detailed in the previous section, and enters the market clearing conditions under the label "plant demand".

In the following, we present methods for modeling other groups of consumers in a more aggregated sense that focuses on habits and partially flexible demand. Causally, the general equilibrium model determines the yearly level of demand for electricity and district heating given yearly average prices. Essentially, this means that demand flexibility should in general be formulated in ways that rearrange demand intertemporally, but do not change the overall level of yearly demand. The current version separates these two margins explicitly: an *hourly allocation* of a given yearly level (section 3.1) and an *endogenous yearly level* that responds to the consumer price index (section 3.2).

3.1 Discrete choice like model with Independence of Irrelevant Alternatives

Let $D^E_{t,c}$ denote yearly consumption and $\mu_{h,c}$ exogenous habits, normalised such that $\sum_h n_h \mu_{h,c} = 8760$.[18] Let $\underline{p}^E < 0$ denote a global price floor in the model. The hourly share of demand is

$$\frac{D^E_{t,h,c}}{D^E_{t,c}} = \frac{n_h \mu_{h,c} \left(\frac{p^E_{t,h,g(c)} - \underline{p}^E}{\overline{p}^E_{t,c} - \underline{p}^E} \right)^{-\sigma_{t,c}}}{\sum_{h'} n_{h'} \mu_{h',c} \left(\frac{p^E_{t,h',g(c)} - \underline{p}^E}{\overline{p}^E_{t,c} - \underline{p}^E} \right)^{-\sigma_{t,c}}}, \tag{57}$$

where $\overline{p}^E_{t,c}$ is a normalising price index. Three remarks:

- With constant hourly prices, the demand shares are exactly equal to $n_h \mu_{h,c}/8760$.
- As long as the individual hour is relatively insignificant compared to the yearly aggregate, $\sigma_{t,c}$ becomes the (approximate) elasticity of substitution for hourly demand.
- We allow for negative prices, but never crossing the threshold set by the price floor. The floor is imposed as a lower bound on the price variables themselves, $p^E_{t,h,g} \geq \underline{p}^E + 10^{-4}$, so that the expression is always well defined.

$\overline{p}^E_{t,c}$ is held *exogenous* within a solve (`pEt_C_cIdx_exo`); since it appears in both the numerator and the denominator it cancels out of (57) and only serves as a numerical normalisation. It is re-anchored to the model-consistent consumer price index between solution stages, cf. section 9. The corresponding expressions for district heating replace E by DH and $\underline{p}^E$ by $\underline{p}^{DH}$ throughout.

Parameterisation. In the data step, consumers not already covered by section 2 are split into a price elastic share (1/3) and a price inelastic share (2/3). In line with evidence from Danish

[18] Note that in this formulation, we can also rescale habit parameters $\mu_{h,c}$ to instead sum to one; the demand specification remains the same.

residential demand for electricity [Andersen and Dietrich, 2026], the elasticity of the price elastic part is set to 2.6 and the inelastic part to 0.1; for district heating, all consumers are given an elasticity of 0.5. *However*, the baseline currently overrides these values and sets $\sigma_{t,c} = \sigma_{t,c}^{DH} = 0.1$ for all consumer segments. The elastic/inelastic split therefore exists in the database and in the consumer index, but is not active in the baseline reported here; the split is retained so that it can be switched on without rebuilding the database. Similarly, the price floors are set to -100 DKK/MWh in the data step and overridden to -500 DKK/MWh in the baseline.

Note: It is important to adjust parameters if we implement new consumers, e.g. smart charging of electric vehicles, because a fair share of the flexibility comes through this consumer segment. Also, we note that the Independence of Irrelevant Alternatives assumption here is not very plausible; it is only adopted to keep things simple. A more dynamic view of demand flexibility is a relevant extension. Also, we do not model any costs related to consumers switching demand around.

3.2 Endogenous yearly demand

In the previous version, the yearly demand level $D_{t,c}^{E}$ was strictly exogenous to the energy module. The current version makes it a function of the endogenous consumer price index, so that the energy module can be run on its own with a demand response at the yearly margin:

$$D_{t,c}^{E} = u_{t,c}^{E} \left(\frac{\overline{p}_{t,c}^{E}}{p_{t,c}^{E,C}} \right)^{\sigma_{t,c}^{E,t}} \tag{58}$$

$$D_{t,c}^{DH} = u_{t,c}^{DH} \left(\frac{\overline{p}_{t,c}^{DH}}{p_{t,c}^{DH,C}} \right)^{\sigma_{t,c}^{DH,t}}, \tag{59}$$

where $p_{t,c}^{E,C}$ is the endogenous, quantity-weighted consumer price index of section 3.4, $\overline{p}_{t,c}^{E}$ is the same exogenous reference index that appears in (57), $\sigma_{t,c}^{E,t}$ is the elasticity of yearly demand, and $u_{t,c}^{E}$ is a level parameter. Note the direction: a rise in the endogenous index relative to the reference reduces yearly demand.

The level parameters $u_{t,c}^{E}, u_{t,c}^{DH}$ can either be set directly (in the baseline they are set to the exogenous yearly demand levels $D_{t,c}^{E}$ from the data) or calibrated so that the model reproduces a given yearly demand path. The latter is available as a separate block,

$$D_{t,c}^{E,\text{data}} = u_{t,c}^{E} \left(\frac{\overline{p}_{t,c}^{E}}{p_{t,c}^{E,C}} \right)^{\sigma_{t,c}^{E,t}} + J_{t,c}^{E}, \tag{60}$$

which can be run in two states: one where $J_{t,c}^{E}$ is endogenous and $u_{t,c}^{E}$ exogenous (measuring the residual), and one where $u_{t,c}^{E}$ is endogenous and $J_{t,c}^{E}$ exogenous (calibrating the level parameter). Neither state is used in the current baseline, where $\overline{p}_{t,c}^{E}$ is instead re-anchored to $p_{t,c}^{E,C}$ so that the ratio in (58) equals one at the point of introduction and $J_{t,c}^{E} = 0$ by construction.

In the baseline, $\sigma_{t,c}^{E,t}$ and $\sigma_{t,c}^{DH,t}$ are initialised at zero – which makes yearly demand exogenous and reproduces the behaviour of the previous version – and then raised to 0.1 for the consumer segments that have an hourly variation profile, as one of the stages of the solution procedure.

3.3 Distribution losses

Consumer demand is measured at the point of consumption, whereas the market clears at the level of the electricity area. The two differ by distribution losses. We let $\delta_{t,g} \in (0, 1]$ denote the distribution efficiency (one minus the loss rate) and let demand enter the market clearing condition as $D^E_{t,h,g}/\delta_{t,g}$; the implied loss volume,

$$L^E_t = \sum_{g \in DK} \left(\frac{1}{\delta_{t,g}} - 1 \right) D^E_{t,g}, \tag{61}$$

is reported as a separate item in the yearly energy balance. Losses apply to the aggregated consumer segments only, not to the demand of energy-producing plants, which are assumed to be connected at transmission level.

3.4 Price indices

The model computes a set of quantity-weighted price indices, which serve three purposes: they enter the yearly demand functions (58), they are the natural object to hand back to the CGE model, and they are the main object of interest when reporting results. For a consumer segment c the index is

$$p^{E,C}_{t,c} \sum_h D^E_{t,h,c} = \sum_h p^E_{t,h,g(c)} D^E_{t,h,c}, \tag{62}$$

i.e. the consumption-weighted average price faced by that segment. Analogous definitions give

- consumer price indices by area, $p^{E,C}_{t,g}$ and $p^{DH,C}_{t,g}$, weighted by the demand of the aggregated consumer segments in that area;
- producer price indices by area, $p^{E,S}_{t,g}$ and $p^{DH,S}_{t,g}$, weighted by production;
- the corresponding Danish aggregates $p^{E,C}_{t,DK}$, $p^{DH,C}_{t,DK}$, $p^{E,S}_{t,DK}$, $p^{DH,S}_{t,DK}$;
- producer price indices split by production sector, $p^{E,S}_{t,DK,s}$ and $p^{DH,S}_{t,DK,s}$ for $s \in \{\text{Waste}, \text{NonWaste}\}$, where the waste sector consists of the incineration plants (`IncinerationBH`, `IncinerationKV`);
- and the average export and import prices for Denmark,

$$p^{E,X}_{t,DK}\, X^E_{t,DK} = \sum_{h,l,g \in DK, g' \notin DK} p^E_{t,h,g}\, X_{t,h,l,g \to g'} \tag{63}$$

$$p^{E,M}_{t,DK}\, M^E_{t,DK} = \sum_{h,l,g \in DK, g' \notin DK} p^E_{t,h,g'}\, M_{t,h,l,g' \to g}, \tag{64}$$

  where the import price is evaluated at the price in the *sending* area, so that the difference between the two indices is the congestion rent plus transmission losses.

The segment-level indices $p^{E,C}_{t,c}, p^{DH,C}_{t,c}$ are always endogenous, since they feed back into demand. The remaining indices are pure reporting variables and are only made endogenous in the `priceIndex_regional` state, which the baseline uses.

4 Trade in electricity

Electricity areas are connected via transmission lines l. Transmission lines are characterized by an efficiency rate $\eta_{t,l} \in (0, 1)$ and export capacities in each direction. Rather than a single capacity with a relative reverse capacity, the current implementation stores a directed capacity $Q_{t,l,g \to g'}$ for each ordered pair of areas that the line connects, so asymmetric lines are handled directly in the data.

Transmission capacities may vary over the course of a year. We let $CapF^X_{t,l}$ denote the capacity factor for exports and $\gamma^X_{h,v(l)}$ the hourly-to-average availability.

Short run optimization problem. In the short run within a year t, a system operator utilizes a given transmission line l to maximize available arbitrage income given hourly prices and installed capacities. Assuming that short run costs of utilising the line are normally distributed, the utilization rate in each direction is

$$x_{t,h,l,g\to g'} = \Phi\left(\frac{\eta_{t,l}p^E_{t,h,g'} - p^E_{t,h,g}}{\sigma_{t,l}}\right), \tag{65}$$

and gross flows are

$$X_{t,h,l,g\to g'} = x_{t,h,l,g\to g'} \times CapF^X_{t,l} \times \gamma^X_{h,v(l)} \times n_h \times Q_{t,l,g\to g'}. \tag{66}$$

The gross flows in the two directions are then netted at the level of the individual line, so that a line never carries flow in both directions at once:

$$NX_{t,h,l,g\to g'} = X_{t,h,l,g\to g'} - \eta_{t,l}X_{t,h,l,g'\to g} \tag{67}$$

$$X^{net}_{t,h,l,g\to g'} = \max\left(NX_{t,h,l,g\to g'}, 0\right), \qquad M^{net}_{t,h,l,g\to g'} = \max\left(-NX_{t,h,l,g\to g'}, 0\right). \tag{68}$$

Total exports and imports of an area are obtained by summing the net line flows over lines and counterpart areas,

$$X^E_{t,h,g} = \sum_{l,g'} X^{net}_{t,h,l,g\to g'}, \qquad M^E_{t,h,g} = \sum_{l,g'} M^{net}_{t,h,l,g'\to g}, \tag{69}$$

and these enter the market clearing condition of section 6. Note that the netting is applied before aggregation, which means that $X^E_{t,h,g}$ and $M^E_{t,h,g}$ can both be positive for an area with several interconnectors, as is the case for Denmark. We interpret the profit of the line operator, $\Pi_{t,l}$, as the congestion rent on line l.

Lines that connect two areas which are aggregated into the same area are removed from the data (section 8.3).

5 Markets for constrained domestic resources

Some of the fuels used in the Danish energy system are not available at an exogenous world market price. Waste, hazardous waste and biogas are available in quantities that are largely determined outside the energy system – by household and industrial waste generation and by the agricultural sector – and the energy system competes for a fixed supply. Treating them as perfectly elastic at a fixed price, as in the previous version, means that any policy that raises the value of these fuels leads to an implausible expansion of their use.

The current version therefore introduces a market for each of these resources. Let r index resource zones, currently $r \in \{\texttt{DK\_Affald}, \texttt{DK\_Farligt\_Affald}, \texttt{DK\_Biogas}\}$, and let $b(r)$ denote the base fuel associated with the zone (waste for the first two, biogas for the third). Each plant is mapped to at most one resource zone. We let $\overline{R}_{t,r}$ denote the available quantity of the resource in year t, taken from the Ramses resource-zone data, and $F_{t,r,b}$ the model's total fuel use of base fuel b by plants in zone r. The market clearing condition is

$$F_{t,r,b(r)} = \overline{R}_{t,r} - J^r_{t,r,b(r)}, \tag{70}$$

where J^r is a residual term used in the calibration (see below). The associated shadow price $p^r_{t,r,b}$ is added to the fuel cost of every plant in the zone,

$$c^r_{t,i} = \sum_b \alpha_{f(i),b}\, p^r_{t,r(i),b}, \tag{71}$$

and enters marginal costs alongside the exogenous fuel price index, cf. (1). Since $\alpha_{f(i),b}$ are the fuelmix shares, a plant that only partially relies on the constrained resource only pays the shadow price on that part of its fuel input.

Economically, $p^r_{t,r,b}$ is the scarcity rent on the resource. In the baseline it is calibrated so that the model reproduces the observed use of each resource: the equilibrium condition (70) is first imposed with p^r fixed at its data value and J^r endogenous, and J^r is subsequently driven to zero with p^r endogenous (section 9). Two implementation details are worth noting:

- The resource zone is treated as a categorical attribute of the plant, so plants in different resource zones are never aggregated (section 8.4).
- The biogas market is deliberately *not* cleared in the terminal year 2050. Imposing the constraint in 2050 with the current data leads to a corner where the shadow price has to rise without bound; leaving the terminal year out lets the model treat biogas as elastically supplied at the exogenous price in that year.

6 Market clearing

The market clearing conditions for electricity and district heating are

$$D^{E,\text{plants}}_{t,h,g} + \frac{D^{E,\text{cons}}_{t,h,g}}{\delta_{t,g}} + X^E_{t,h,g} = S^E_{t,h,g} + M^E_{t,h,g} + J^E_{t,h,g} \tag{72}$$

$$D^{DH,\text{plants}}_{t,h,g} + D^{DH,\text{cons}}_{t,h,g} = S^{DH}_{t,h,g} + J^{DH}_{t,h,g}, \tag{73}$$

where $D^{\cdot,\text{plants}}$ is the demand of energy-producing plants (storage charging, electrical heaters, PtX) and $D^{\cdot,\text{cons}}$ is the demand of the aggregated consumer segments of section 3. The distribution efficiency $\delta_{t,g}$ applies only to the latter, and only in areas for which yearly consumer demand data exist.

The terms $J^E_{t,h,g}, J^{DH}_{t,h,g}$ are residuals. They are used in exactly the same way as J^r in the resource-zone market: in the first stage of the solution procedure, prices are fixed at their starting values and the J-terms absorb any excess demand; they are then driven to zero while prices become endogenous. In a converged baseline the J-terms are zero and (72)–(73) are ordinary market clearing conditions. There is no market clearing condition for hydrogen: hydrogen is priced at the exogenous world market price p^H_t and any excess supply is implicitly exported.

Note that the market clearing conditions are only imposed for *active* areas, i.e. pairs (t, g) for which at least one plant exists in the data.

