muse-documentation

Release 0.1

Sustainable Gas Institute

Nov 05, 2020

CONTENTS:

CHAPTER

ONE

INSTALLATION

There are two ways to install MUSE: one for users who do not wish to modify the source code of MUSE, and another for developers who do.

Note: Windows users and developers may need to install Windows Build Tools. These tools include C/C++ compilers which are needed to build some python dependencies.

MacOS includes compilers by default, hence no action is needed for Mac users.

Linux users may need to install a C compiler, whether GNU gcc or Clang, as well python development packages, depending on their distribution.

1.1 For users

MUSE is developed using python, an open-source programming language, which means that there are two steps to the installation process. First, python should be installed. Then so should MUSE.

The simplest method to install python is by downloading the Anaconda distribution. Make sure to choose the appropriate operating system (e.g. windows), python version 3.7, and the 64 bit installer. Once this has been done follow the steps for the anaconda installer, as prompted.

After python is installed we can install MUSE. MUSE can be installed via the Anaconda Prompt (or any terminal on Mac and Linux). This is a command-line interface to python and the python eco-system. In the anaconda prompt, run:

```
python -m pip install --user git+https://github.com/SGIModel/StarMuse
```

It should now be possible to run muse. Again, this can be done in the anaconda prompt as follows:

```
python -m muse --help
```

Note: Although not strictly necessary, users are encouraged to create an Anaconda virtual environment and install MUSE there, as shown in *For developers*.

1.2 For developers

Although not strictly necessary, creating an Anaconda virtual environment is highly recommended. Anaconda will isolate users and developers from changes occuring on their operating system, and from conflicts between python packages. It also ensures reproducibility from day to day.

Create a virtual env including python with:

```
conda create -n muse python=3.7
```

Activate the environment with:

```
conda activate muse
```

Later, to recover the system-wide "normal" python, deactivate the environment with:

```
conda deactivate
```

The simplest approach is to first download the muse code with git:

```
git clone https://github.com/SGIModel/StarMuse.git muse
```

For interested users, there are plenty of good tutorials for git. Next, it is possible to install the working directory into the conda environment:

```
# On Linux and Mac
cd muse
conda activate muse
python -m pip install -e ".[dev,docs]"

# On Windows
dir muse
conda activate muse
python -m pip install -e ".[dev,docs]"
```

The quotation marks are needed on some systems or shells, and do not hurt on any. The downloaded code can then be modified. The changes will be automatically reflected in the conda environment.

Tests can be run with the command pytest, from the testing framework of the same name.

The documentation can be built with:

```
python setup.py docs
```

The main page for the documentation can then be found at *build\sphinx\html\index.html* (or *build/sphinx/html/index.html* on Mac and Linux). The file can viewed from any web browser.

CHAPTER

TWO

RUNNING YOUR FIRST EXAMPLE

In this section we run an example simulation of MUSE and visualise the results. There are a number of different examples in the source code, which can be found INSERT LINK HERE.

Once python and MUSE have been installed, we can run an example. To do this open anaconda prompt. Then change directory to where you have downloaded the MUSE source code.

Navigate to the following link for MacOS or Linux based operating systems:

`{MUSE_download_location}/StarMuse/run/example/default/`

Change `{MUSE_download_location}` to the location you downloaded MUSE to, for example `Users/{my_name}/Documents/` using the `cd` command, or "change directory" command. Once we have navigated to the directory containing the example settings `settings.toml` we can run the simulation using the following command in the anaconda prompt or terminal:

```
python -m muse settings.toml
```

If running correctly, your prompt should output text similar to that which can be found here.

It is also possible to run MUSE directly in python using the following code:

```
[ ]: from muse import examples
     model = examples.model("default")
     model.run()
```

2.1 Results

If the default MUSE example has run successfully, you should now have a folder called `Results` in the same directory as `settings.toml`.

This directory should contain results for each sector (`Gas`,`Power` and `Residential`) as well as results for the entire simulation in the form of `MCACapacity.csv` and `MCAPrices.csv`.

- `MCACapacity.csv` contains information about the capacity each agent has for each technology per year.
- `MCAPrices.csv` has the price of each commodity per year and timeslice. eg. the cost of electricity at night for electricity in 2020.

Within each of the sector result folders, there is an output for `Capacity` for each commodity in each year. The years into the future, which the simulation has not run to, refers to the capacity as it retires. Within the `Residential` folder there is also a folder for `Supply` within each year. This refers to how much end-use commodity was output.

It is possible to extend the outputs using hooks. To see how to do this refer to the developer guide here.

2.2 Visualisation

Now, we can visualise the results of our first simulation! For this we will need to install a data management package called `pandas`, as well as data visualisation software, `matplotlib` and `seaborn`.

This can be done with the following command:

```
python -m pip install --user pandas
```

To install `matplotlib` and `seaborn` replace `pandas` in the previous command with the respective package name.

Firstly, we import the packages required to visualise the results.

[1]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Next, we load the dataset of interest to us for this example: the `MCACapacity.csv` file. We do this using pandas.

[2]:
```
capacity_results = pd.read_csv("Results/MCAcapacity.csv")
capacity_results.head()
```

[2]:

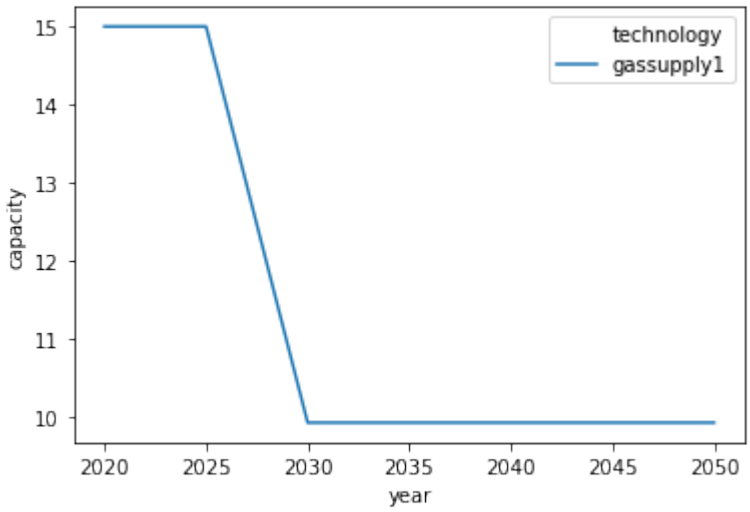
	technology	region	agent	type	sector	capacity	year
0	gasboiler	R1	A1	retrofit	residential	10.0	2020
1	gasCCGT	R1	A1	retrofit	power	1.0	2020
2	gassupply1	R1	A1	retrofit	gas	15.0	2020
3	gasboiler	R1	A1	retrofit	residential	5.0	2025
4	heatpump	R1	A1	retrofit	residential	19.0	2025

Using the `head` command we print the first five rows of our dataset. Next, we will visualise each of the sectors, with capacity on the y-axis and year on the x-axis.

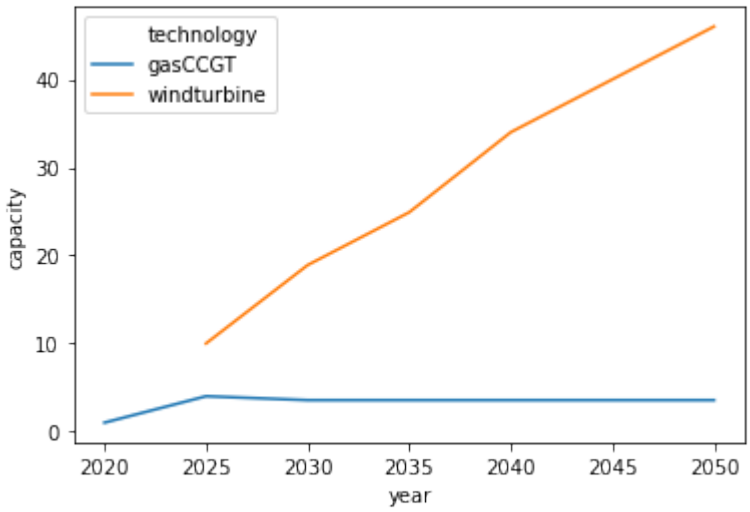
Don't worry too much about the code if some of it is unfamiliar. We effectively split the data into each sector and then plot a line plot for each.