6.1 What clearing looks like

The plant-by-plant treatment of Section 2 is easier to hold onto if one sees what it adds up to. Figures 3 and 4 are the same state of the same area — West Denmark in 2035, a June state with 6.4 GW of solar and 2.3 GW of wind available at once — drawn first as a supply curve and then as a clearing diagram on the same axes.

Figure 3 shows the supply side. The coloured blocks are the plants, each placed at its effective bid price and given the width of the capacity the state makes available to it, ordered from cheapest to dearest: the classical merit order. The black line is the model's actual supply function, $\sum_i Q^E_{t,i} s^E_{t,h,i}(p)$, which is what the market clearing condition above sees. Two features of the model are visible in the gap between the two. The line is smooth where the blocks are a staircase, which is the dispatch smoothing of Section 2.2. And the line lies below the blocks at the right-hand end, because the storage plants' blocks are gross offers while the line is net of their charging.

Figure 4 adds the demand side. The dashed curve is domestic demand alone; the solid curve adds net exports, which are a function of the local price through the smoothed export rule of Section 4. Where the solid curve crosses supply is the price the model solves for, and the marker confirms that it does. The right-hand panel splits net exports by trading partner: at the solved price West Denmark is importing 2.1 GW from Norway, Sweden and East Denmark while simultaneously exporting 5.8 GW to Germany, the Netherlands and Britain. Importing and exporting at once is not an artefact of averaging hours within a state: the neighbours differ in price, and each line responds to its own price difference.

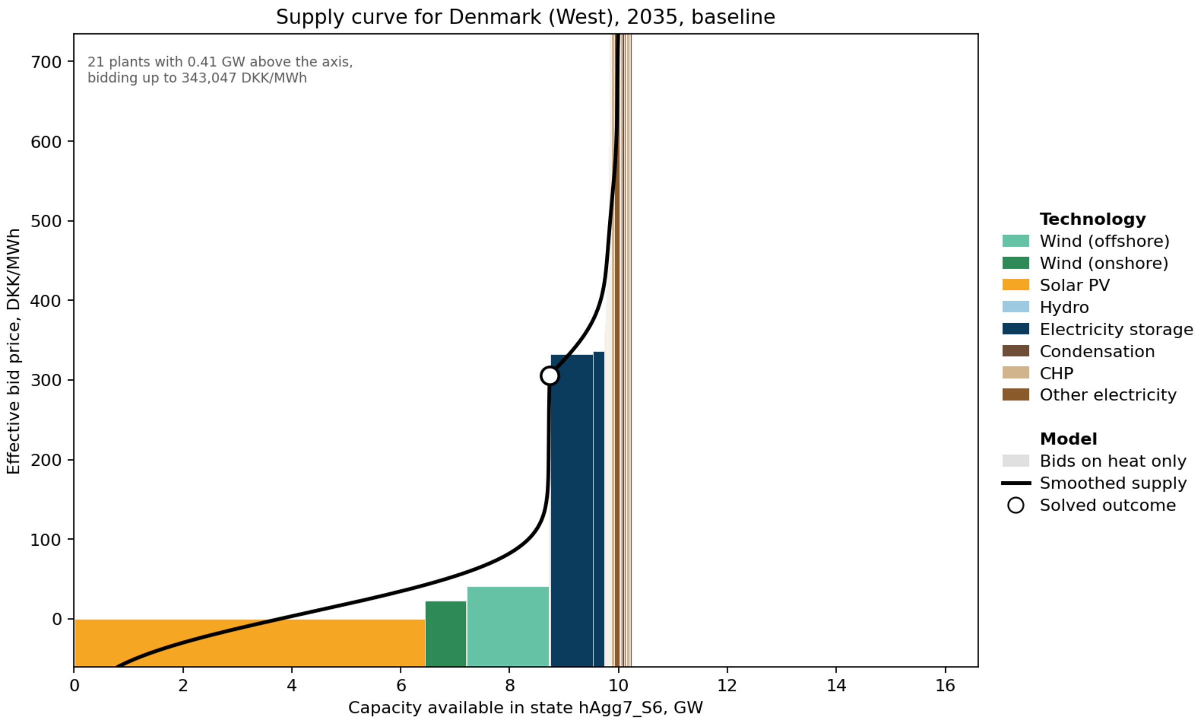


Figure 3: Supply in West Denmark in one state of 2035. Blocks: individual plants at their effective bid price, widths equal to available capacity. Line: the model's supply function. The quantity axis is the one used in Figure 4, so the two can be read against each other.

The same picture in the tightest state of the year, Figure 5, shows what the trade block is carrying. With no solar and 0.5 GW of wind available, domestic supply tops out at 2.5 GW against a domestic demand that does not fall below 5.3 GW at any price in the range: the two curves never cross, so the area has no autarky price at all, and the price that clears the market — 1,229 DKK/MWh — is set entirely by what the interconnectors will deliver. States of this kind are a small share of the year but a large share of what a question about security of supply or about the value of transmission turns on.

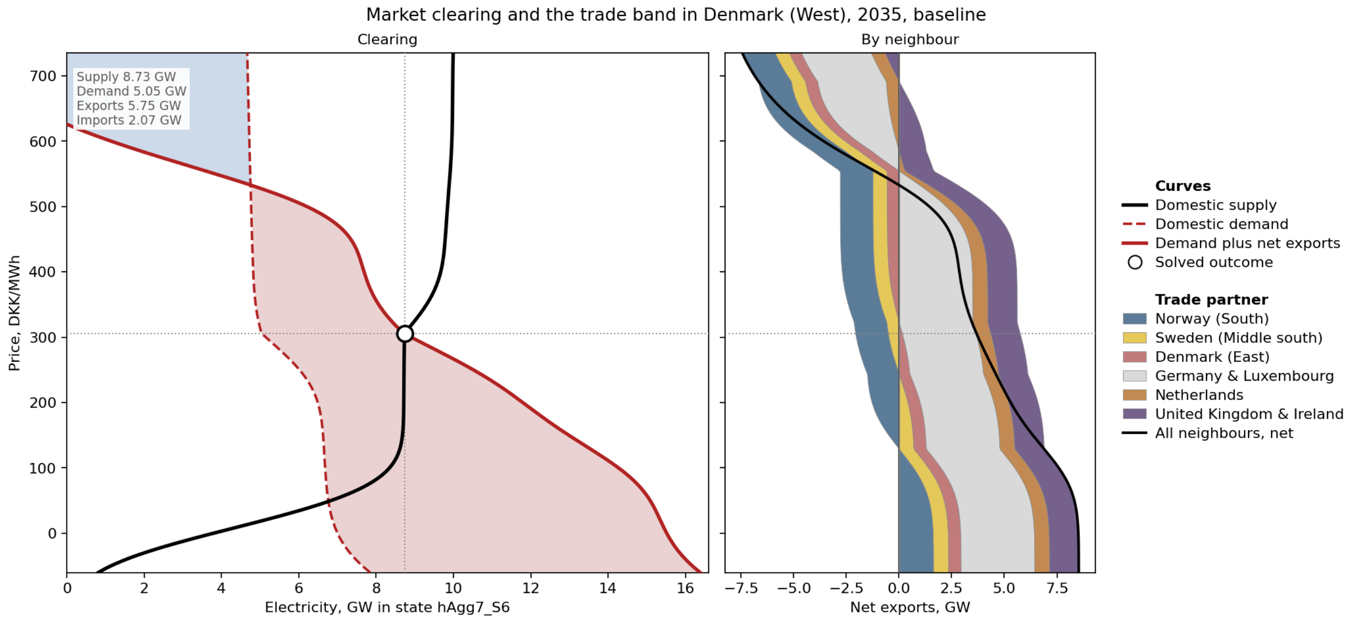


Figure 4: Market clearing in West Denmark, same state as Figure 3. Left: supply, domestic demand, and demand plus net exports. Right: net exports decomposed by trading partner, ordered by the price at which each switches between buying and selling.

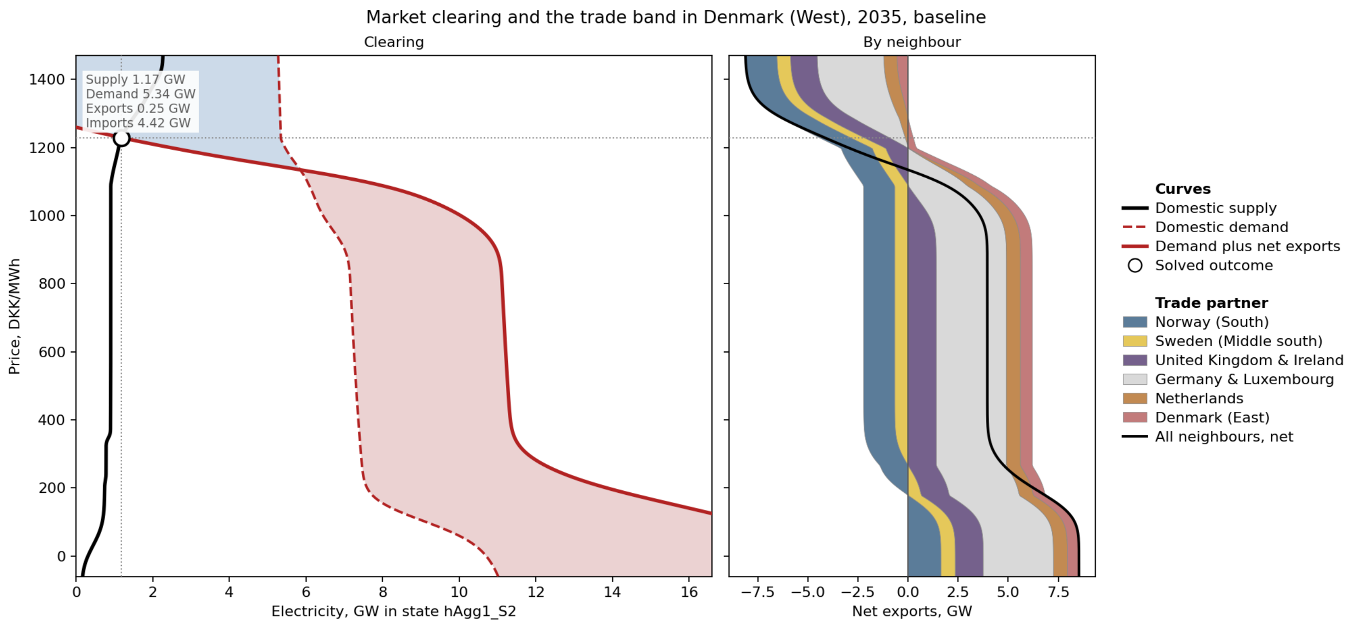


Figure 5: Market clearing in West Denmark in the dearest of the 120 states of 2035. Domestic supply and domestic demand do not cross within the price range shown.

7 Aggregates and the interface with the CGE model

A separate module collects the quantities that the rest of the model, and the reporting layer, need. It contains no behaviour; every equation is an accounting identity. It is included here because the fuel accounting in particular is not a trivial identity for CHP plants.

Production and demand. Hourly production of electricity, district heat and hydrogen is summed over plants within each area, and yearly totals are formed by summing over hours. The same is done for demand, split by whether the consumer is an energy-producing plant or an aggregated consumer segment. Danish totals sum the areas `DK1` and `DK2`. Production is additionally split by production sector $s \in \{\text{Waste}, \text{NonWaste}\}$, where the waste sector consists of the incineration plants.

Fuel use. Total use of base fuel b in area g is

$$F_{t,g,b} = \sum_h \left[\sum_{i \in \{CD,EPE,BP\}} \alpha_{f(i),b} \frac{S^E_{t,h,i}}{\eta_{t,i}} + \sum_{i \in \{EX,BB\}} \alpha_{f(i),b} \frac{n_h Q^E_{t,i} s^E_{t,h,i}}{\eta_{t,i}} + \sum_{i \in \{BH,EPH\}} \alpha_{f(i),b} \frac{S^{DH}_{t,h,i}}{\eta_{t,i}} \right]. \tag{74}$$

The three groups differ in what the conversion efficiency is defined relative to. For CD, EPE and BP plants, $\eta_{t,i}$ is the efficiency of electricity generation, so fuel use follows from electricity output. For BH and EPH plants it is the efficiency of heat generation. For EX and BB plants, neither output alone is proportional to fuel use, since the electricity-to-heat mix varies with $u^{BP}_{t,h,i}$; fuel use is instead proportional to the *composite* output $Q^E_{t,i} s^E_{t,h,i}$, which is why these plants appear with their own term. The same expression restricted to plants in a given resource zone gives $F_{t,r,b}$ in (70).

Fuel use by tax purpose. For the interface with the tax model, fuel use is further split into the part attributed to electricity and the part attributed to heat, using the allocation factors of section 2.8. For a BP plant under the E-rule, for instance, the electricity-attributed fuel is $S^E_{t,h,i}/\overline{\varepsilon}$; under the V-rule it is $S^E_{t,h,i}/\eta_{t,i} - (S^E_{t,h,i}/\kappa_{t,i})/\overline{\nu}$. For BB and EX plants the split is formed mode by mode and weighted by $u^{BP}_{t,h,i}$. The heat-attributed part is the residual, $F^H_{t,s,b} = F_{t,s,b} - F^E_{t,s,b}$, which guarantees that the two parts add up to total fuel use by construction. These quantities are computed per production sector s.

What is handed to the CGE model. The energy module delivers yearly quantities (production, demand, net exports, fuel use by base fuel and by tax purpose, distribution losses) and yearly price indices (consumer and producer, by area and for Denmark, plus export and import prices). It receives yearly demand levels, fuel prices, tax rates, capacities and the ETS price paths.

8 Aggregation

Model data can be aggregated in various ways to reduce computational costs. Currently, the implementation allows for reduction of the following dimensions: fuel types f, variation types v, geography g, producers i, consumers c, and hours h. In some instances the aggregation is straightforward (e.g. fuel types, variation types), whereas other areas are more complicated and allow for multiple approaches (e.g. hours).

8.1 Fuel types

In the raw data, we have roughly 200 f types (latest update from Climate Outlook 2025). The only symbol that depends on f is the fuel mix parameter $\alpha_{f,b}$ that defines the share of base fuels b that each fuel type consists of.

Aggregation of fuel types is done using KMeans clustering [Pedregosa et al., 2011] with 80 clusters; with the latest data this ensures that the largest difference compared to original data is less than 1 percentage point. The number of clusters is currently set manually. The direct effect of the size of f on runtime is negligible, but it also changes how producers/consumers may be aggregated; this aggregation is essential to reduce computational costs (see 8.4).

8.2 Variation types

In the current version, our input data has close to a thousand different base variation types v. In a version where hours are clustered, the impact of reducing the number of base variation types is negligible; for detailed hourly models this can, however, be important. We can use archetypal analysis to reduce the number of base variations [Alcacer, 2025]. This produces a number of new base types v and coefficients representing linear combinations that are mapped to producer and consumer sets. This step is not active in the current data pipeline.

8.3 Geography

We operate with geography sets for electricity and heat areas; they identify the relevant markets that producers/consumers engage in. The current raw data includes roughly 30 European electricity areas and 85 Danish district heat areas.

Aggregation is done by specifying many-to-one mappings manually. The baseline keeps twelve electricity areas explicitly – `DK1`, `DK2`, `SE1–SE4`, `NON`, `NOM`, `NOS`, `DELU`, `NL`, `GBNIIE` – and aggregates the remainder into a single "rest of world" area, excluding a handful of areas that are dropped entirely (`ROHRRS`, `RU`, `SIGR`, `UA`). District heat areas cannot be dropped, since they are all in Denmark; the default is to aggregate them to the corresponding electricity area, which leaves two heat areas. This is the most aggressive aggregation in the model, and Section 10.2.1 argues that it is the likely source of the largest discrepancy between the baseline and the projection: with one heat market per price zone, a plant in any town can serve heat demand in any other, and the cheapest heat in the zone sets the margin everywhere.

In most instances this aggregation is straightforward: categorical data like the location of a producer remains unique, $g(i)$. For transmission lines, however, aggregation of electricity areas results in lines that now operate within zones. These lines are currently simply removed from the data, but we could look into handling potential bottlenecks within zones in some way.

Consumer indices are aggregated after the geographic aggregation, grouping consumers of the same type within each area.

8.4 Producers

In the raw model data, we have roughly 20,000 different producers. The aggregation of this set is of first order importance, because almost all equations and parameters in the module are defined over this set (a few exceptions are equilibrium conditions and definitions of yearly outcomes).

The aggregation method ultimately relies on many-to-one mappings from full to aggregated ids. However, the many-to-one mapping has to be aligned with categorical data, so that ids that differ on any of the following are never aggregated: technology type (Pt), electricity area, heat area, fuel type, subsidy type, duration type, bidding type, technology label (`TechType`, which identifies incineration plants), CHP tax rule (section 2.8), and resource zone (section 5). The last three categories are new relative to the previous version and raise the lowest attainable number of aggregate ids somewhat; with the current data the model is solved with a few hundred producer ids.

The duration type is a categorical index that maps individual storage plants into "types" based on common duration levels. Within each group defined by the categories above, we check whether the durations d_i of the plants in the group satisfy $\max(d)/\min(d) \leq \overline{d}_{rel}$ or $\max(d) - \min(d) \leq \overline{d}_{abs}$, with $\overline{d}_{rel} = 2$ and $\overline{d}_{abs} = 3$ in the baseline. If not, we apply KMeans to the transformed durations

$$\tilde{d}_i = \log\left(1 + d_i/\text{lb}\right), \qquad \text{lb} = 3, \tag{75}$$

increasing the number of clusters by one at a time until the tolerance is met. The transformation puts weight on relative rather than absolute differences, and de-emphasises differences in the region 0–lb.

Aggregation procedure. At the time of aggregation, the database contains roughly 30 symbols defined over producer ids. For a few variables the aggregation defines new aggregate capacities as the sum of the underlying individual ones (`ECap`, `HCap`, `HyCap`, `StorageCap`). In most instances, however, variables represent ratios and efficiencies; for these we aggregate using weighted averages based on relative capacities. For combined heat and power plants, the weights are the sum of capacities across heat and power. For storage plants, we rely on a nonlinear combination of storage and dispatch capacities, $Q_{t,i}\left[1 + \log\left(1 + \zeta_{t,i}\right)\right]$, where $\zeta_{t,i}$ is the duration; storage types with vastly different durations should generally not be aggregated. For a few variables – duration and charging speed – we recompute the value from the aggregated capacity definitions instead of using a weighted average.

For producers that are also consumers, we subsequently aggregate the relevant consumer indices in a similar fashion: hourly demand capacities are added together, while other symbols are weighted together based on relative capacities.

Clustering and identification of smoothing parameters. Currently not implemented – the smoothing parameters are set to $\sigma_{t,i} = 50$ for all plants and $\sigma_{t,l} = 50$ for all transmission lines, except lines connecting to the rest-of-world area, where $\sigma_{t,l} = 200$. The smoothing parameter governing the mode choice of flexible CHP plants is set to $\sigma_{t,i}^{EH} = 10$. Two ideas for doing better:

- Estimate the smoothing parameter based on aggregation. Compute the static supply curve (take dynamics out of it) on a grid of prices based on aggregate ids and on the disaggregated curve, and use the σ that minimizes the distance between the two.
- For storage plants in particular, we likely need a different type of estimation (e.g. based on Berg and Eskildsen, 2021b).
- We can use smoothing estimation to determine the number of clusters within each categorical type: if σ is larger than some threshold, try to increase the number of clusters by one and check again.

8.5 Hours

Currently, we perform one of two types of aggregation: adjacent chronological aggregation or KMeans clustering. In both instances, we map the original 8760 hours to new aggregate short run states k. Let S_k denote the subset of hours h that are mapped to state k and $n_k = |S_k|$ the number of hours in the state. Before clustering, the parameter $\gamma_{h,v}$ indicates hourly-to-average variations for different types v (thus with mean one).

In both instances, we define new cluster values $\gamma_{k,v}$ in one of three ways:

- **Mean aggregation:** $\gamma_{k,v}$ is defined by averaging over hours mapped to the new state k.
- **Fitting distributions separately:** we can fit the distribution of $\gamma_{k,v}$ to match the distribution in the original data $\gamma_{h,v}$. This is done by sorting the clustered and original data in ascending order for each v separately, and assigning values to $\gamma_{k,v}$ based on the sorting

of the original data. This ensures that the data are valid in terms of the volatility of the original data, but comes at the expense of a worse representation of correlations between different variation types v.

- **Fix extreme values:** in addition to fitting the distribution, we can target extreme values of the variation data. For the clusters that represent the min/max of a given variation type, we set $\gamma_{k,v}$ to the min/max from the full hourly distribution instead of the mean within the lowest/highest cluster. Afterwards the profile is rescaled so that $\sum_k n_k \gamma_{k,v} = 8760$.

Both options are switched on in the baseline.

Chronological aggregation. Blocks of `nAdjacent` consecutive hours are collapsed into one state. The baseline chronological configuration uses $8760/24 = 365$ adjacent hours per block, i.e. 24 states, each representing 365 consecutive hours.

Non-chronological clustering. The year is first split into seasons, and KMeans is then applied within each season separately. The baseline uses calendar months as seasons (12 seasons) with 10 clusters per season, i.e. 120 states. The clustering data consist of the standardised variation profiles $\gamma_{h,v}$, augmented – when `useAux` is switched on, as in the baseline – with four aggregate series computed at full hourly resolution: variable renewable electricity supply, solar heat supply, habitual electricity demand, and habitual heat demand, together with the two residual-demand series formed from them. Adding these makes the clustering sensitive to the hours that matter for prices, rather than only to the shape of the underlying variation profiles.

Transition dynamics with non-chronological clustering. In the case where data are clustered non-chronologically, we use the mapping from hours to clusters to identify transition probabilities between the k new clusters, using relative frequencies. Let $n_{k,k'}$ denote the number of hours where $h \in S_k$ and $h+1 \in S_{k'}$. The probability that cluster k transitions to cluster k' is

$$\Pr\left(k \to k' \,\middle|\, k\right) = n_{k,k'}/n_k, \tag{76}$$

and, given that the current cluster is k, the probability that the preceding cluster was k' is

$$\Pr\left(k' \to k \,\middle|\, k\right) = n_{k',k}/n_k. \tag{77}$$

The transition count is cyclical: the last hour of the year is treated as transitioning to the first, so that the storage problem closes on itself over the year. Optionally, transition probabilities below a threshold can be dropped and the matrix renormalised, which reduces the density of the storage equations; season-boundary transitions are always preserved so that the chronological ordering of the seasons survives. This is not used in the baseline.

These probabilities define a data generating process for cluster values $\gamma_{k,v}$ under a simple Markov assumption. However, while this provides a simple way to generate variation data, there are hundreds if not thousands of endogenous state variables in the model for each cluster, which makes a full dynamic optimization problem practically infeasible. Instead, we choose state averages as sample vectors, cf. section 2.12.3.

Figure 6 shows the resulting matrix for the baseline clustering. It is nearly block-diagonal: because the seasons are calendar months and clustering is done within each month, almost all probability mass sits in the ten-by-ten block belonging to the current month, with a thin band at the month boundaries carrying the year forward. This is what allows a storage plant to move energy through the year at all in the non-chronological version — and equally, it is why the model represents seasonal storage far better than it represents a battery cycling within a day, since a transition out of a state and back into it within the same month is the only thing a daily cycle can be built from.

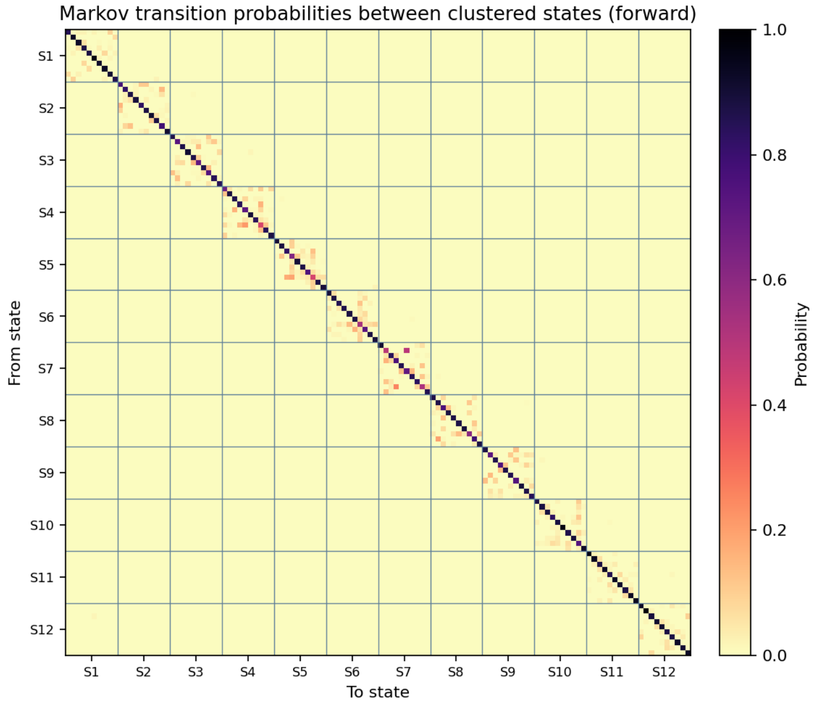


Figure 6: Transition probabilities between the 120 clustered states, $\Pr(k \to k')$. The block structure is the twelve calendar months.

What the aggregation costs. Clustering hours is not free, and the cost falls almost entirely on the tail of the price distribution. Averaging within a state cannot preserve an extreme hour: the most extreme cluster is at most as extreme as the mean of its members. This can be measured without reference to the model at all, by taking the published hourly prices and averaging them within the model's own states, and Section 10.2 does exactly that. The result is that roughly seven tenths of the model's understatement of the highest prices is attributable to the clustering rather than to the economics. Anything that depends on the upper tail — scarcity rents, peak capacity values, the profitability of assets that earn their income in a few hours a year — should be read with that in mind, and it is the main argument for the chronological version and for the projects in Sections 11.6 and 11.8.

9 Solution and calibration

The model is a large square system of nonlinear equations, solved with GAMS as a constrained nonlinear system (CNS). Two features of the model make a direct solve from an arbitrary starting point unreliable: the dispatch decisions are compositions of normal CDFs, which are flat far from the switching point and therefore give poor Newton directions; and the storage problem couples all hours of the year through the continuation value. The solution procedure addresses this in two ways: a homotopy that walks the model into equilibrium one closure at a time, and a fallback that relaxes the smoothing parameters when a stage fails.

9.1 The homotopy

Every stage solves the full model; what changes between stages is which variables are endogenous. Each stage is warm-started from the checkpoint left by the previous one, so the solution of one stage is the starting point of the next.

1. `exo_prices`. All prices – hourly electricity and district heat prices and the resource-zone

shadow prices – are fixed at values loaded from a previous solution, and the residual terms J^E, J^{DH}, J^r are endogenous. Every equation in the model except the market clearing conditions is imposed. This stage is a pure evaluation of the supply and demand side at given prices; it is the only stage that is solved cold (i.e. without a checkpoint), and it always solves, since the J-terms can absorb any imbalance. Its purpose is to produce a consistent set of levels for all remaining variables.

2. `endo_prices`. The electricity and district heat prices are made endogenous and J^E, J^{DH} are set to zero, so that (72)–(73) become genuine market clearing conditions. The resource-zone prices remain fixed and J^r remains endogenous. The implementation multiplies the J-terms by a step factor, so the stage can be run as a sequence of partial steps if the jump to zero is too large; in the current baseline a single step suffices.

3. `endo_prices_rZone`. The resource-zone shadow prices are made endogenous and J^r is driven to zero, so that the resource markets (70) clear. Again this can be done in steps.

4. `SigmaCXt`. The elasticity of yearly demand is raised from zero to 0.1 for the consumer segments with an hourly variation profile. At the same time, the exogenous reference price index $\overline{p}^E_{t,c}$ is re-anchored to the endogenous consumer price index $p^{E,C}_{t,c}$ from the previous stage, so that yearly demand is unchanged at the point of introduction and the elasticity only bites for deviations from the baseline. The resource-zone markets are held fixed at their calibrated prices during this stage.

5. `baseline`. The final model definition, solved at the original parameter values with no fallback allowed. Its solution is the baseline that shocks are run against, and the corresponding checkpoint is the starting point for the shock notebook.

The stages differ only in the module states they combine; the underlying equations are the same throughout. Because each stage inherits the previous solution as its starting point, the sequence is a homotopy in the closure of the model rather than in its parameters.

9.2 Fallback for failed solves

Even with the homotopy, individual stages fail for some model versions – typically the versions with disaggregated electricity areas, and typically in years where a large capacity change makes the initial price guess poor. Rather than hand-tuning parameters for each version, every stage is called through one routine, `robustSolve`, which does the tuning itself.

The idea is that a stage which will not solve at the parameters one wants will usually solve at parameters that make the model flatter and better behaved — more smoothing, softer kinks, looser price floors — and that the solution found there is a good enough starting point to reach the parameters one wants after all. `robustSolve` therefore relaxes, solves, and then walks the relaxation back out, reporting whether it got all the way. Since every stage is called the same way, no version needs its own hand-written recipe; the routine finds the least relaxation each stage happens to need.

The routine is given a set of parameters it may adjust, together with a multiplier to apply at full relaxation. In the baseline these are: the dispatch smoothing parameters $\sigma_{t,i}$ ($\times 5$) and $\sigma_{t,l}$ ($\times 5$); the hourly demand elasticities $\sigma_{t,c}$ and $\sigma^{DH}_{t,c}$ ($\times 3$), which flatten the demand allocation; the storage cost parameters $x^{\Delta}_{t,i}$ ($\times 100$, a softer kink in the continuation value) and $x^{out}_{t,i}$ ($\times 0.2$, a milder penalty for exceeding storage bounds); and the price floors $\underline{p}^E, \underline{p}^{DH}$ ($\times 3$). Multipliers are geometric in a scalar relaxation level $\lambda \in [0,1]$: a parameter with factor θ is scaled by θ^λ, so $\lambda = 0$ restores the original values.

The search proceeds as follows. The stage is first solved at $\lambda = 0$. If that fails, the routine walks up a ladder of relaxation levels ($\lambda \in \{0.25, 0.5, 1\}$) until the model solves. It then attempts

$\lambda = 0$ again from the relaxed solution, which often succeeds even though it did not from the point where the stage started. If it does not, the routine bisects between the largest failing and the smallest succeeding λ to tighten the relaxation as far as possible. The number of extra attempts is capped by a solve count and a time budget, with the guarantee that every stage gets at least one relaxed attempt and one attempt at returning to the original values.