[3]:
```
for sector_name, results in capacity_results.groupby("sector"):
    print("{} sector".format(sector_name))
    sns.lineplot(data=results, x="year", y="capacity", hue="technology")
    plt.show()
    plt.close()
```

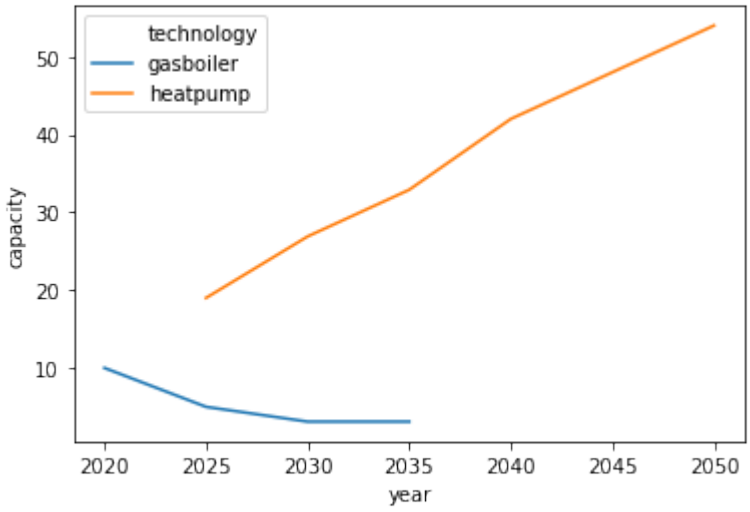
```
gas sector
```




```
power sector
```




```
residential sector
```



In this toy example, we can see that the end-use technology of choice in the residential sector becomes a heatpump. The heatpump displaces the gas boiler. Therefore, the supply of gas crashes due to a reduced demand. To account for the increase in demand for electricity, the agent invests heavily in wind turbines.

2.3 Next steps

If you want to jump straight into customising your own example scenarios, head to the link *here*. If you would like a little bit of background based on how MUSE works first, head to the next section!

CHAPTER

THREE

MUSE OVERVIEW

Note: TODO: Potentially find introductory image to place here.

MUSE is an open source agent-based modelling environment that can be used to simulate change in an energy system over time. An example of the type of question MUSE can help in answering is:

- How may a carbon budget affect investments made in the power sector over the next 30 years?

MUSE can incorporate residential, power, industrial and conversion sectors, meaning many questions can be explored using MUSE, as per the wishes of the user.

MUSE is an agent-based modelling environment, where the agents are investors and consumers. In MUSE, this means that investment decisions are made from the point of view of the investor and consumer. These agents can be heterogenous, enabling for differering investment strategies between agents, as in the real world.

MUSE is technology rich and can model energy production, conversion and end-use technologies. So, for example, MUSE can enable the user to develop a power sector with solar photovoltaics, wind turbines and gas power plants which produce energy for appliances like electric stoves, heaters and lighting in the residential sector. Agents invest within these sectors, investing in technologies such as electric stoves in the residential sector or gas power plants in the power sectors. The investments made depend on the agent's investment strategies.

Every sector is a user configurable module. This means that a user can configure any number of sectors, cointaining custom, user-defined technologies. In practice, this configuration is carried out using a selection of *Input Files*. In addition, MUSE can model any geographical region around the world and over any time scale, from a single year through to 100 years or more. Within a year, MUSE allows for a user-defined temporal granularity. This allows for the year to be split into different seasons and times, where energy demand may differ.

MUSE differs from the vast majority of energy systems models, which are intertemporal optimisation, by allowing agents to have "limited foresight". This enables these agents to invest under uncertainty of the future, as in the real world. In addition, MUSE is a "partial equilibrium" model, in the sense that it balances supply and demand of each energy commodity in the system.

3.1 What questions can MUSE answer?

MUSE allows for users to investigate how an energy system may evolve over a time period, based upon investors using different decision metrics or objectives such as the net present value, levelized cost of electricity or a custom-defined function. In addition to this, it can simulate how investors search for technology options, and how different objectives are combined to reach an investment decision.

The search for new technologies can depend on several factors such as agents' budgets, technology maturity or preferences on the fuel-type. For instance, an investor in the power sector may decide that they want to focus on renewable energy, whereas another may prefer the perceived most profitable option.

Examples of the questions MUSE can answer include:

- How may India's steel industry decarbonise?
- How might residential consumers change their investment decisions over time?
- How might a carbon tax impact investments made in the power sector?

3.2 How to use MUSE

There are a huge number of ways that MUSE could be used. The energy field is varied and diverse, and many different scenarios can be explored. Users can model the impact of changes in technology prices, demand, policy instruments, sector interactions and much, much more. People are always thinking of new ways that MUSE can be used. So, get creative!

A simulation model of a geographical region or world can be developed and is made up of the following features:

1. **Sectors** such as the power sector, gas production sector and the residential sector.
2. **Agents** such as a high-income subsection of the population in the UK or a risk-averse generation company. These agents are responsible for making investments in energy technologies.
3. **Technologies** which the agents choose to adopt. Technologies either produce an energy commodity (e.g. electricity), or a service demand (e.g. building space heating).
4. **Service demands** are demands that must be serviced such as lighting, heating or steel production.
5. **Market clearing algorithm** is the algorithm which determines global commodity prices based upon the balancing of supply and demand from each of the sectors.
6. **Equilibrium prices** are the prices determined by the market clearing algorithm and can determine the investments made by agents in various sectors. This allows for the model to project how the system may develop over a time period.

These features are described in more detail in the rest of this documentation.

3.3 What are MUSE's unique features?

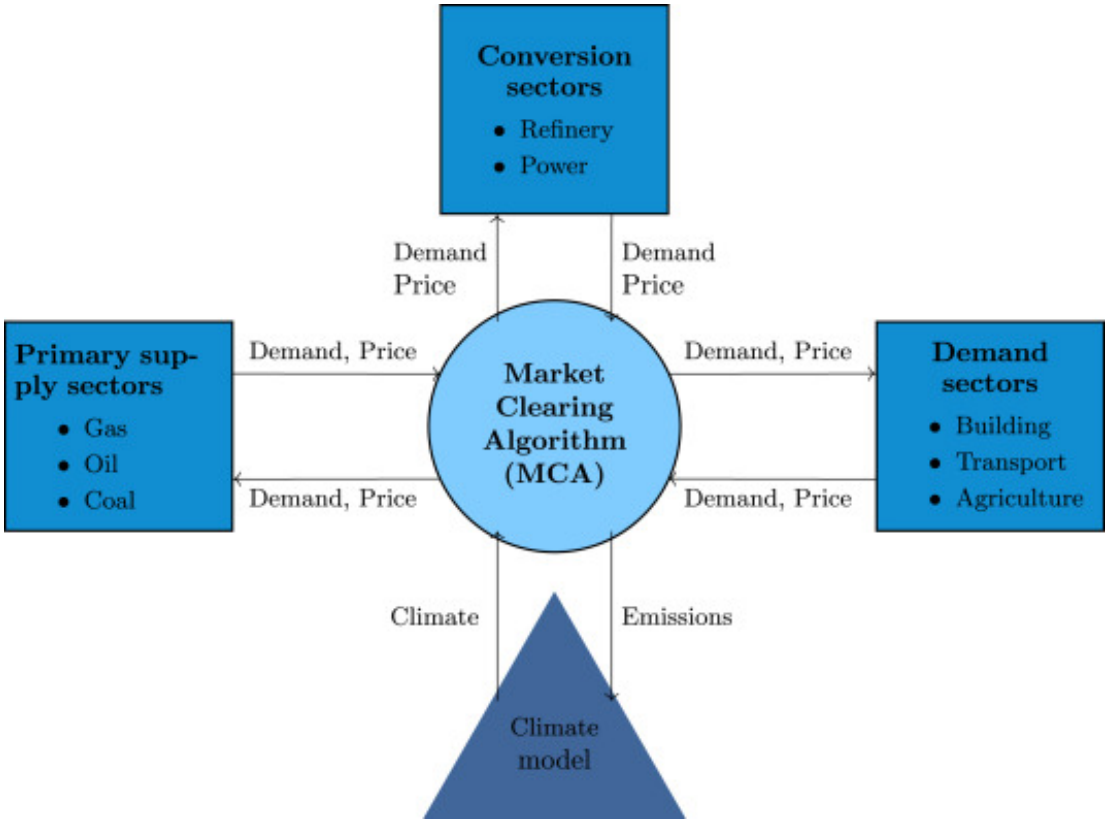
MUSE is a generalisable agent-based modelling environment and simulates energy transitions from the point of view of the investor and consumer agents. This means that users can define their own agents based upon their needs. The fact that MUSE is an agent-based model means that each of these agents can have different investment behaviours.

Additionally, agent-based models allow for agents to model imperfect information and limited foresight. An example of this is the ability to model the uncertainty residential users face when predicting the price of gas over the next 25 years. This is a unique feature to agent-based models when compared to intertemporal optimisation models and more closely models the real world. Many energy systems models are intertemporal optimisation models, which consider the viewpoint of a single benevolent decision maker, with perfect foresight and knowledge. These models optimise energy system investment and operation.