Three implementation points matter for interpreting results:

- Relaxation multipliers are applied relative to the parameter's own value, so heterogeneity in the underlying data is preserved (for instance, $\sigma_{t,l} = 200$ on rest-of-world lines stays five times larger than $\sigma_{t,l} = 50$ elsewhere at any λ).
- The multiplier applied by the previous solve is stored in the GAMS state and divided out again, so a relaxation inherited through a checkpoint is undone when a later stage is solved at $\lambda = 0$.
- The final `baseline` stage is called with the fallback disabled, so the reported baseline is always a solution at the original parameter values. If an earlier stage ended at $\lambda > 0$, the `baseline` stage undoes that relaxation, and the model either solves at the original values or the run is flagged.

9.3 Where the starting values come from

The first stage needs a price vector to start from. This is supplied by a stored dictionary of prices from a previous solution of the same model version (`price_dict`), which is refreshed whenever a converged baseline is produced. This means that the model is, in practice, always warm-started from a nearby solution, and that a substantial change in the data or in the model requires the first stage to be re-solved from a price vector that may be some distance away. The `exo_prices` stage is robust to this, since it does not impose market clearing, but the `endo_prices` stage may then need the fallback.

9.4 Running a shock

A shock is a re-solve of the `baseline` model definition from the baseline checkpoint with some parameters changed. Because the checkpoint holds the full converged state, nothing has to be rebuilt: the model definition, the data and the starting point are all inherited, and only the perturbed parameters differ.

The perturbation is written as GAMS statements that are prepended to the solve. This is the reason the mechanism is general rather than a fixed menu of shocks: any parameter or variable bound in the database can be assigned to, including in a way that depends on sets, so that "halve the capacity of every transmission line with a Danish endpoint" or "raise the CO_2 tax on non-ETS heat from 2030" are both a line or two of GAMS. Parameters that the model treats as fixed — capacities, in the baseline closure — are shocked by moving their bounds rather than their levels, since a level assignment on a fixed variable would be overwritten.

Two features are worth knowing. First, a large perturbation can be applied in steps, each solved from the previous one, which is the same warm-start logic as the homotopy applied to parameters instead of to the closure. Second, the shocked solve goes through `robustSolve` like any other, so a shock that puts the model somewhere awkward will be met with the relaxation ladder rather than a failure — with the same caveat as in the baseline, that the final solve is required to return to the original parameter values or the run is flagged.

The shock section of `2_Analysis.ipynb` is set up to halve the net transfer capacity on the interconnectors between Denmark and its neighbours, and to compare the result with the baseline. It is included as a worked example of the mechanism rather than as a piece of analysis, and the comparison figures it draws are the same ones used in Section 10.

10 Results

This section documents what the model produces when it is solved on the KF25 data. It is not a policy analysis. The purpose is twofold: to show that the baseline is a reasonable representation of the system it is calibrated against, and to illustrate the kind of output the model delivers. We first state the dimensions of the version being solved (Section 10.1), then compare the baseline with Climate Outlook 2025 on prices and on quantities (Section 10.2), and then describe the baseline itself (Section 10.3).

Everything below is the `clusterAvg` version on the `aggEA` geographical aggregation. The pipeline also solves the chronological version, the version with disaggregated electricity areas, and the KF26 overlay of Appendix B, and the solved databases for those are distributed alongside this one; we do not report on them here. All figures come from the diagnostic and plotting routines in `pyGRE`, which is what `2_Analysis.ipynb` calls. All prices are in fixed 2025-prices unless stated otherwise.

10.1 The version that is solved

The `clusterAvg` version on the `aggEA` geographical aggregation is the configuration used for the interface with the CGE model. Its dimensions are:

Dimension	**Size**
Years t	31 (2020–2050)
Short-run states h	120, of 20–178 hours each ($\sum_h n_h = 8760$)
Electricity areas g_E	13, of which DK1 and DK2 are domestic
District heat areas g_{DH}	2, one per Danish price zone
Producers i	390, of which 295 active in 2035
Consumers c	77
Transmission lines l	46, of which 41 active in 2035
Binding resource zones r	3: waste, hazardous waste, biogas

The 91 CHP plants are distributed over the three tax rules of Section 2.8 as follows: 43 plants on the V-rule (35 BP, 8 EX), 16 on the E-rule (12 BP, 4 BB), and 32 on the guarded V-rule (17 BP, 11 BB, 4 EX). The rule parameters are the statutory ones, $\overline{\nu} = 1.25$, $\overline{\varepsilon} = 0.65$ and $\overline{\gamma} = 0.35$.

Note that the 120 states are not hours: each is a cluster of hours with weight n_h, and transitions between clusters are governed by the Markov matrices of Section 8.5. This matters for how the results below should be read, and in particular for the price tail discussed next.

10.2 Validation against Climate Outlook 2025

The model is calibrated on the KF25 data, but it is not fitted to the KF25 price projection: prices are an equilibrium outcome. Comparing the two is therefore informative about whether the supply curve, the demand system and the trade block together reproduce the price formation in the projection they are built from.

Figure 7 compares the model's yearly consumer price index for electricity with the published KF25 prices. The two agree closely through 2035 in all regions. After 2035 they diverge: KF25 has prices rising steadily to 2050, whereas the model has them flat or falling until the late 2040s. The divergence is largest in the areas with a lot of hydro (Norway, Sweden), where the model settles at roughly half the KF25 level by 2050.

Table 1 puts numbers on the comparison at the hourly level, by mapping each of the 8760 hours to the state it belongs to and giving it that state's price. Three things stand out. First, the yearly averages line up well up to 2035: the model is 5–6 % above KF25 in 2030 and within 2–5 % in 2035. Second, the correlation across hours is 0.70–0.72 through 2035 and 0.64 in 2050,

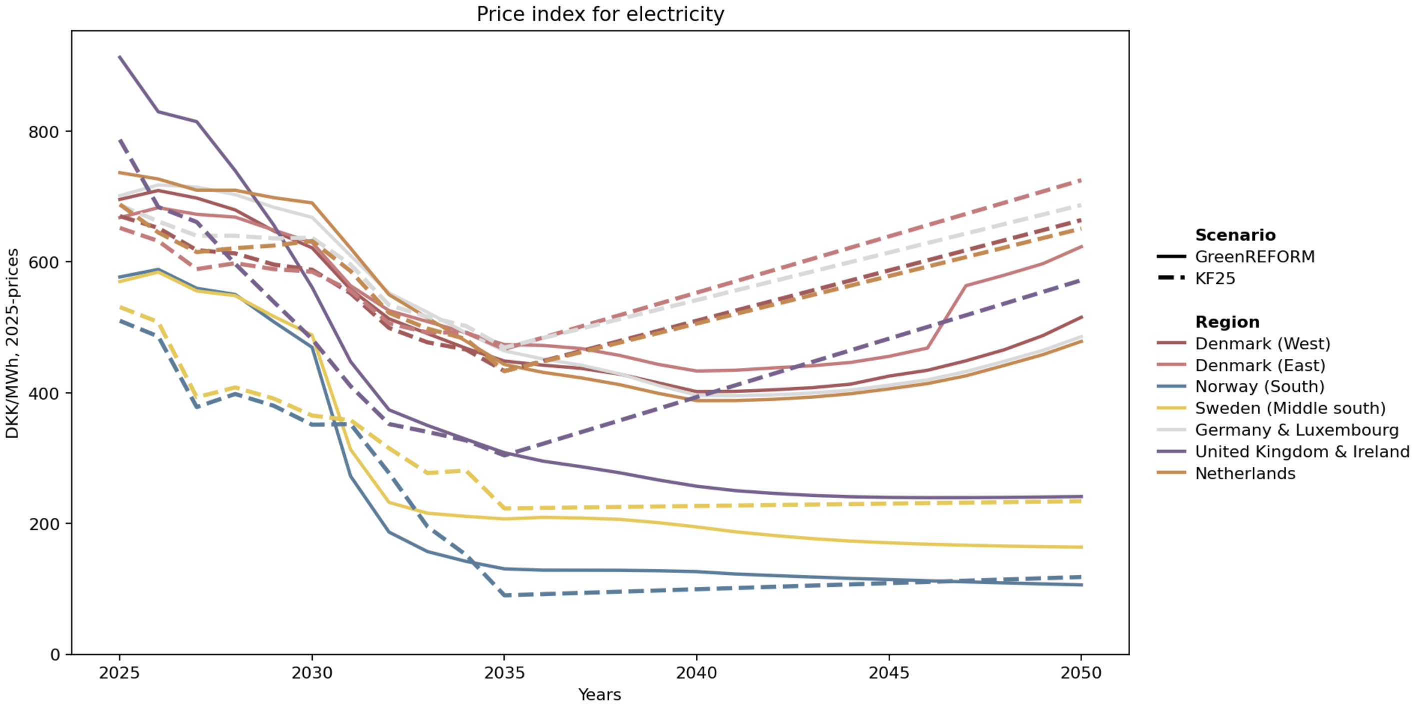


Figure 7: Yearly consumer price index for electricity. Solid: GreenREFORM baseline. Dashed: published KF25 prices.

so the model gets the timing of high and low prices broadly right, which is what one would hope for given that the clustering is estimated on the same underlying variation profiles. Third, the model understates the top of the price distribution: averaged over the 100 most expensive hours, it is at 1,199 DKK/MWh against 2,129 in KF25 in West Denmark in 2035.

Table 1: Model baseline against published KF25 hourly prices, DKK/MWh. "Top 100" is the mean over the 100 most expensive hours of the year in each series separately. "Clustered" is the published series after averaging within the model's own states, i.e. the best the model could do if its economics were exactly right.

		Yearly mean			Median		Top 100		
Area	Year	Model	KF25	Corr.	Model	KF25	Model	Clustered	KF25
DK1	2030	617	588	0.71	675	707	1055	1258	1716
	2035	455	433	0.71	407	231	1199	1490	2129
	2050	545	664	0.64	238	238	2845	2729	3966
DK2	2030	618	585	0.72	671	709	1058	1258	1716
	2035	475	467	0.70	397	232	1247	1644	2555
	2050	640	725	0.64	443	307	2880	3121	4096

The third column of the "Top 100" block is there to say how much of that understatement is the model's fault. Averaging the published prices within the model's own states, which involves no behaviour at all, already brings the top-100 mean in West Denmark in 2035 down from 2,129 to 1,490 DKK/MWh. Of the 930 DKK/MWh the model is short, 639 — about seven tenths — is therefore lost before the economics is consulted. The same split holds in the other rows: the clustering accounts for roughly 70 % of the gap in 2030 and in 2035, and for all of it in 2050, where the model in fact prices the top 100 hours slightly *above* the cluster-averaged series.

Figure 8 shows this as three duration curves. The published series and the cluster-averaged series lie on top of each other for the first 8,000 hours and separate only in the last few hundred, and the model's curve tracks the cluster-averaged one rather than the raw one.

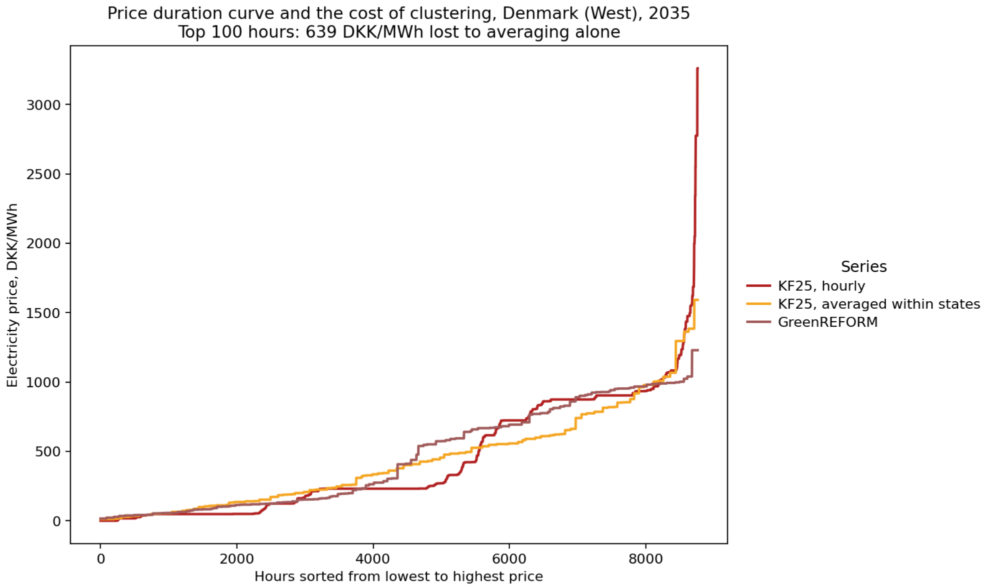
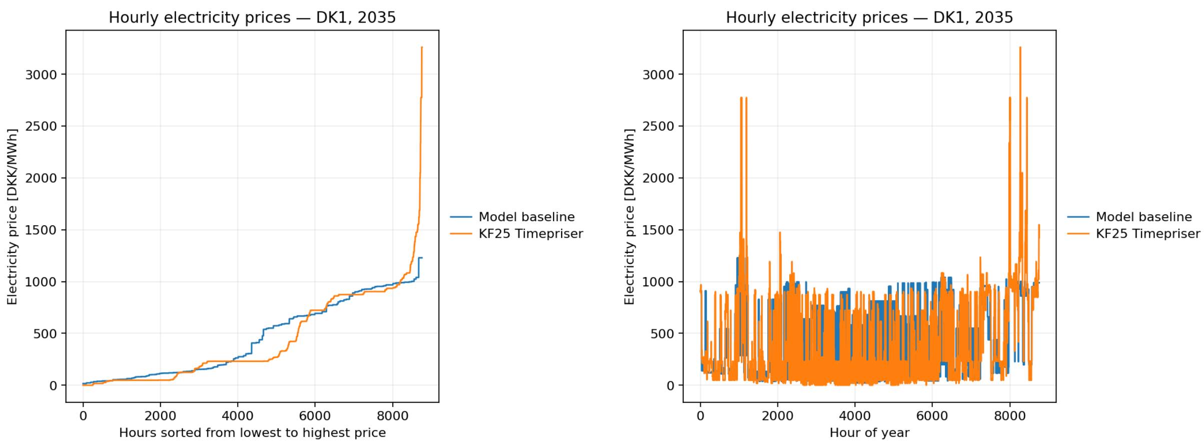


Figure 8: Price duration curve in West Denmark, 2035: published hourly prices, the same prices averaged within the model's 120 states, and the model.

The consequence is worth stating in the form a user needs it. The model's compression of the price tail is close to the best that *any* 120-state model could do on this clustering, so it is not repaired by improving the economics; it is repaired by changing the time representation. Results that depend on the upper tail — scarcity rents, peak capacity values, the profitability of assets that earn their income in a few hours a year — are understated, and the chronological version exists for exactly those questions.

Figure 9 shows the same year sorted by price and in chronological order. The second panel is the more demanding test, and the model tracks the seasonal shape well while missing the short-lived spikes, which is the tail effect again.



(a) Sorted by price
(b) Chronological

Figure 9: Hourly electricity prices in DK1, 2035: model baseline against the KF25 hourly projection.

For district heating there is no directly comparable published hourly price, so Figure 10 simply reports the model's yearly consumer price index. It is flat at around 210–260 DKK/MWh throughout, with the two price zones close together. The stability is a consequence of the structure rather than an assumption: heat demand is largely inflexible, heat cannot be traded between areas, and the marginal heat producer is a boiler or a heat pump whose costs move slowly.

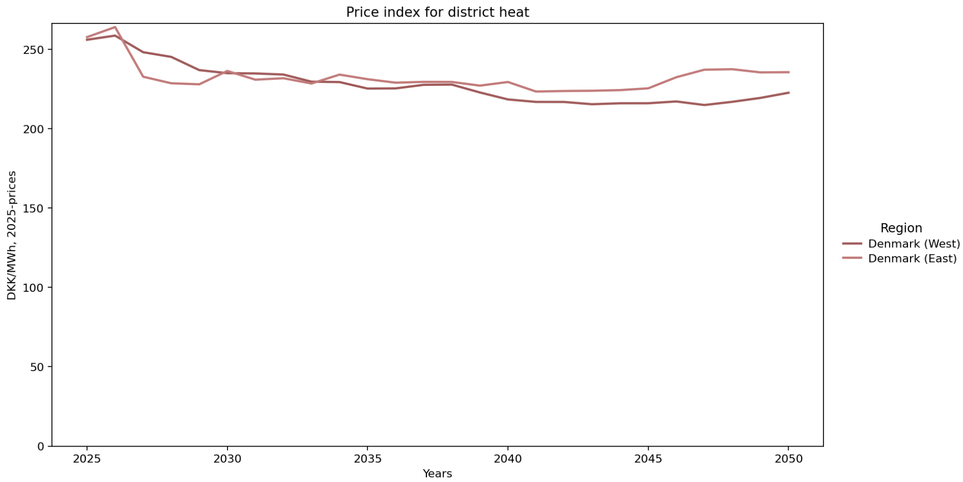


Figure 10: Yearly consumer price index for district heat, model baseline.

10.2.1 Quantities

Prices can come out right for offsetting reasons, so the composition is worth checking separately. Figure 12 compares Danish production by technology with the projection, on both the electricity and the district heat side.

Installed capacity is not the issue. The model is given the projection's capacity paths and reproduces them (Figure 11): 15.1 GW of solar in 2030, 25.1 in 2035 and 43.9 in 2050, matching KF25 to the first decimal, and the same for onshore and offshore wind and for condensation plants. What differs is what the plants do with that capacity.

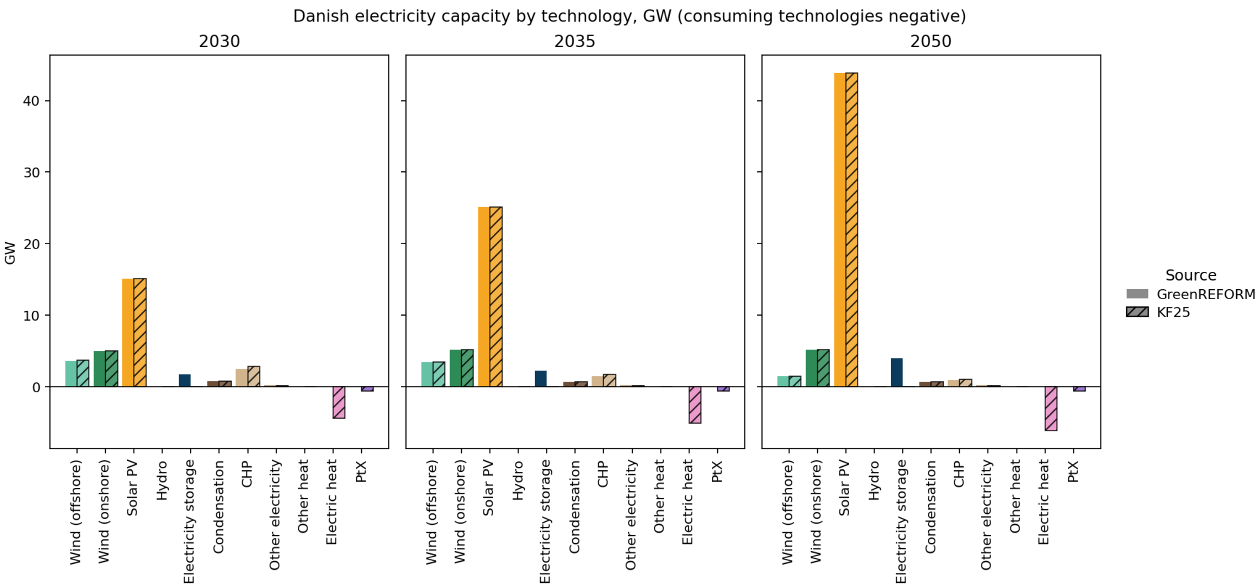
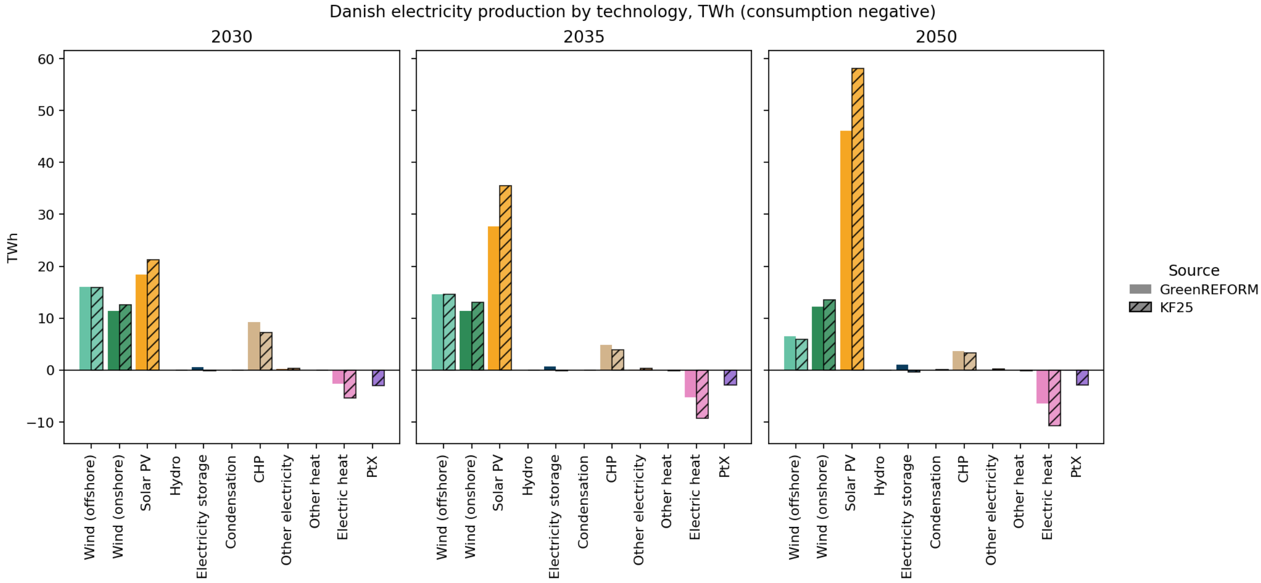
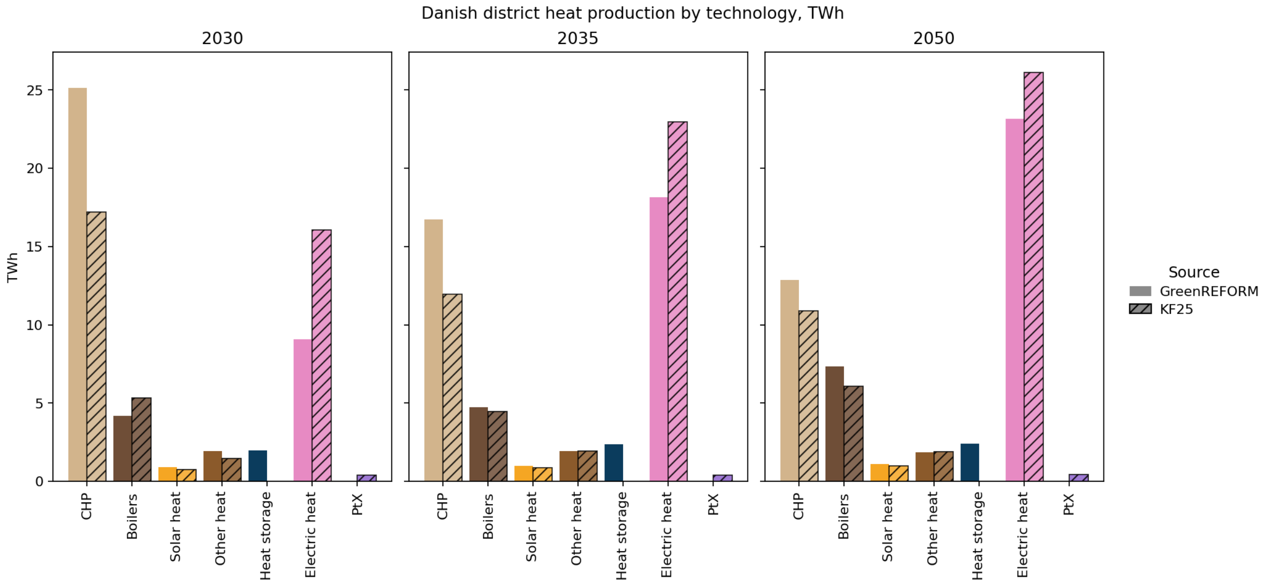


Figure 11: Danish electricity capacity by technology group, model baseline against Climate Outlook 2025, GW. The model's electric heating capacity is held as a consumption capacity and is not picked up by this comparison.



(a) Electricity



(b) District heat

Figure 12: Danish production by technology group, model baseline against Climate Outlook 2025, TWh.

Two discrepancies are large enough to matter, and they turn out to be the same discrepancy.

CHP crowds out electric heat. On the heat side the model produces more heat from CHP than the projection does and correspondingly less from heat pumps and electric boilers, and the two are almost exactly equal and opposite (Table 2). This is the signature of substitution on the heat margin rather than of two independent errors: the same demand is being met, by a different technology. The fuel accounting agrees, with Danish biomass use 8.8 TWh above the projection in 2030 and 5.3 TWh above in 2035 (Figure 13), while waste — whose quantity is fixed by the resource-zone constraint of Section 5 — matches to the decimal.

The leading explanation is the geographical aggregation of the heat side. The 85 Danish district heating areas are collapsed into one per price zone (Section 8.3), which means there is a single heat merit order for all of Jutland and Funen and another for Zealand. A biomass CHP

Table 2: Danish district heat production and fuel use, model against KF25, TWh.

	2030		2035		2050	
	Model	KF25	Model	KF25	Model	KF25
Heat from CHP	25.1	17.1	16.7	12.0	12.9	10.8
Heat from heat pumps and electric boilers	9.1	16.0	18.2	23.0	23.2	26.1
Biomass use	32.4	23.6	17.6	12.3	14.9	12.2

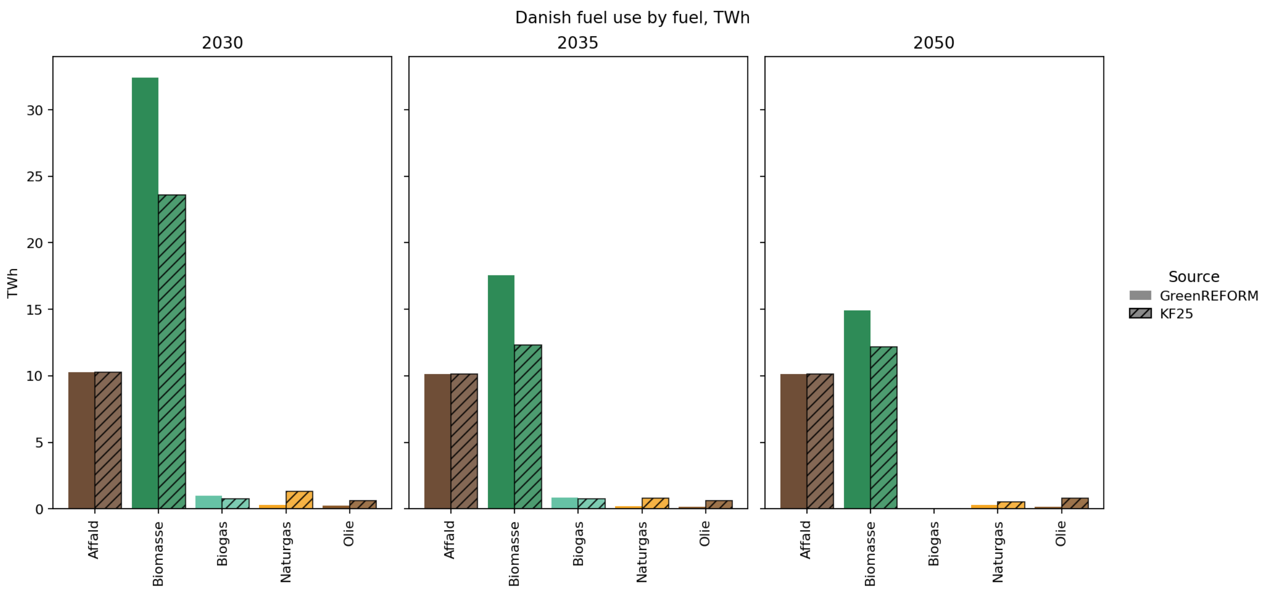


Figure 13: Danish fuel use by fuel, model baseline against Climate Outlook 2025, TWh.

unit in one town can therefore serve heat demand in any other, and the cheapest heat in the zone sets the margin everywhere. In reality each network is supplied locally, and a town whose own plant is a heat pump runs the heat pump whatever a CHP unit two hundred kilometres away would have charged. Consistent with this, the model's CHP fleet runs 4,350–4,650 full-load hours on heat, which is high for a load as seasonal as district heating. This remains a hypothesis: testing it requires solving a version with the heat areas disaggregated, which the current pipeline does not produce, and Section 11.3 records it as a project.

And that shows up in solar. The substitution takes electricity demand with it: the 4.8 TWh of heat the model does not make electrically in 2035 is electricity it does not consume, which is why the electric heat bar in Figure 12a falls short of the projection's by a similar order. On the supply side the model produces 27.7 TWh of solar in 2035 against 35.5 in the projection. Capacity being identical, the gap has to be yield, and it splits into two roughly equal parts in that year: the model's own profiles make 31.7 TWh available where the projection implies 35.5, and of the 31.7 available the model dispatches 27.7, curtailing 4.0 TWh or 12.5 %. By 2050 curtailment dominates — 10.1 TWh, 18 % of available solar, against essentially none in the projection — which is what one would expect of a fleet growing to 44 GW in a country whose peak demand is a fraction of that, but the effect is amplified by the electricity demand that the heat side did not deliver. Wind is curtailed 5–7 % over the same period.

Neither discrepancy is fatal, but both should be known before the model is used. Questions about the electrification of heat will understate how much electric heating runs; questions about the value of solar will understate its output and overstate its curtailment; and questions about biomass demand will overstate it.