Whilst such intertemporal optimisation models are certainly useful, MUSE is different in that it models the incentives and challenges faced by investors. It can, therefore, be used to investigate different research questions, from the point of view of the investor and consumer. These questions are up to you, so impress us!

MUSE is completely open source, and ready for development.

3.4 Visualisation of MUSE



The figure above displays the key sectors of MUSE:

- Primary supply sectors
- Conversion sectors
- Demand sectors
- Climate model
- Market clearing algorithm (MCA)

3.5 How MUSE works

MUSE works by iterating between sectors shown above to ensure that energy service demands are met by the technologies chosen by the agents. Next, we detail the calculations made by MUSE throughout the simulation.

1. The energy service demand is calculated. For example, how much electricity, gas and oil demand is there for cooking, building space heating and lighting in the residential sector?
2. **A demand sector is solved. That is, agents choose which end-use technologies to serve the demands in the sector. For exam**

    i. Search space (which technologies are they willing to consider?)

    ii. Their objectives (which metrics do they consider important?)

    iii. Their decision rules (how do they choose to combine their metrics if they have multiple?)

3. The decisions made by the agents in the demand sectors then leads to a certain level of demand for energy commodities, such as electricity, gas and oil, as a whole. This demand is then passed to the MCA.
4. The MCA then sends these demands to the sectors that supply these energy commodities (supply or conversion sectors).
5. The supply and conversion sectors are solved: agents in these sectors use the same approach (i.e. search space, objectives, decision rules) to decide which technologies to investment in to serve the energy commodity demand. For example, agents in the power sector may decide to invest in solar photovoltaics, wind turbines and gas power plants to service the electricity demand.
6. As a result of these decisions a price for each energy commodity is formed based upon supply and demand. This is passed to the MCA.
7. The MCA then sends these prices back to the demand sectors, which are solved again as above.
8. This process repeats itself until commodity supply and demand converges for each energy commodity. Once these converge, the model has found a "partial equilibrium" and it moves forward to the next time period.

CHAPTER

FOUR

KEY MUSE COMPONENTS

MUSE is made up of five key components:

- Service Demand
- Technologies
- Sectors
- Agents
- Market Clearing Algorithm

In this section we will briefly explore what these components do and and how they interact.

4.1 Service Demand

The service demand is a user input which defines the demand that an end-use sector has. An example of this is the service demand commodity of heat or cooling that the residential sector requires. End-use in this case, refers to the energy which is utilised at the very final stage, after both extraction and conversion.

The estimate of the energy service is the first step. This estimate can be an exogenous input derived from the user, or correlations of GDP and population which reflect the socio-economic development of a region or country.

4.2 Technologies

Users are able to define any technology they wish for each of the energy sectors. Examples include power generators such as coal power plants, buses in the transport sector or lighting in the residential sector.

Each of the technologies are placed in their regions of interest, such as the USA or India. They are then defined by the following, but not limited to, variables:

- Capital costs
- Fixed costs
- Maximum capacity limit
- Maximum capaicty growth
- Lifetime of the technology
- Utilization factor
- Interest rate

Technologies, and their parameters are defined in the Technodata.csv file. For a full description of the input files, please refer to the *Techno-data* file.

4.3 Sectors

Sectors typically group areas of economic activity together, such as the residential sector, which might include all energy conusming activies of households. Possible examples of sectors are:

- Gas sector
- Power sector
- Residential sector
- Industrial sector

Each of these sectors contain their respective technologies which consume energy commodities. For example, the residential sector may consume electricity, gas or oil for a variety of different energy demands such as lighting, cooking and heating.

Each of the technologies, which consume a commodity, also output a different commodity or service demands. For example, a gas boiler consumes gas, but outputs heat and hot water.

4.4 Agents

Agents represent the investment decision makers in an energy system, for example consumers or companies. They invest in technologies that meet service demands, like heating, or produce other needed energy commodities, like electricity. These agents can be heterogenous, meaning that their investment priorities have the ability to differ.

As an example, a generation company could compare potential power generators based on their levelized cost of electricity, their net present value, by minimising the total capital cost, a mixture of these and/or any user-defined approach. This approach more closely matches the behaviour of real-life agents in the energy market, where companies, or people, have different priorities and constraints.

4.5 Market Clearing Algorithm

The market clearing algorithm (MCA) is the central component between the different supplies and demands of the energy system in question. The MCA iterates between the demand and supply of each of these sectors. Its role is to govern the endogenous price of commodities over the course of a simulation.

For a hypothetical example, the price of electricity is set to $70/MWh. However, at this price, the majority of residential agents prefer to heat their homes using gas. As a result of this, residential agents consume less electricity and more gas. This reduction in demand reduces the electricity price to $50/MWh. However, at this lower electricity price, some agents decide to invest in electric heating as opposed to gas. Eventually, the price converges on $60/MWh, where supply and demand for both electricity and gas are equal.

This is the principle of the MCA. It finds an equilibrium by iterating through each of the different sectors until an overall equilibrium is reached for each of the commodities. It is possible to run the MCA in a carbon budget mode, as well as exogenous mode. The carbon budget mode ensures that a carbon price limits the amount of carbon produce by the market. Whereas, the exogenous mode allows the carbon price to be set by the user.

CHAPTER

FIVE

CUSTOMISING MUSE TUTORIALS

Next, we show you how to customise MUSE to create your own scenarios.

We recommend following the tutorials step by step, as the files build on the previous example. If you prefer to jump straight in, your results may be different to the ones presented. To help you, we have provided the code to generate the various examples, in case you want to compare your code to ours.

5.1 Adding a new technology

5.1.1 Input Files

MUSE is made up of a number of different *input files*. These, however, can be broadly split into two:

- *Simulation settings*
- *Simulation data*

Simulation settings specify how a simulation should be run. For example, which sectors to run, for how many years and what to output.

Whereas, simulation data parametrises the technologies involved in the simulation, or the number and kinds of agents.

To create a customised case study it is necessary to edit both of these file types.

Simulation settings are specified in a TOML file. TOML is a simple, extensible and intuitive file format well suited for specifying small sets of complex data.

Simulation data is specified in CSV. This is a common format used for larger datasets, and is made up of columns and rows, with a comma used to differentiate between entries.

MUSE requires at least the following files to successfully run:

- a single *simulation settings TOML file* for the simulation as a whole
- a file indicating initial market price *projections*
- a file describing the *commodities in the simulation*
- for generalized sectors:
  - a file descring the *agents*
  - a file descring the *technologies*
  - a file descring the *input commodities* for each technology
  - a file descring the *output commodities* for each technology
  - a file descring the *existing capacity* of a given sector

- for each preset sector:
  - a csv file describing consumption for the duration of the simulation

For a full description of these files see the *input files section*. To see how to customise an example, continue on this page.

5.1.2 Addition of solar PV

In this section, we will add solar photovoltaics to the default model seen in the *example page*. To achieve this, we must modify some of the input files shown in the above section. These files can be found in the `StarMuse` folder at the following location:

`{muse_install_location}/src/muse/data/example/default`

Change `{muse_install_location}` to the location where you installed MUSE using your file browser. You can modify the files in your favourite spreadsheet editor or text editor such as Excel, Numbers, Notepad or TextEdit.

Technodata Input

Within the default folder there is the `settings.toml` file, input folder and technodata folder. To add a technology within the power sector, we must open the `technodata` folder followed by the `power` folder.

Next, we will edit the `CommIn.csv` file, which specifies the commodities consumed by solar photovoltaics.

The table below shows the original `CommIn.csv` version in normal text, and the added column and row in **bold**.

ProcessName	RegionName	Time	Level	electricity	gas	heat	CO2f	wind	**solar**
Unit	•	Year	•	PJ/PJ	PJ/PJ	PJ/PJ	kt/PJ	PJ/PJ	**PJ/PJ**
gasCCGT	R1	2020	fixed	0	1.67	0	0	0	**0**
windturbine	R1	2020	fixed	0	0	0	0	1	**0**
solarPV	**R1**	**2020**	**fixed**	**0**	**0**	**0**	**0**	**0**	**1**

We must first add a new row at the bottom of the file, to indicate the new solar photovoltaic technology:

- we call this technology `solarPV`
- place it in region `R1`
- the data in this row is associated to the year 2020
- the input type is fixed
- solarPV consumes solar

As the solar commodity has not been previously defined, we must define it by adding a column, which we will call solar. We fill out the entries in the solar column, ie. that neither `gasCCGT` nor `windturbine` consume solar.

We repeat this process for the file: `CommOut.csv`. This file specifies the output of the technology. In our case, solar photovoltaics only output `electricity`. This is unlike `gasCCGT` which also outputs `CO2f`, or carbon dioxide.