10.3 The KF25 baseline

Figure 14 shows the intra-yearly electricity balance for DK1 in 2035, with the 8760 hours sorted by the equilibrium price. Supply is above the axis and demand below it, and the black line is the price on the right-hand axis. The figure is a compact way of reading the dispatch logic of Section 2.1: at the left, prices are low, wind and solar cover most of supply, and the country is exporting and running electric heaters and PtX; moving right, prices rise, renewables recede as a share, thermal plants and imports enter, and the flexible part of demand withdraws. The transition is smooth rather than a merit-order step, which is the effect of the dispatch smoothing $\sigma_{t,i}$.

Figure 15 is the same balance with the traded quantities traced back to the technology that produced them abroad, so that hatched areas indicate energy that crossed a border. It makes visible how much of the Danish balance in the middle of the price distribution is met by imported hydro.

The corresponding district heat balance is in Figure 16. The ordering of technologies with the heat price is the expected one: heat pumps and electric heaters at the bottom, then CHP in back-pressure mode, then boilers at the top. Heat storage is charged in the low-price hours and discharged in the high-price hours, which is the storage block of Section 2.12 doing its work; note that the netting in (43) is what keeps the plant from appearing on both sides at once.

Figure 17 decomposes hourly net exports from DK1 by trading partner. Denmark exports in the low-price hours and imports in the high-price hours, and the switch happens at different points against different partners, reflecting the differences in their own price formation.

Figure 18 shows the effective CO_2 price faced by each category of emissions, that is the combination of quota price and tax that enters $\tau^{CD}_{t,i}$ and $\tau^{BH}_{t,i}$. Two features are worth noting. The Danish rates are well above the EU-ETS price throughout, by about 400 DKK/ton in 2030 rising to 850 DKK/ton in 2050, which is the additional Danish CO_2 tax. And the non-ETS heat rate sits above the non-ETS electricity rate by an amount that grows from 351 DKK/ton in 2030 to 942 DKK/ton in 2050. That gap is exactly the EU-ETS2 price, and it is the double count

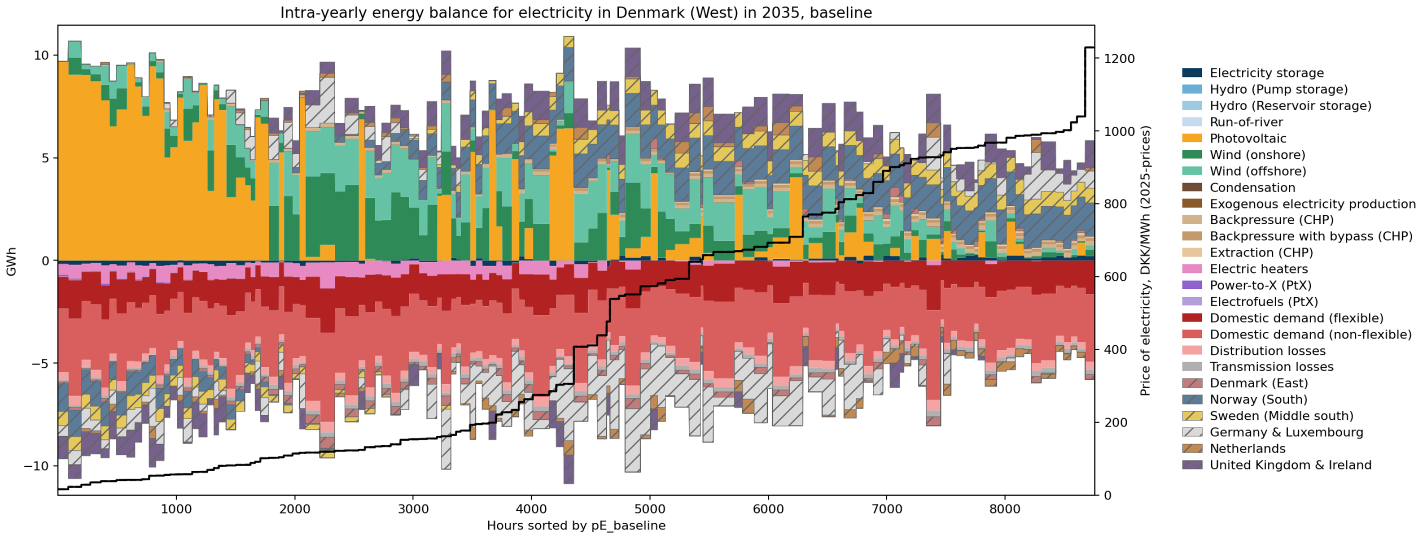


Figure 14: Intra-yearly electricity balance, DK1, 2035, KF25 baseline. Hours sorted by the equilibrium price.

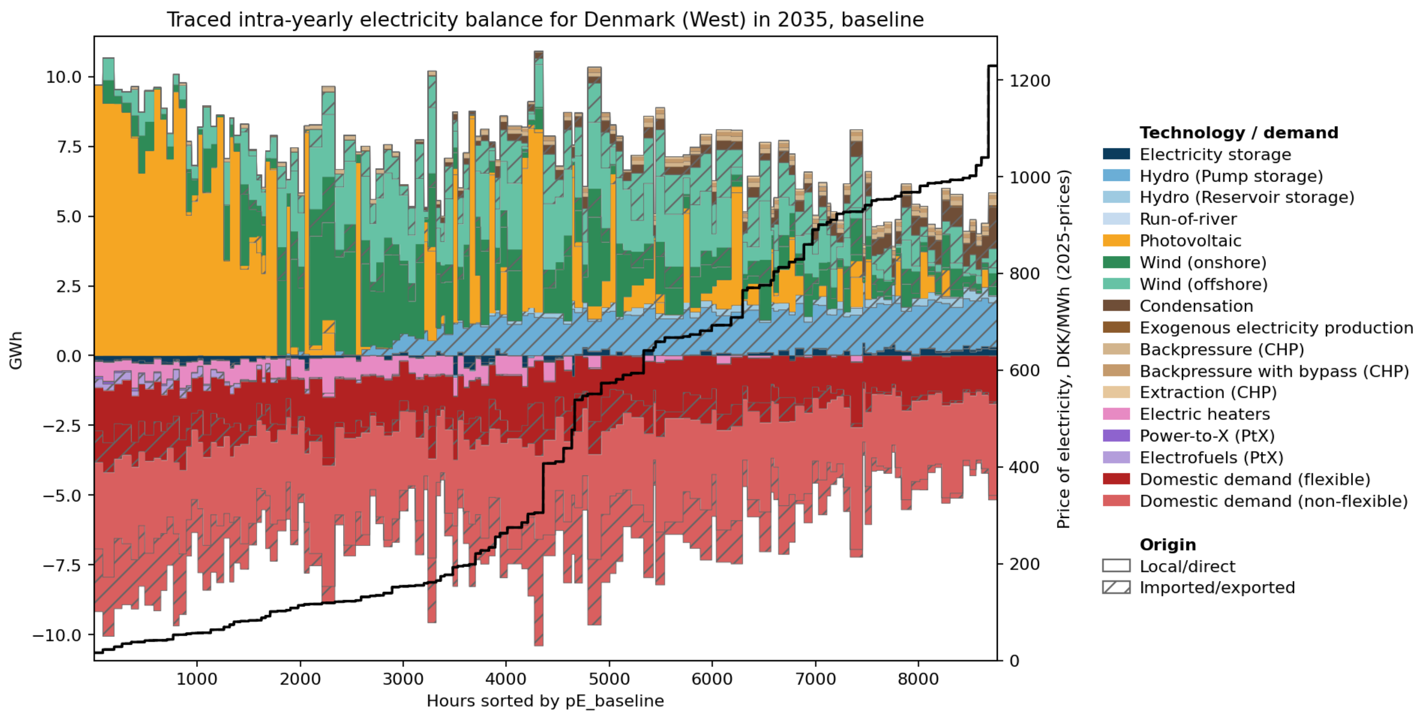


Figure 15: Traced intra-yearly electricity balance, DK1, 2035, KF25 baseline. Hatched: imported or exported.

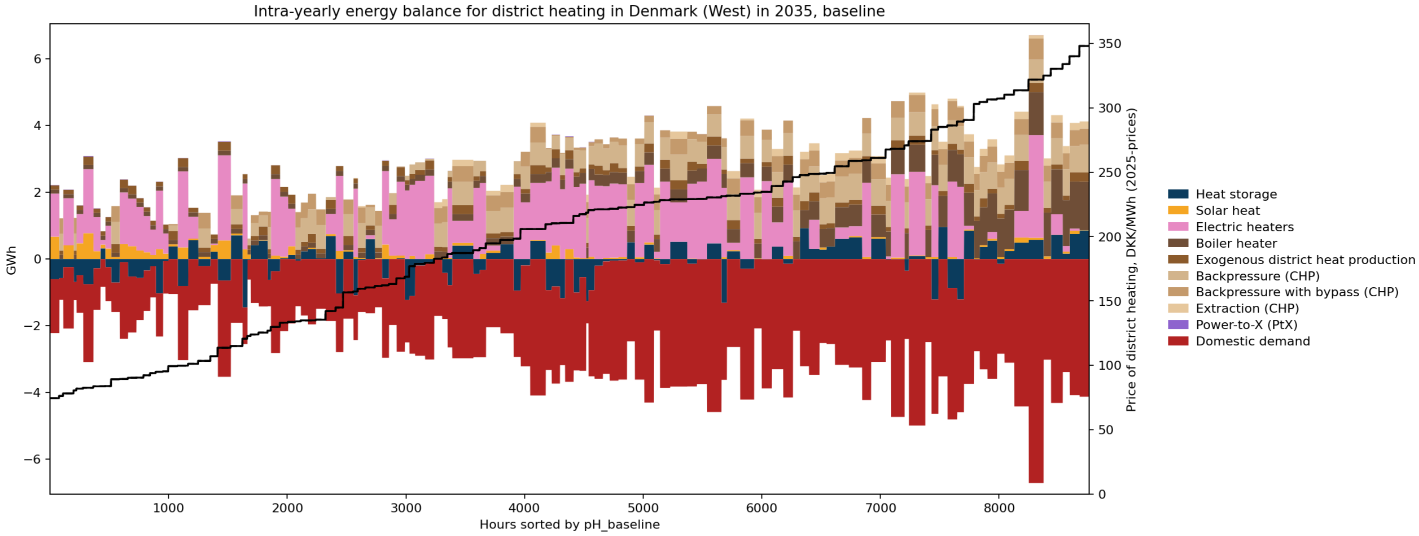


Figure 16: Intra-yearly district heat balance, DK1, 2035, KF25 baseline.

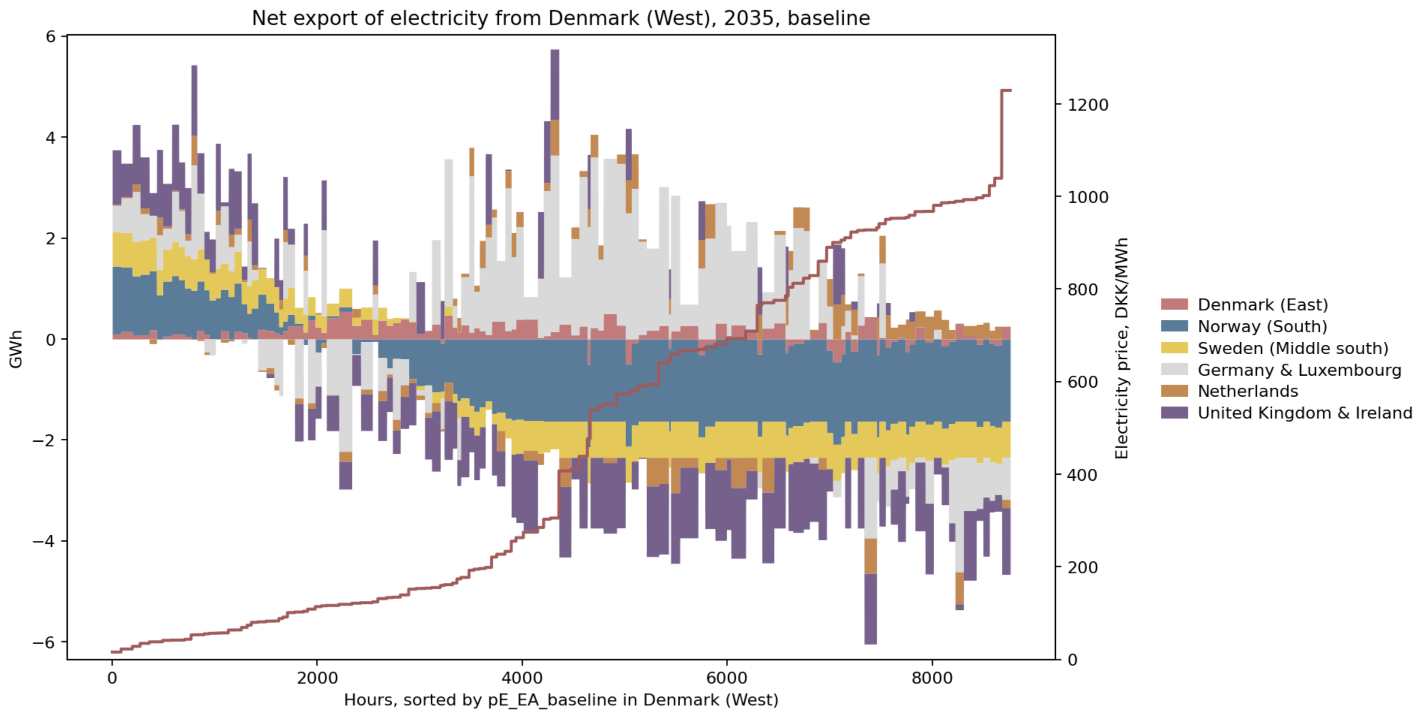


Figure 17: Hourly net exports from DK1 by trading partner, 2035, KF25 baseline.

discussed in the footnote to (11): it is an artefact of the data preparation and the tax macro both adding $\tau_t^{\text{ETS-2}}$, not a feature of Danish tax law. Results that depend on the relative taxation of non-ETS heat and non-ETS electricity should be treated with care until this is resolved.

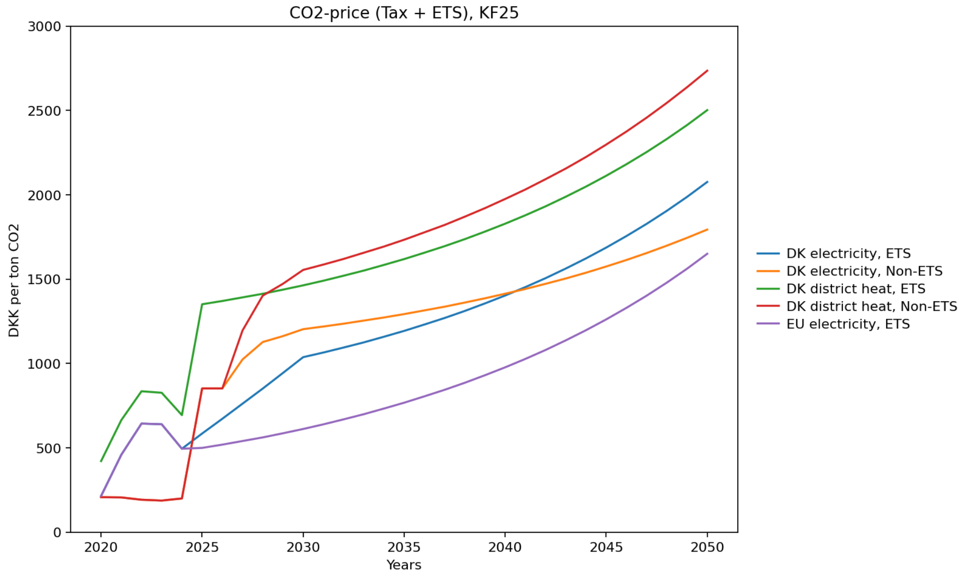


Figure 18: Effective CO_2 price by category, KF25 baseline.