ProcessName	RegionName	Time	Level	electricity	gas	heat	CO2f	wind	**solar**
Unit	•	Year	•	PJ/PJ	PJ/PJ	PJ/PJ	kt/PJ	PJ/PJ	**PJ/PJ**
gasCCGT	R1	2020	fixed	1	0	0	91.67	0	**0**
windturbine	R1	2020	fixed	1	0	0	0	0	**0**
solarPV	**R1**	**2020**	**fixed**	**1**	**0**	**0**	**0**	**0**	**0**

Similar to the the `CommIn.csv`, we create a new row, and add in the solar commodity. We must ensure that we call our new commodity and technologies the same as the previous file for MUSE to successfully run. ie `solar` and `solarPV`

The next file to modify is the `ExistingCapacity.csv` file. This file details the existing capacity of each technology, per year. For this example, we will set the existing capacity to be 0.

ProcessName	RegionName	Unit	2020	2025	2030	2035	2040	2045	2050
gasCCGT	R1	PJ/y	1	1	0	0	0	0	0
windturbine	R1	PJ/y	0	0	0	0	0	0	0
solarPV	**R1**	**PJ/y**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

Finally, the `technodata.csv` containts parametrisation data for the technology, such as the cost, growth constraints, lifetime of the power plant and fuel used. The technodata file is too long for it all to be displayed here, so we will truncate the full version.

Here, we will only define the parameters: `processName`, `RegionName`, `Time`, `Level`,`cap_par`, `Fuel`,`EndUse`,`Agent2` and `Agent1`

We shall copy the existing parameters from the `windturbine` technology for the remaining parameters that can be seen in the `technodata.csv` file for brevity. You can see the full file here INSERT LINK HERE

ProcessName	RegionName	Time	Level	cap_par	cap_exp	...	Fuel	EndUse	Agent2	Agent1
Unit	•	Year	•	MUS$2010/PJ_a	•	...	•	•	Retrofit	New
gasCCGT	R1	2020	fixed	23.78234399		...	gas	electricity	1	0
windturbine	R1	2020	fixed	36.30771182		...	wind	electricity	1	0
solarPV	**R1**	**2020**	**fixed**	**30**	**1**	...	**solar**	**electricity**	**1**	**0**

Global inputs

Next, navigate to the `input` folder, found at `{muse_installation_location}src/muse/data/example/default/input`.

We now must edit each of the files found here to add the new solar commodity. Due to space constraints we will not display all of the entries contained in the input files. The edited files can be viewed here INSERT LINK HERE however.

The `BaseYearExport.csv` file defines the exports in the base year. For our example we add a column to indicate that there is no export for solar. However, it is important that a column exists for our new commodity.

RegionName	Attribute	Time	electricity	gas	heat	CO2f	wind	**solar**
Unit	•	Year	PJ	PJ	PJ	kt	PJ	**PJ**
R1	Exports	2010	0	0	0	0	0	**0**
R1	Exports	2015	0	0	0	0	0	**0**
…	…	…	…	…	…	…	…	**…**
R1	Exports	2100	0	0	0	0	0	**0**

The `BaseYearImport.csv` file defines the imports in the base year. Similarly to `BaseYearExport.csv`, we add a column for solar in the `BaseYearImport.csv` file. Again, we indicate that solar has no imports.

RegionName	Attribute	Time	electricity	gas	heat	CO2f	wind	**solar**
Unit	•	Year	PJ	PJ	PJ	kt	PJ	**PJ**
R1	Imports	2010	0	0	0	0	0	**0**
R1	Imports	2015	0	0	0	0	0	**0**
…	…	…	…	…	…	…	…	**…**
R1	Imports	2100	0	0	0	0	0	**0**

The `GlobalCommodities.csv` file is the file which defines the commodities. Here we give the commodities a commodity type, CO2 emissions factor and heat rate. For this file, we will add the solar commodity, with zero CO2 emissions factor and a heat rate of 1.

Commodity	Commodity-Type	Commodity-Name	CommodityEmissionFactor_CO2	HeatRate	Unit
Electricity	Energy	electricity	0	1	PJ
Gas	Energy	gas	56.1	1	PJ
Heat	Energy	heat	0	1	PJ
Wind	Energy	wind	0	1	PJ
CO2fuelcomsbustion	Environmental	CO2f	0	1	kt
Solar	**Energy**	**solar**	**0**	**1**	**PJ**

The `projections.csv` file details the initial market prices for the commodities. The market clearing algorithm will update these throughout the simulation, however, an initial estimate is required to start the simulation. As solar energy is free, we will indicate this by adding a final column.

RegionName	Attribute	Time	electricity	gas	heat	CO2f	wind	**solar**
Unit	•	Year	MUS$2010/PJ	MUS$2010/PJ	MUS$2010/PJ	MUS$2010/kt	MUS$2010/kt	**MUS$2010/kt**
R1	CommodityPrice	2010	14.814814	26.6759	100	0	0	**0**
R1	CommodityPrice	2015	17.898148	66.914325	100	0.0529138	10	**0**
…	…	…	…	…	…	…	…	**…**
R1	CommodityPrice	2100	21.398148	067.373485	819100	1.8712996	970	**0**

Running our customised simulation

Now we are able to run our simulation, with the new solar power technology.

To do this we run the same run command as previously in the anaconda command prompt:

```
python -m muse settings.toml
```

The output should be similar to the output here. However, expect the simulation to take slightly longer to run. This is due to the additional calculations made.

If the simulation has run successfully, you should now have a folder in the same location as your settings.toml file called Results. The next step is to visualise the results using the python visualisation package `seaborn` as well as the data analysis library `pandas`.

[2]:
```
import seaborn as sns
import pandas as pd
```

Next, we will import the `MCACapacity.csv` file into pandas and print the first 5 lines using the `head()` command.

Make sure to change the file path of `"../Results/MCACapacity.csv"` to where the `MCACapacity.csv` is on your computer, otherwise you will receive an error when you import the csv file.

[6]:
```
mca_capacity = pd.read_csv("../Results/MCACapacity.csv")
mca_capacity.head()
```

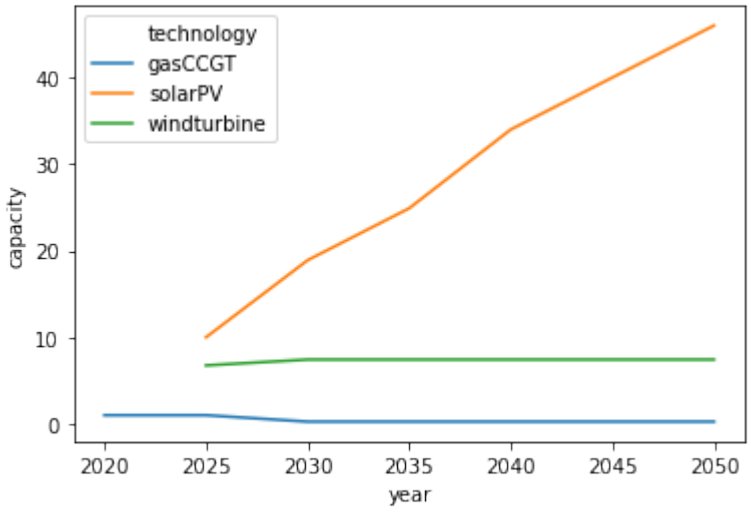
[6]:

	technology	region	agent	type	sector	capacity	year
0	gasboiler	R1	A1	retrofit	residential	10.0	2020
1	gasCCGT	R1	A1	retrofit	power	1.0	2020
2	gassupply1	R1	A1	retrofit	gas	15.0	2020
3	gasboiler	R1	A1	retrofit	residential	5.0	2025
4	heatpump	R1	A1	retrofit	residential	19.0	2025

We will only visualise the power sector in this example, as this was the only sector we changed. We, therefore, filter for this sector, and then visualise it using `seaborn`:

[7]:
```
power_capacity = mca_capacity[mca_capacity.sector=="power"]
sns.lineplot(data=power_capacity, x='year', y='capacity', hue="technology")
```

[7]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fd544a6e760>
```



We can now see that there is solarPV in addition to windturbine and gasCCGT, when compared to the example *here*! That's great and means it worked!

The difference in uptake of `solarPV` compared to `windturbine` is due to the fact that `solarPV` has a lower `cap_par` cost of 30, compared to the `windturbine`. Meaning that `solarPV` outcompetes both `windturbine` and `gasCCGT` in the electricity market.

Change Solar Price

Now, we will observe what happens if we increase the price of solar to be more expensive than wind. To achieve, this we have to modify the `Technodata.csv` file:

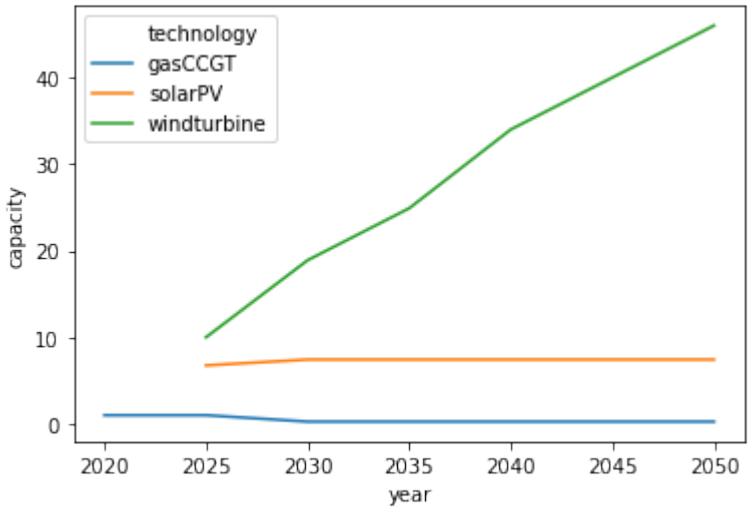
ProcessName	RegionName	Time	Level	cap_par	cap_exp	...	Fuel	EndUse	Agent2	Agent1
Unit	•	Year	•	MUS$2010/PJ_a	•	...	•	•	Retrofit	New
gasCCGT	R1	2020	fixed	23.782343	99	...	gas	electricity	1	0
windturbine	R1	2020	fixed	36.307711	82	...	wind	electricity	1	0
solarPV	R1	2020	fixed	**40**	1	...	solar	electricity	1	0

Here, we increase the `cap_par` variable by 10, to be a total of 40. We will now rerun the simulation, using the same command as previously and visualise the new results.

We must import the new `MCACapacity.csv` file again, and then visualise the results.

[8]:
```
mca_capacity = pd.read_csv("../Results/MCACapacity.csv")
power_capacity = mca_capacity[mca_capacity.sector=="power"]
sns.lineplot(data=power_capacity, x='year', y='capacity', hue="technology")
```

[8]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fd544ef7ee0>
```



Now, we can see that the technology `windturbine` outcompetes `solarPV` and `gasCCGT` due to the difference in price. The possibilities for creating your own scenarios are infinite.

For the full example with the completed input files see here INSERT LINK HERE

5.1.3 Next steps

In the next section we will add a new agent to the simulation.

5.2 Adding an agent

In this section, we will add a new agent called `A2`. This agent will be slightly different to the other agents in the `default` example, in that it will make investments based upon a mixture of levelised cost of electricity (LCOE) and net present value (NPV). These two objectives will be combined by calculating the mean of the two when comparing potential investment options.

To achieve this, we must modify the `Agents.csv` file in the directory:

```
{muse_install_location}/src/muse/data/example/default/technodata/Agents.csv
```

To do this, we will add two new rows to the file. To simplify the process, we copy the data from the first two rows of agent `A1`, changing only the rows: `Name`, `Objective1`, `Objective2`, `ObjData1`, `ObjData2` and `DecisionMethod`. The values we changed can be seen below. Again, we only show some of the rows due to space constraints, however see here for the full file.

AgentShare	Name	Agent-Number	Region-Name	Objective1	Objective2	Objective3	ObjData1	ObjData2	...	DecisionMethod	...	Type
Agent1	A1	1	R1	LCOE			1		...	singleObj	...	New
Agent2	A1	2	R1	LCOE			1		...	singleObj	...	Retrofit
Agent1	**A2**	**1**	**R1**	**LCOE**	**NPV**		**1**	**1**	**...**	**mean**	**...**	**New**
Agent2	**A2**	**2**	**R1**	**LCOE**	**NPV**		**1**	**1**	**...**	**mean**	**...**	**Retrofit**

We will now save this file and run the new simulation model using the following command:

```
python -m muse settings.toml
```

Again, we use seaborn and pandas to analyse the data in the `Results` folder.

[6]:
```
import pandas as pd
import seaborn as sns
```

[19]:
```
mca_capacity = pd.read_csv("../Results/MCACapacity.csv")
power_sector = mca_capacity[mca_capacity.sector=="power"]
power_sector.head()
```

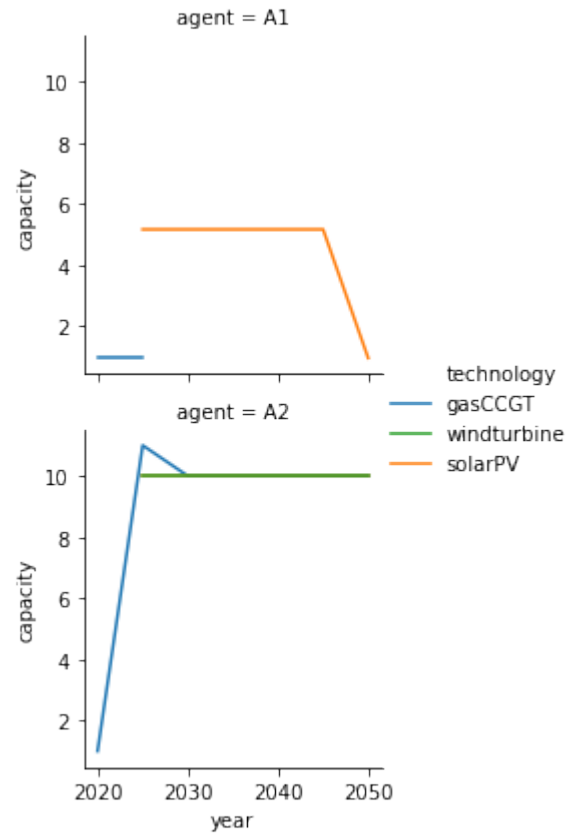
[19]:

	technology	region	agent	type	sector	capacity	year
2	gasCCGT	R1	A1	retrofit	power	1.000	2020
3	gasCCGT	R1	A2	retrofit	power	1.000	2020
10	gasCCGT	R1	A1	retrofit	power	1.000	2025
11	windturbine	R1	A1	retrofit	power	5.172	2025
12	gasCCGT	R1	A2	retrofit	power	11.000	2025

This time we can see that there is data for the new agent, `A2`. Next, we will visualise the investments made by each of the agents using seaborn's facetgrid command.

[20]:
```
g=sns.FacetGrid(power_sector, row='agent')
g.map(sns.lineplot, "year", "capacity", "technology")
g.add_legend()
```

[20]:
```
<seaborn.axisgrid.FacetGrid at 0x7f95819b4730>
```



In this scenario, agent `A1` is investing using LCOE, whereas agent `A2` is investing based on the mean of the objectives: LCOE and NPV in the same region. A different strategy is employed by these agents with `A2` investing in gasCCGT and windturbines, whereas `A1` invests in solarPV.

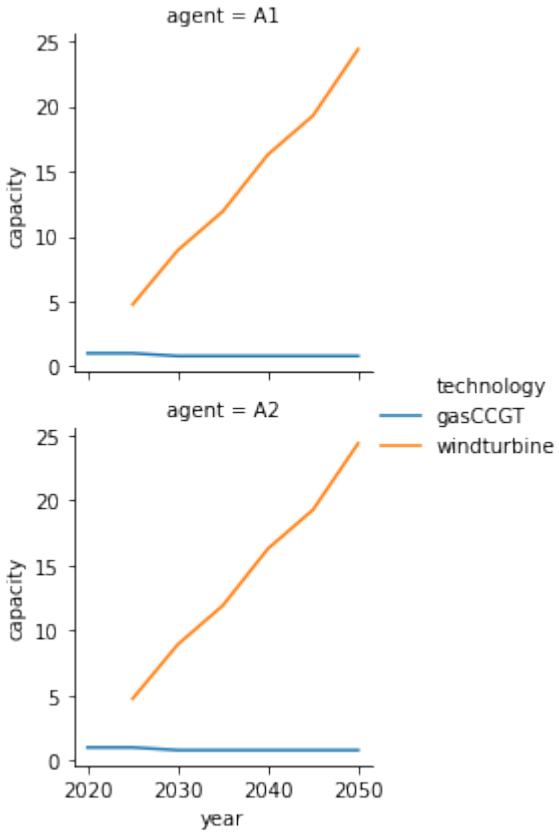
Next, we will see what occurs if the agents invest based upon the same investment strategy, with both investing using NPV. This requires to edit the `Agents.csv` file once more, to look like the following:

AgentShare	Name	Agent-Number	Region-Name	Objective1	Objective2	Objective3	ObjData1	ObjData2	...	DecisionMethod	...	Type
Agent1	A1	1	R1	LCOE			1		...	singleObj	...	New
Agent2	A1	2	R1	LCOE			1		...	singleObj	...	Retrofit
Agent1	A2	1	R1	**LCOE**			**1**		...	**singleObj**	...	New
Agent2	A2	2	R1	**LCOE**			**1**		...	**singleObj**	...	Retrofit

Again, this requires the re-running of the simulation, and visualisation like before:

[21]:
```
mca_capacity = pd.read_csv("../Results/MCACapacity.csv")
power_sector = mca_capacity[mca_capacity.sector=="power"]
g=sns.FacetGrid(power_sector, row='agent')
g.map(sns.lineplot, "year", "capacity", "technology")
g.add_legend()
```

[21]:
```
<seaborn.axisgrid.FacetGrid at 0x7f957e5f0970>
```



In this new scenario, with both agents running the same objective, very similar results can be seen, with a high investmwent in `windturbine`, none in `solarPV` and low `gasCCGT`. Have a play around with the files to see if you can come up with different scenarios!