Finally, Table 3 reports the shadow prices on the constrained domestic resources of Section 5. The waste prices are negative, and increasingly so, which is the correct sign: waste is a liability rather than an input, the plant is paid to take it, and the shadow price measures how much. Hazardous waste carries a larger negative price than ordinary waste throughout. Biogas has a positive shadow price that falls to zero by 2050, at which point the constraint stops binding — the available quantity in 2050 exceeds what the plants want at the going fuel price.

Table 3: Shadow prices on constrained domestic resources, $p^r_{t,r,b}$, DKK/MWh.

Resource zone	2025	2030	2035	2040	2050
Waste	−164	−158	−233	−288	−414
Hazardous waste	−186	−318	−369	−426	−560
Biogas	364	367	184	77	0

10.4 Summary

The baseline reproduces the KF25 price projection well at the yearly level through 2035 — within 2–6 % in both Danish price zones — and gets the within-year timing broadly right, with an hourly correlation of about 0.7. Below the level of the yearly average, four things are worth knowing before the model is used.

1. **The price tail is compressed, and mostly not by the economics.** Roughly seven tenths of the model's understatement of the highest prices survives if one simply averages the published prices within the model's own states. It is a property of a 120-state representation rather than a mis-specification, and it is repaired by changing the time representation, not by changing behaviour.
2. **After 2035 the price path flattens where the projection rises**, most visibly in the hydro-dominated areas, where the model settles at roughly half the KF25 level by 2050. This is worth understanding before the model is used for long-horizon analysis.

3. **Heat from CHP displaces heat from electric boilers and heat pumps**, one for one, and Danish biomass use is correspondingly above the projection. The heat-side aggregation is the leading suspect, and the knock-on is that electricity demand from electric heating is understated and solar curtailment overstated.
4. **Short-duration storage is the weakest part of the non-chronological representation.** The stock of a three-hour battery is the state average of a quantity that cycles within the day, which the representation cannot pin down; the solved stock strays a few per cent outside its physical bounds on average and more in individual states.

Separately, the EU-ETS2 double count on non-ETS heat noted above should be resolved before the baseline is used for questions involving the relative taxation of heat and electricity.

11 Projects

11.1 Endogenous investments in generating capacity

In the version documented here, generating capacity follows the projection's path and does not respond to prices. Every result is therefore a short-run result: it says what the existing and already-planned fleet would do at the prices the model solves for, not what would be built if those prices persisted. For a shock of any size this is a strong restriction, and it biases the answer in a known direction, since the price effect that would have been competed away by entry is instead left in the price.

A working extension has been built and used. Vestergaard and Carlsen [2026] add an investment margin for solar, onshore wind and electricity storage, with capacity chosen against the model's own price path and calibrated so that the baseline reproduces the projection's capacities. Applied to a halving of transmission capacity, the extension shows a substantial part of the price effect being absorbed by entry, with storage supplying the largest single share of the capacity that gets built. The work establishes both that the margin can be added without disturbing the baseline and that it changes the answer enough to matter.

It is not part of this release. The implementation is built on a separate vintage of the model, with its own plant aggregation, and consolidating it means reconciling that with the version documented here. The intention is to publish it in a future update, on the KF26 data rather than on KF25, so that the consolidation is done once.

11.2 Bringing the model onto the KF26 data

This version is built on the Climate Outlook 2025 data. Appendix B describes an overlay that applies a number of the major KF26 revisions on top of the KF25 database, which is useful for gauging the direction of the revision but is explicitly not a KF26 version of the model: it operates on the aggregated database and so cannot revise the plant population, the fuel types, the tax rules, the resource-zone assignments or the clustering of hours.

Work on a proper KF26 version, built from the full KF26 Ramses input data through the same pipeline that produced this one, is under way. It will be published as an update, and the extensions above are intended to land on it rather than on KF25.

11.3 Disaggregated district heating areas

Section 10.2.1 identifies the collapse of the 85 Danish district heating areas into one per price zone as the leading explanation for the model producing too much CHP heat, too little heat from electric boilers and heat pumps, and burning too much biomass. The hypothesis is testable: solve a version that keeps the heat areas separate and see whether the gap closes. The aggregation machinery already supports an arbitrary heat-area mapping, so what is needed is a database at that resolution and the patience for a much larger model, since the number of heat market clearing conditions grows by a factor of forty. If the hypothesis holds, the question becomes whether an intermediate aggregation — grouping networks by the technology that supplies them rather than by price zone — recovers most of the accuracy at a fraction of the cost.

11.4 Grid costs and tariffs

Currently, we implement a simple loss in transporting energy using transmission lines, and an additional loss from an exchange node to the consumer. However, grid tariffs and endogenous costs/losses in the distribution grid is a relevant extension to consider.

11.5 Variation in net transfer capacities on transmission lines

Hourly variation in net transfer capacities on transmission lines is currently an essential part of the ability to balance the grid and smooth out prices. In the current approach, there is no variation included, however. We should look into how to model future bottlenecks in the transmission grid and/or the investments required to avoid them.

A larger version of the same question is whether to move from net transfer capacities to a flow-based representation. European and Nordic capacity allocation is moving towards flow-based market coupling, in which the constraints are on the physical flows across critical network elements rather than on bilateral exchange between areas, so that what one border can carry depends on what the others are carrying. The distinction matters for the questions the model is most often asked: with net transfer capacities each line is constrained independently, which will overstate simultaneous flows in the hours when the grid is tight — the very hours in which the value of transmission is decided.

Adopting it is not a modelling problem so much as a data problem. A flow-based representation needs power transfer distribution factors and remaining available margins for each critical network element, which are not in the Ramses input data the model is built from and are not published on a horizon that reaches 2050. Until a projected flow-based domain exists, the honest position is that the model gives an upper bound on how much trade can relieve a tight hour, and that shocks to transmission capacity should be read as shocks to that upper bound.

11.6 A design-days approach with uncertainty

Create specific short run, medium run, and long run variation as follows:

- Create 4 seasons ($n_s = 4$) and $n_{w,s}$ "week" types for each season. Each week type consists of $n_{d,w,s}$ days of n_{id} intraday blocks of states (e.g. 4 blocks representing night, morning, midday, evening).
- Seasons, days, and intraday blocks are chronologically ordered.
- Weeks are chosen to illustrate archetypes – catching different important extremes (and a few average states as well). The transition between weeks, however, is probabilistic with known transition probabilities.

11.7 Estimating ramp-up constraints from the chronological model

We could run a dynamic version of the chronological model with ramp-up constraints, and use hourly capacity factors from that version to approximate the ramp-up constraints in a simple model without them.

11.8 Using the full hourly model to inform non-chronological aggregation

We could implement a version where the endogenous price variation from the hourly model informed the aggregation of hours. The auxiliary series used in the clustering (section 8.5) are a first step in this direction, but they are computed from exogenous data rather than from model output.

11.9 Demand flexibility with dynamic optimization

We could consider implementing a version of demand flexibility where energy can be reallocated in the short run, but less so in the long run, with the cost of reallocation increasing in the duration of time that the energy is postponed.

References

Alcacer, Aleix (2025). *Archetypes*. Version 0.11. Python package.

Andersen, Per and Alexander M. Dietrich (2026). "Price response in residential electricity demand: Evidence from Danish smart meter data". *Energy Economics* 153, p. 109087.

Berg, Rasmus K. and Janek Eskildsen (2019). "Modelling the energy sector in a computable general equilibrium framework: A new approach to integrated bottom-up and top-down modelling".

Berg, Rasmus K. and Janek Eskildsen (2021a). "Modelling the Energy Sector in a Computable General Equilibrium Framework".

Berg, Rasmus K.S. and Janek Eskildsen (2021b). "The Value and Potential of Energy Storage".

Pedregosa, F. et al. (2011). "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research* 12, pp. 2825–2830.

Vestergaard, Elin and Marcus Emil Carlsen (2026). "Flexibility and price formation in the electricity market: implications for solar and wind investments". Master's thesis, University of Copenhagen.

A Dictionary: from LaTeX to GAMS

The tables below map the notation used in this document to the symbol names used in the GAMS implementation. Domains are given with the GAMS index names; `t` is years, `h` hours or clustered states, `id` producers, `cIdx` consumers, `EA`/`HA` electricity/heat areas, `Ft`/ `BFt` fuel types/base fuel types, `hvt` variation types, `line` transmission lines, `rZone` resource zones, `St` subsidy types, and `Pt` plant types.

A.1 Sets and mappings

Math	GAMS	Description
t	`t`	Years, 2020–2050
h, k	`h`, `hh`	Hours (8760) or clustered short-run states
i	`id`	Producer ids
c	`cIdx`	Consumer ids
g_E	`EA`, `EAEA`	Electricity areas (price zones)
g_{DH}	`HA`	District heat areas
b	`BFt`	Base fuel types
f	`Ft`	Fuel types (linear combinations of base fuels)

Math	GAMS	Description
v	hvt	Hourly variation types
l	line	Transmission lines
r	rZone	Resource zones
s	prod_sector	Production sector (Waste / NonWaste)
—	Pt	Plant types (CD, BP, BB, EX, BH, ...)
—	St	Subsidy types
—	Dt	Duration types for storage plants
$g(i)$	id2EA, id2HA	Area of producer i
$f(i)$	id2Ft	Fuel type of producer i
$c(i)$	id2cIdxE, id2cIdxH	Consumer index of producer i
$r(i)$	id2rZone	Resource zone of producer i
$\mathcal{R}(i)$	id2TaxRule	CHP fuel allocation rule of producer i
$b(r)$	rZone2BFt	Base fuel associated with resource zone r
$\mathcal{B}_1$	bidT1	Producers on fixed settlement contracts
—	DKId, DK_EA	Danish producers / Danish electricity areas
—	activeId, activeEA	Active (t, i) / (t, g_E) combinations
—	activeCE, activeCH	Active consumers of electricity / heat
—	activeLine, activeLineEA	Active lines / directed line-area pairs
—	EA_dem	Areas with exogenous yearly consumer demand
—	cIdx_DCIIA	Consumers described by the DC-IIA demand system
—	VREId, storageId, chargeId	Variable renewable / storage / charging producers
—	CHPId, CHPIdFlex, CHHyId	CHP producers / flexible CHP (BB, EX) / CHP with hydrogen
—	UPId	Producers for which unit profits are computed

A.2 Exogenous parameters

Math	GAMS	Domain	Description / unit
$Q^E_{t,i}$	ECap	t,id	Electricity generating capacity, GW
$Q^{DH}_{t,i}$	HCap	t,id	Heat generating capacity, GW
$Q^H_{t,i}$	HyCap	t,id	Hydrogen generating capacity, GW
$Q^E_{t,c}$	EDCap	t,cIdx	Demand capacity for electricity, GW
$CapF_{t,i}$	capFCap	t,id	Capacity factor
$CapF^I_{t,i}$	capFInf	t,id	Capacity factor of inflow to reservoirs
$CapF^X_{t,l}$	capFExp, capFImp	t,line	Capacity factor on line, export/import direction
$\gamma_{h,v}$	hvar	h,hvt	Hourly-to-average ratio, mean one
$\mu^{cap}_{t,i,v}$	uHvtCap	t,id,hvt	Weights on base variation profiles, capacity
$\mu^{inf}_{t,i,v}$	uHvtInf	t,id,hvt	Weights on base variation profiles, inflow
$\mu_{h,c}$	uHvtDemE, uHvtDemH	t,cIdx,hvt	Weights on base variation profiles, demand

Math	GAMS	Domain	Description / unit
—	uHvtExp, uHvtImp	t,line,hvt	Weights on base variation profiles, lines
n_h	hWeight	h	Number of full hours represented by state h
$\eta_{t,i}$	Eff	t,id	Conversion efficiency (definition varies by plant type)
$\kappa_{t,i}$	Cm	t,id	Electricity-to-heat ratio (hydrogen-to-heat for PX)
$\xi_{t,i}$	Cv	t,id	Electricity loss from a marginal increase in heat
$v_{t,i}$	VOM	t,id	Variable O&M costs, DKK/MWh
$c_{t,c(i)}$	VOM_charge	t,id	Variable O&M costs of charging, DKK/MWh
$FOM_{t,i}$	FOM	t,id	Fixed O&M costs, million DKK/GW/year
$\zeta_{t,i}$	duration	t,id	Storage capacity over dispatch capacity, hours
$\chi_{t,i}$	chargeSpeed	t,id	Charge-to-discharge capacity ratio
$x^{out}_{t,i}$	parMCOut	t,id	Cost of exceeding storage bounds, DKK/MWh
$x^{in}_{t,i}$	parMCIn	t,id	Cost of adjusting storage within bounds, DKK/MWh
$x^{\Delta}_{t,i}$	parMCSmooth	t,id	Smoothing between the two storage cost regimes
$\sigma_{t,i}$	sigmaID	t,id	Dispatch smoothing parameter, DKK/MWh
$\sigma^{EH}_{t,i}$	sigmaID_EH	t,id	Mode-choice smoothing parameter for BB/EX
$\sigma_{t,l}$	sigmaLine	t,line	Line utilisation smoothing parameter
$\sigma_{t,c}$	sigmaCE, sigmaCH	t,cIdx	Hourly demand elasticity
$\sigma^{E,t}_{t,c}$	sigmaCEt, sigmaCHt	t,cIdx	Yearly demand elasticity
$\alpha_{f,b}$	fm	Ft,BFt	Fuelmix coefficients
$p_{t,b}$	fp_t	t,BFt	Base fuel prices, DKK/MWh
$\tau^{DH}_{t,b}$	heatTax	t,BFt	Fuel tax on inputs to heat production, DKK/MWh
$\mu^{CO_2}_{t,b}$	CO2prop	t,BFt	CO_2 intensity of base fuel, ton/MWh
$\mu^{SO_2}_{t,b}$	SO2prop	t,BFt	SO_2 intensity of base fuel, kg/MWh
$\tilde{\mu}^{CO_2}_{t,i}$	uCO2e_id	t,id	Plant CO_2e intensity beyond fuelmix, ton/MWh
$\mu^{NO_x}_{t,i}$	uNOx	t,id	NO_x intensity, kg/MWh input
$DS_{t,i}$	Desulp	t,id	Degree of desulphurization
$CO2_{t,i}$	CO2Cap	t,id	Share of emissions covered by EU-ETS
$\tau^{\text{ETS-1}}_{t}$	taxCO2_ETS	t	EU-ETS price, DKK/ton
$\tau^{\text{ETS-2}}_{t}$	taxCO2_ETS2	t	EU-ETS2 price, DKK/ton
$\tau^{CO_2,E}_{t}$	taxCO2_E	t	Additional CO_2 tax, ETS electricity, DKK/ton