5.2.1 Next steps

In the next section we will show you how to add a new region.

5.3 Adding a region

The next step is to add a region which we will call `R2`, however, this could equally be called `USA` or `India`. This requires a similar process to before of modifying the input simulation data. However, we will also have to change the `settings.toml` file to achieve this.

The process to change the `settings.toml` file is relatively simple. We just have to add our new region to the `regions` variable, in the 4th line of the `settings.toml` file, like so:

```
regions = ["R1", "R2"]
```

The process to change the input files, however, takes a bit more time. To achieve this, there must be data for each of the sectors for the new region. This, therefore, requires the modification of every *input file*.

Due to space constraints, we will not show you how to edit all of the files. However, you can access the modified files here INSERT LINK HERE.

Effectively, for this example, we will copy and paste the results for each of the input files from region `R1`, and change the name of the region for the new rows to `R2`.

However, as we are increasing the demand by adding a region, as well as modifying the costs of technologies, it may be the case that a higher growth in technology is required. For example, there may be no possible solution to meet demand without increasing the `windturbine` maximum allowed limit. We will therefore increase the allowed limits for `windturbine` in region `R2`.

We have placed two examples as to how to edit the residential sector below. Again, the edited data are highlighted in **bold**, with the original data in normal text.

The following file is the modified `/technodata/residential/CommIn.csv` file:

ProcessName	RegionName	Time	Level	electricity	gas	heat	CO2f	wind
Unit	•	Year	•	PJ/PJ	PJ/PJ	PJ/PJ	kt/PJ	PJ/PJ
gasboiler	R1	2020	fixed	0	1.16	0	0	0
heatpump	R1	2020	fixed	0.4	0	0	0	0
gasboiler	**R2**	**2020**	**fixed**	**0**	**1.16**	**0**	**0**	**0**
heatpump	**R2**	**2020**	**fixed**	**0.4**	**0**	**0**	**0**	**0**

Whereas the following file is the modified `/technodata/residential/ExistingCapacity.csv` file:

ProcessName	RegionName	Unit	2020	2025	2030	2035	2040	2045	2050
gasboiler	R1	PJ/y	10	5	0	0	0	0	0
heatpump	R1	PJ/y	0	0	0	0	0	0	0
gasboiler	**R2**	**PJ/y**	**10**	**5**	**0**	**0**	**0**	**0**	**0**
heatpump	**R2**	**PJ/y**	**0**	**0**	**0**	**0**	**0**	**0**	**0**

Below is the reduced `/technodata/power/technodata.csv` file, showing the increased capacity for `windturbine` in `R2`. For this, we highlight only the elements we changed from the rows in `R1`. The rest of the elements are the same for `R1` as they are for `R2`.

ProcessName	RegionName	...	MaxCapacityAddition	MaxCapacityGrowth	TotalCapacityLimit	...	Agent2	Agent1
Unit		...	PJ	%	PJ	...	Retrofit	New
gasCCGT	R1	...	2	0.02	60	...	1	0
windturbine	R1	...	2	0.02	60	...	1	0
solarPV	R1	...	2	0.02	60	...	1	0
gasCCGT	R2	...	2	0.02	60	...	1	0
windturbine	R2	...	**5**	**0.05**	**100**	...	1	0
solarPV	R2	...	2	0.02	60	...	1	0

Now, go ahead and amend all of the other input files for each of the sectors by copying and pasting the rows from `R1` and replacing the `RegionName` to `R2` for the new rows. All of the edited input files can be seen here.

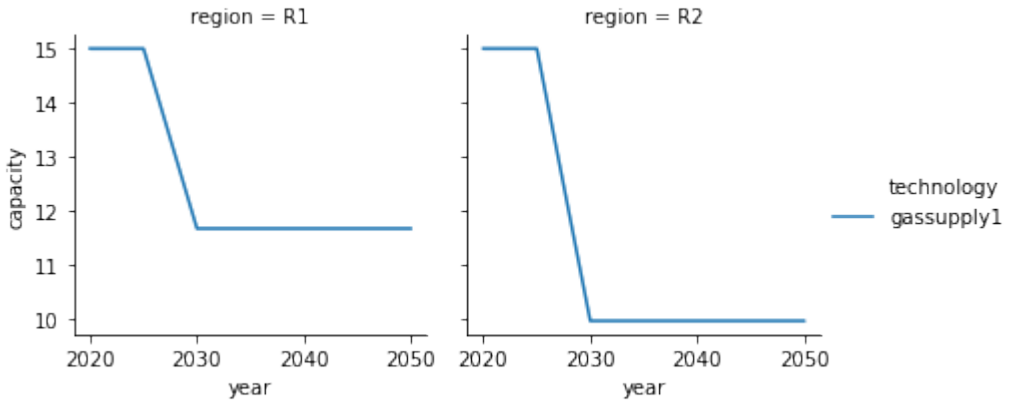
Again, we will run the results using the `python -m pip muse settings.toml` in anaconda prompt, and analyse the data as follows:

[1]:
```
import seaborn as sns
import pandas as pd
import matplotlib.pyplot as plt
```

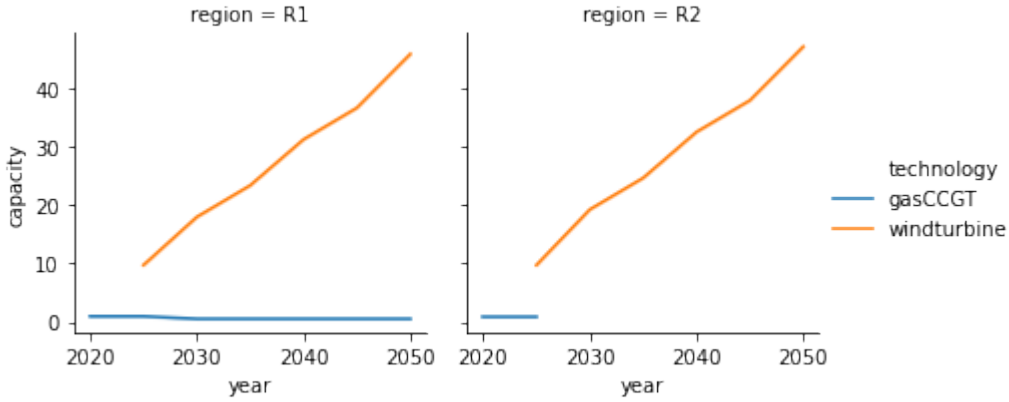
[2]:
```
mca_capacity = pd.read_csv("../tutorial-code/add-region/Results/MCACapacity.csv")

for name, sector in mca_capacity.groupby("sector"):
    print("{} sector:".format(name))
    g = sns.FacetGrid(data=sector, col="region")
    g.map(sns.lineplot, "year", "capacity", "technology")
    g.add_legend()
    plt.show()
```

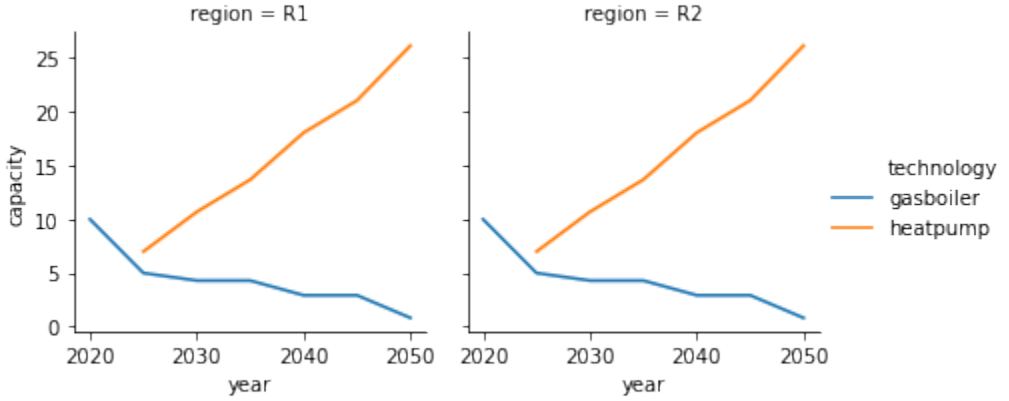
```
gas sector:
```




```
power sector:
```




```
residential sector:
```



Due to the similar natures of the two regions, with the parameters effectively copied and pasted between them, the results are very similar in both `R1` and `R2`. `gassupply1` drops significantly within the gas sector, due to the increasing demand of `heatpump` and falling demand of `gasboiler` in both region `R1` and `R2`. `windturbine` increases significantly to match this `heatpump` demand.

Have a play around with the various costs data in the technodata files for each of the sectors and technologies to see if different scenarios emerge. Although be careful. In some cases, the constraints on certain technologies will make it impossible for the demand to be met. Therefore you may have to relax these constraints.

5.3.1 Next steps

In the next section we modify the `settings.toml` file to change the timeslicing arrangements as well as project until 2040, instead of 2050, in two year timeslices.

5.4 Modification of time

In this section we will show you how to modify the timeslicing arrangement as well as change the time horizon and year intervals by modifying the `settings.toml` file.

5.4.1 Modify timeslicing

Timeslicing is the division of a single year into multiple different sections. For example, we could slice the year into different seasons, make a distinction between weekday and weekend or a distinction between morning and night. We do this as energy demand profiles can show a difference between these timeslices. eg. Electricity consumption is lower during the night than during the day.

To achieve this, we have to modify the `settings.toml` file, as well as the files within the preset folder: `Residential2020Consumption.csv` and `Residential2050Consumption.csv`. This is so that we can edit the demand for the residential sector for the new timeslices.

First we edit the `settings.toml` file to add two additional timeslices: early-morning and late-afternoon. We also rename afternoon to mid-afternoon. These settings can be found at the bottom of the `settings.toml` file.

An example of the changes is shown below:

```
[timeslices]
all-year.all-week.night = 1095
all-year.all-week.morning = 1095
all-year.all-week.mid-afternoon = 1095
all-year.all-week.early-peak = 1095
all-year.all-week.late-peak = 1095
all-year.all-week.evening = 1095
all-year.all-week.early-morning = 1095
all-year.all-week.late-afternoon = 1095
level_names = ["month", "day", "hour"]
```

Next, we modify both Residential Consumption files. Again, we put the text in bold for the modified entries. We must add the demand for the two additional timelsices, which we call timeslice 7 and 8. We make the demand for heat to be 2 for both of the new timeslices.

Below is the modified `Residential2020Consumption.csv` file:

	RegionName	ProcessName	Timeslice	electricity	gas	heat	CO2f	wind
0	R1	gasboiler	1	0	0	1	0	0
1	R1	gasboiler	2	0	0	1.5	0	0
2	R1	gasboiler	3	0	0	1	0	0
3	R1	gasboiler	4	0	0	1.5	0	0
4	R1	gasboiler	5	0	0	3	0	0
5	R1	gasboiler	6	0	0	2	0	0
6	**R1**	**gasboiler**	**7**	**0**	**0**	**2**	**0**	**0**
7	**R1**	**gasboiler**	**8**	**0**	**0**	**2**	**0**	**0**
0	R2	gasboiler	1	0	0	1	0	0
1	R2	gasboiler	2	0	0	1.5	0	0
2	R2	gasboiler	3	0	0	1	0	0
3	R2	gasboiler	4	0	0	1.5	0	0
4	R2	gasboiler	5	0	0	3	0	0
5	R2	gasboiler	6	0	0	2	0	0
6	**R2**	**gasboiler**	**7**	**0**	**0**	**2**	**0**	**0**
7	**R2**	**gasboiler**	**8**	**0**	**0**	**2**	**0**	**0**

We do the same for the `Residential2050Consumption.csv`, however this time we make the demand for heat in 2050 to both be 5 for the new timeslices. See here INSERT LINK HERE for the full file.

Once the relevant files have been edited, we are able to run the simulation model using `python -m muse settings.toml`.

Then, once run, we import the necessary packages:

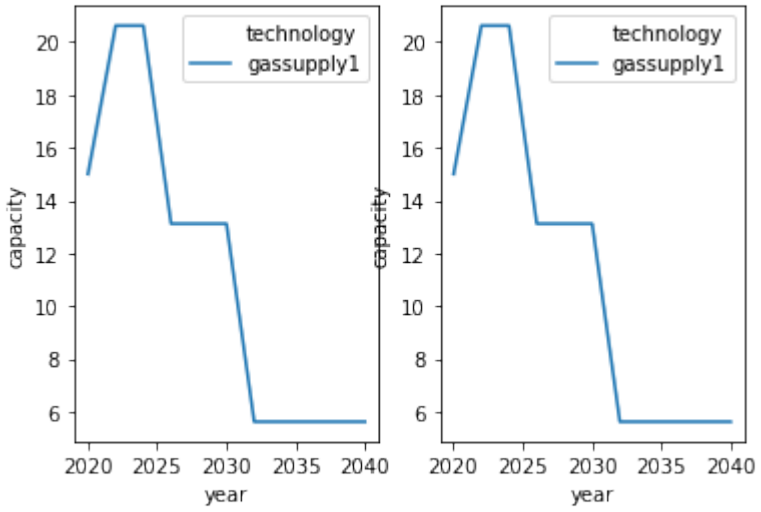
[1]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

and visualise the relevant data:

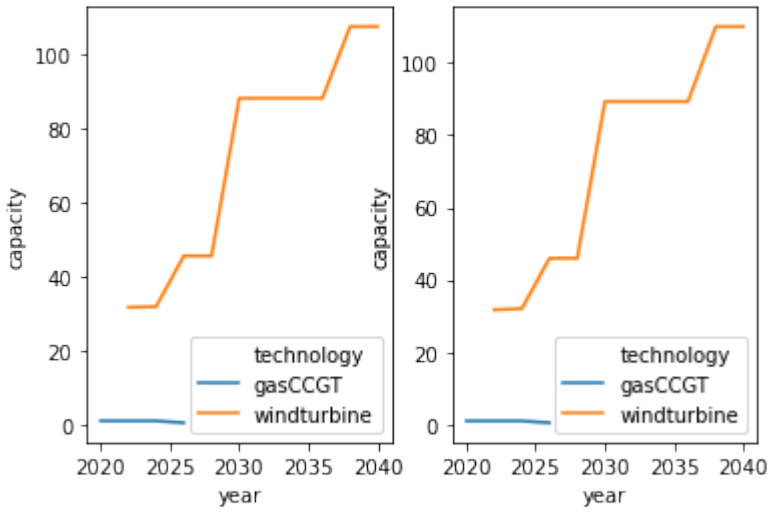
[2]:
```
mca_capacity = pd.read_csv("../tutorial-code/modify-timing-data/modify-time-framework/
↪Results/MCACapacity.csv")

for name, sector in mca_capacity.groupby("sector"):
    print("{} sector:".format(name))
    fig, ax =plt.subplots(1,2)
    sns.lineplot(data=sector[sector.region=="R1"], x="year", y="capacity", hue=
↪"technology", ax=ax[0])
    sns.lineplot(data=sector[sector.region=="R2"], x="year", y="capacity", hue=
↪"technology", ax=ax[1])
    plt.show()
    plt.close()
```

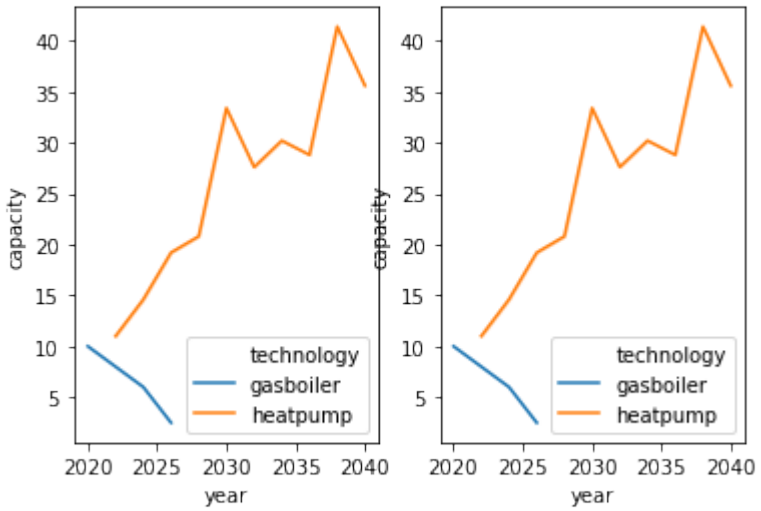
```
gas sector:
```




```
power sector:
```




```
residential sector:
```



Compared to the scenario where we added a *region*, there is a slight increase in solarPV in the power sector. However, the rest remains unchanged.