Math	GAMS	Domain	Description / unit
$\tau_t^{\text{Non-ETS},E}$	taxCO2_xETS_E	t	CO_2 tax, non-ETS electricity, DKK/ton
$\tau_t^{CO_2,DH}$	taxCO2_H	t	Additional CO_2 tax, ETS heat, DKK/ton
$\tau_t^{\text{Non-ETS},DH}$	taxCO2_xETS_H	t	CO_2 tax, non-ETS heat, DKK/ton
$\tau_t^{SO_2}$	taxSO2	t	SO_2 tax, DKK/kg
$\tau_t^{NO_x}$	taxNOx	t	NO_x tax, DKK/kg
$\tilde{T}_{t,i}^{E}$	subsidy	t,St	Subsidy per MWh energy output (enters with $-$)
$\overline{\nu}$	taxVRule	scalar	Heat-efficiency rule, 1.25
$\overline{\varepsilon}$	taxERule	scalar	Electricity-efficiency rule, 0.65
$\overline{\gamma}$	taxGuardRule	scalar	Guard on the V-rule, 0.35
$Q_{t,l,g\to g'}$	trCap	t,line,EA,EAEA	Directed transmission capacity, TW
$\eta_{t,l}$	trEff	t,line	Transmission line efficiency
$\delta_{t,g}$	distrEff	t,EA	Distribution efficiency
p_t^H	pHy	t	World market price of hydrogen, DKK/MWh
$\underline{p}^E, \underline{p}^{DH}$	pE_min, pH_min	scalar	Price floors, DKK/MWh
$\overline{D}_{t,c}^{E,\text{data}}$	DEt, DHt	t,cIdx	Exogenous yearly consumer demand, TWh
$\overline{p}_{t,c}^{E}$	pEt_C_cIdx_exo	t,cIdx	Reference consumer price index, electricity
$\overline{p}_{t,c}^{DH}$	pHt_C_cIdx_exo	t,cIdx	Reference consumer price index, heat
$\overline{R}_{t,r}$	rDemand	t,rZone	Available quantity of the resource, TWh
$\Pr^+, \Pr^-$	markovPr_plus, markovPr_minus	h,hh	Forward/backward transition probabilities

A.3 Endogenous variables

Math	GAMS	Domain	Description / unit
$c_{t,i}$	MC	t,id	Marginal cost, DKK/MWh (or per composite output)
$c_{t,i}^{BP}$	MC_BP	t,id	Marginal cost in back-pressure mode
$c_{t,i}^{BB}$	MC_BB	t,id	Marginal cost in bypass mode
$c_{t,i}^{EX}$	MC_EX	t,id	Marginal cost in extraction mode
$c_{t,c(i)}$	MC_charge	t,id	Marginal cost of charging
$c_{t,f}^{fuel}$	fuelCost	t,Ft	Fuel price index, DKK/MWh input
$c_{t,i}^{r}$	fuelCost_rZone	t,id	Resource shadow price per MWh input
—	fuelTaxH	t,Ft	Heat fuel tax index, DKK/MWh input
$\mu_{t,f}^{CO_2}$	uFuelCO2	t,Ft	CO_2 intensity of fuel type, ton/MWh
$\mu_{t,f}^{SO_2}$	uFuelSO2	t,Ft	SO_2 intensity of fuel type, kg/MWh
$s_{t,h,i}^{E}$	dUtil	t,h,id	Dispatch rate per installed GW
$d_{t,h,c(i)}^{E}$	dUtilC	t,h,id	Charge rate per installed GW

Math	GAMS	Domain	Description / unit
$u^{BP}_{t,h,i}$	`uCHP_BP`	`t,h,id`	Share of flexible CHP plant in back-pressure mode
$\Delta^{q}_{t,h,i}$	`pEH`	`t,h,id`	Short run profit margin per composite output
$\lambda_{t,h,i}$	`contVal`	`t,h,id`	Continuation value of stored energy, DKK/MWh
$r_{t,h,i}$	`stored`	`t,h,id`	Stored energy relative to installed GW
$\Pi_{t,i}$	`UP`	`t,id`	Unit profits, million DKK/GW
$S^{E}_{t,h,i}$	`SEh`	`t,h,id`	Electricity supply, GWh
$S^{DH}_{t,h,i}$	`SHh`	`t,h,id`	District heat supply, GWh
$S^{H}_{t,h,i}$	`SHyh`	`t,h,id`	Hydrogen supply, GWh
$D^{E}_{t,h,c}$	`DEh`	`t,h,cIdx`	Hourly electricity demand, GWh
$D^{DH}_{t,h,c}$	`DHh`	`t,h,cIdx`	Hourly district heat demand, GWh
—	`DEscale`, `DHscale`	`t,cIdx`	Denominator of the DC-IIA share function
$D^{E}_{t,c}$	`DEt_endo`	`t,cIdx`	Endogenous yearly electricity demand, TWh
$D^{DH}_{t,c}$	`DHt_endo`	`t,cIdx`	Endogenous yearly heat demand, TWh
$u^{E}_{t,c}$, $u^{DH}_{t,c}$	`uEt`, `uHt`	`t,cIdx`	Level parameters in yearly demand
$J^{E}_{t,c}$, $J^{DH}_{t,c}$	`J_DEt`, `J_DHt`	`t,cIdx`	Residuals in the yearly demand calibration
$p^{E}_{t,h,g}$	`pE`	`t,h,EA`	Hourly electricity price, DKK/MWh
$p^{DH}_{t,h,g}$	`pH`	`t,h,HA`	Hourly district heat price, DKK/MWh
$J^{E}_{t,h,g}$, $J^{DH}_{t,h,g}$	`jE`, `jH`	`t,h,EA/HA`	Residuals in market clearing
$p^{r}_{t,r,b}$	`p_rZone`	`t,rZone,BFt`	Resource shadow price, DKK/MWh
$J^{r}_{t,r,b}$	`j_rZone`	`t,rZone,BFt`	Residual in resource market clearing
$X_{t,h,l,g\to g'}$	`lineExport`	`t,h,line,EA,EAEA`	Gross exports through a line, TWh
$X^{E}_{t,h,g}$	`qExportsE`	`t,h,EA`	Net exports of an area, TWh
$M^{E}_{t,h,g}$	`qImportsE`	`t,h,EA`	Net imports of an area, TWh
—	`qExportsE_t`, `qImportsE_t`	`t,EA`	Yearly exports/imports, TWh
$S^{E}_{t,h,g}$	`qProdE`	`t,h,EA`	Electricity production by area, TWh
$S^{DH}_{t,h,g}$	`qProdH`	`t,h,HA`	District heat production by area, TWh
$S^{H}_{t,h,g}$	`qProdHy`	`t,h,EA`	Hydrogen production by area, TWh
$D^{E,\text{cons}}_{t,h,g}$	`qDemandE_DCIIA`	`t,h,EA`	Consumer electricity demand, TWh
$D^{E,\text{plants}}_{t,h,g}$	`qDemandE_plants`	`t,h,EA`	Plant electricity demand, TWh
L^{E}_{t}	`qDemandE_loss_t_DK`	`t`	Distribution losses in Denmark, TWh
$F_{t,g,b}$	`qFuel_t`	`t,EA,BFt`	Fuel use by area and base fuel, TWh
$F_{t,r,b}$	`qFuel_t_DK_rZone`	`t,rZone,BFt`	Fuel use by resource zone, TWh
$F^{E}_{t,s,b}$	`qFuelUseE_t_DK_prod_sector`	`t,prod_sector,BFt`	Fuel attributed to electricity, TWh
$F^{H}_{t,s,b}$	`qFuelUseH_t_DK_prod_sector`	`t,prod_sector,BFt`	Fuel attributed to heat, TWh
$p^{E,C}_{t,c}$	`pEt_C_cIdx`	`t,cIdx`	Consumer price index by segment, DKK/MWh

Math	GAMS	Domain	Description / unit
$p_{t,g}^{E,C}$	`pEt_C`	`t,EA`	Consumer price index by area, DKK/MWh
$p_{t,g}^{E,S}$	`pEt_S`	`t,EA`	Producer price index by area, DKK/MWh
$p_{t,DK}^{E,C}$	`pEt_C_DK`	`t`	Danish consumer price index, DKK/MWh
$p_{t,DK,s}^{E,S}$	`pEt_S_DK_prod_sector`	`t,prod_sector`	Producer price index by sector, DKK/MWh
$p_{t,DK}^{E,X}$	`pEt_Export_DK`	`t`	Average export price, DKK/MWh
$p_{t,DK}^{E,M}$	`pEt_Import_DK`	`t`	Average import price, DKK/MWh

A.4 Macros and functions in `functions.gms`

GAMS macro	Corresponds to
`F_normPdf(x)`	$\phi(x)$, standard normal density
`errorf(x)`	$\Phi(x)$, standard normal CDF (GAMS built-in)
`F_HvarCap`	$\gamma_{h,v(i)}$ for capacity, $\sum_v \mu_{t,i,v}^{cap}\gamma_{h,v}$
`F_HvarInf`	$\gamma_{h,v(i)}^{I}$ for reservoir inflow
`F_HvarDemE(h)`, `F_HvarDemH(h)`	$\mu_{h,c}$, the demand variation profile
`F_uCO2`, `F_uSO2`	$\mu_{t,i}^{CO_2}$, $\mu_{t,i}^{SO_2}$
`F_taxCO2_E`, `F_taxCO2_H`	$\mu_{t,i}^{CO_2}\tau_{t,i}^{CD}$ and $\mu_{t,i}^{CO_2}\tau_{t,i}^{BH}$
`F_EmTax_E`, `F_EmTax_H`	$T_{t,i}^{M}$ for electricity / heat producers
`F_EmTax_CHP`	$T_{t,i}^{M,CHP}$, excluding CO_2
`F_EmTax_For`	$T_{t,i}^{M}$ for foreign plants (EU-ETS only)
`F_ETax`, `F_HTax`	$\tilde{T}_{t,i}^{E}$ and $T_{t,i}^{DH}$
`F_FuelE_BP`, `F_FuelH_BP`	$a_{t,i}^{E,BP}$, $a_{t,i}^{H,BP}$
`F_FuelH_BB`	$a_{t,i}^{H,BB}$
`F_FuelE_EX`, `F_FuelH_EX`	$a_{t,i}^{E,EX}$, $a_{t,i}^{H,EX}$
`F_ETax_BP`, `F_HTax_BP`	$T_{t,i}^{E,BP}$, $T_{t,i}^{DH,BP}$
`F_HTax_BB`, `F_ETax_EX`	$T_{t,i}^{DH,BB}$, $T_{t,i}^{E,EX}$
`F_RatioE`, `F_RatioH`	$(p-c)/\sigma$ for electricity / heat producers
`F_RatioBT1`	$-c/\sigma$, the type-1 (fixed contract) counterpart
`F_RatioEH`	$\Delta_{t,h,i}^{q}/\sigma_{t,i}$
`F_pEH(dE,dH,MC)`	$d_E p^E + d_H p^{DH} - c$, the composite margin
`F_dUtilCHP`	$\Phi(\Delta^q/\sigma)\, CapF\, \gamma$
`F_Delta_BP_BB`, `F_Delta_BP_EX`	$\Delta_{t,h,i}$, mode-choice margin
`F_CW(x)`	$\sigma\phi(x) CapF\, \gamma$, the cost wedge in unit profits
`F_CW_EH(x)`	$\sigma\phi(x)$, cost wedge for EH/EF/PX
`F_RatioE_CHR`, `F_RatioCE_CHR`	Dispatch and charge ratios including $\lambda_{t,h,i}$
`F_contVal_CHR`	$\Lambda(r)$, cf. (39)
`F_NetDischarge`, `F_NetCharge`	$\widetilde{s}_{t,h,i}$, $\widetilde{d}_{t,h,i}$, cf. (43)
`F_COP`, `F_COPh`	$\eta_{t,h,i}$, the hourly coefficient of performance

GAMS macro	Corresponds to
`F_RatioExport`, `F_lineExport`	$x_{t,h,l,g\to g'}$ and $X_{t,h,l,g\to g'}$
`F_NXLine_h`	$NX_{t,h,l,g\to g'}$
`F_ExportLine_h`, `F_ImportLine_h`	X^{net}, M^{net} after netting
`F_SECTOR_COND`	Selects plants in the Waste / NonWaste sector
`F_FM`	$\alpha_{f(i),b}$ for the plant's own fuel type

A.5 Module states

Module	State	Meaning
`exoInputs`	`chronological`	Hours ordered in time; no Markov matrices
	`clusterAvg`	Clustered states; Markov matrices declared
`HY`, `HYS`, `ES`, `HS`	`chronological`	$\lambda_{t,h}$ and $r_{t,h}$ use $h \pm 1$
	`clusterAvg`	$\lambda_{t,h}$ and $r_{t,h}$ use the Markov matrices
`equilibrium`	`exo_prices`	Prices fixed, J-terms endogenous
	`endo_prices`	p^E, p^{DH} endogenous; p^r fixed, J^r endogenous
	`endo_prices_rZone`	All prices endogenous; all J-terms fixed
`demand_DCIIA`	`base`	$u^E_{t,c}$ and $J^E_{t,c}$ both exogenous
	`DCIIA_endo_J`	$J^E_{t,c}$ endogenous (residual measurement)
	`DCIIA_endo_u`	$u^E_{t,c}$ endogenous (level calibration)
`aggregates`	`base`	Hourly aggregates only
	`aggregates_yearly`	Adds yearly production, demand and fuel aggregates
`transmission`	`base`	Hourly line flows only
	`transmission_yearly`	Adds yearly exports and imports
`priceIndex`	`base`	Segment-level consumer price indices only
	`priceIndex_regional`	Adds all regional, sectoral and trade price indices

B The KF26 overlay

Provisional. The distribution contains an ad-hoc implementation of a number of the major changes in Climate Outlook 2026 (KF26), applied on top of the KF25 database. It is activated by setting `kf = 'KF26'` in the baseline notebook and implemented in a single Python routine; solved databases for it are distributed alongside the KF25 ones. It is *not* a KF26 version of the model, and nothing in Section 10 uses it. A proper KF26 version will be built from the full KF26 Ramses data (Section 11.2), and numbers produced with this overlay should be read as an indication of the direction and rough magnitude of the KF26 revision only, not as a KF26 projection.

The overlay adjusts the KF25 database along a small number of dimensions before the model is solved, broadly covering revisions to demand levels, to capacity paths, and to tax rates. Because it operates on the aggregated database rather than on the raw data, it cannot revise anything that is fixed by the aggregation – the set of plants, their fuel types, their tax rules, their resource zones, or the mapping of hours to states. Any KF26 change that would alter those dimensions is therefore not captured.

Practical consequences worth keeping in mind when reading results produced with the overlay:

- The plant population, and hence the composition of the supply curve, is the KF25 one.

- The clustering of hours is the KF25 one, so the correlation between demand and variable renewable supply reflects KF25 capacities.
- The demand revisions do not reach the DC-IIA consumer segments of Section 3.1 in the aggregated database. Consumer electricity demand is therefore identical in the two runs to the last decimal, and any difference in total demand comes from the plants' own consumption — electric heaters, PtX and storage responding to the changed price path. This is the clearest single illustration of why the overlay is provisional: it must not be read as a statement about how KF26 revises Danish electricity consumption.
- The calibration of the resource-zone shadow prices is redone as part of the solution procedure, so those prices do respond to the overlay.
- The starting price vector is stored separately for the KF25 and KF26 configurations, so the two are warm-started from their own previous solutions.