5.4.2 Modify time horizon and time periods

For the previous examples, we have run the scenario from 2020 to 2050, in 5 year time steps. This has been set at the top of the `settings.toml` file. However, we may want to run a more detailed scenario, with 2 year time steps, and up until the year 2040.

Making this change is quite simple as we only have two lines to change. We will modify line 2 and 3 of the `settings.toml` file, as follows:

```
# Global settings - most REQUIRED
time_framework = [2020, 2022, 2024, 2026, 2028, 2030, 2032, 2034, 2036, 2038, 2040]
foresight = 2   # Has to be a multiple of the minimum separation between the years in␣
↪time
```

The `time_framework` details each year in which we run the simulation. The `foresight` variable details how much foresight an agent has when making investments.

As we have modified the timeslicing arrangements there will be a change in the underlying demand for heating. This may require more electricity to service this demand. Therefore, we relax the constraints for growth in the power sector for all technologies and constraints in the `technodata/power/technodata.csv`, as is shown below:

ProcessName	RegionName	...	MaxCapacityAddition	MaxCapacityGrowth	TotalCapacityLimit	...	Agent1
Unit	•	...	PJ	%	PJ	...	New
gasCCGT	R1	...	**40**	**0.2**	**120**	...	0
windturbine	R1	...	**40**	**0.2**	**120**	...	0
solarPV	R1	...	**40**	**0.2**	**120**	...	0
gasCCGT	R2	...	**40**	**0.2**	**120**	...	0
windturbine	R2	...	**40**	**0.2**	**120**	...	0
solarPV	R2	...	**40**	**0.2**	**120**	...	0

We also modify the constraints defined in the `technodata.csv` file for the residential sector, as shown below:

ProcessName	RegionName	...	MaxCapacityAddition	MaxCapacityGrowth	TotalCapacityLimit	...	Agent1
Unit	•	...	PJ	%	PJ	...	New
gasboiler	R1	...	**60**	**0.5**	**120**	...	0
heatpump	R1	...	**60**	**0.5**	**120**	...	0
gasboiler	R2	...	**60**	**0.5**	**120**	...	0
heatpump	R2	...	**60**	**0.5**	**120**	...	0

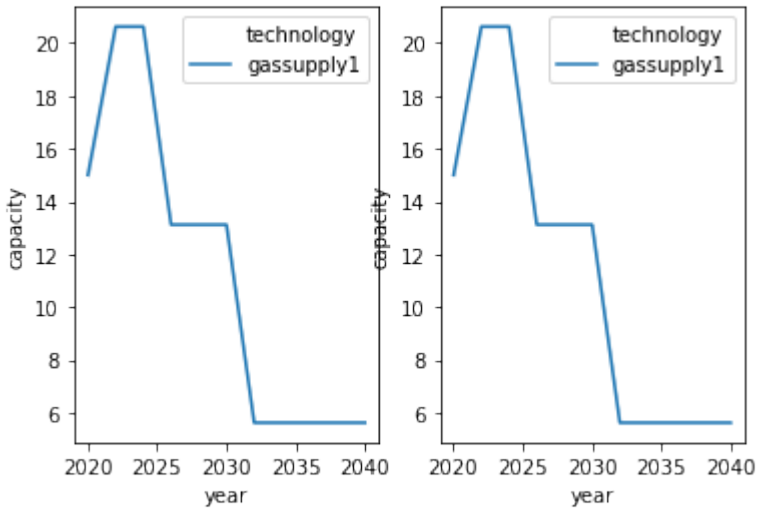
It must be noted, that this is a toy example. For modelling a real life scenario, data should be sought to ensure there remain realistic constriants.

For the full power sector `technodata.csv` file click here INSERT LINK HERE, and for the full residential sector `technodata.csv` file click here INSERT LINK HERE.

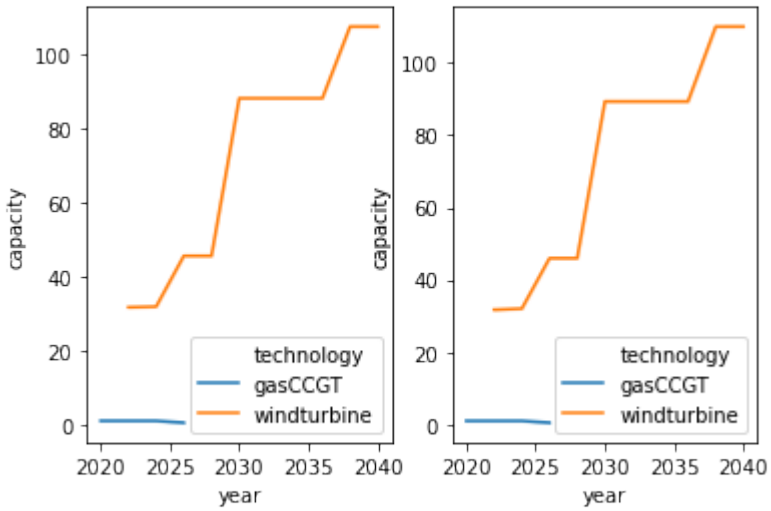
[3]:
```
mca_capacity = pd.read_csv("../tutorial-code/modify-timing-data/modify-time-framework/
↪Results/MCACapacity.csv")

for name, sector in mca_capacity.groupby("sector"):
    print("{} sector:".format(name))
    fig, ax =plt.subplots(1,2)
    sns.lineplot(data=sector[sector.region=="R1"], x="year", y="capacity", hue=
↪"technology", ax=ax[0])
    sns.lineplot(data=sector[sector.region=="R2"], x="year", y="capacity", hue=
↪"technology", ax=ax[1])
    plt.show()
    plt.close()
```

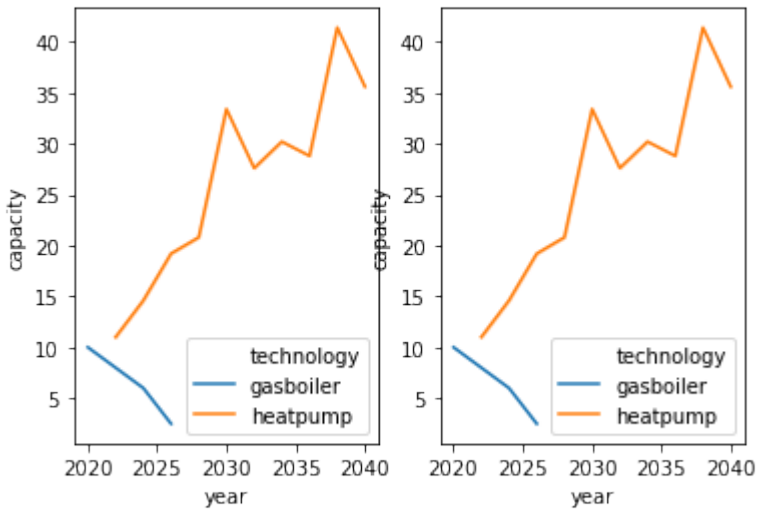
```
gas sector:
```




```
power sector:
```




```
residential sector:
```



5.4.3 Next steps

In the next section we detail how to add an exogenous service demand, such as demand for heating or cooking.

5.5 Adding a service demand

In this section, we will detail how to add a service demand to MUSE.

A service demand is an end-use demand. For example, in the residential sector, a service demand could be cooking. Houses require energy to cook food and a technology to service this demand, such as an electric stove.

This process consists of setting a demand, either through inputs derived from the user or correlations of GDP and population which reflect the socio-economic decvelopment of a region or country. In addition, a technology must be added to service the demand.

5.5.1 Addition of cooking demand

Firstly, we must add the demand section. In this example, we will add a cooking preset demand. To achieve this, we will now edit the `Residential2020Consumption.csv` and `Residential2050Consumption.csv` files, found within the `technodata/preset/` directory.

The `Residential2020Consumption.csv` file allows us to specify the demand in 2020 for each region and technology per timeslice. The `Residential2050Consumption.csv` file does the same but for the year 2050. The datapoints between these are interpolated.

Firstly, we must add the new service demand: `cook` as a column in these two files. Next, we add the demand. Again, the modified entries are in bold:

	RegionName	ProcessName	Timeslice	electricity	gas	heat	CO2f	wind	**cook**
0	R1	gasboiler	1	0	0	1	0	0	**0**
…	…	…	…	…	…	…	…	…	**…**
15	R2	gasboiler	8	0	0	2	0	0	**0**
16	**R1**	**electric_stove**	**1**	**0**	**0**	**0**	**0**	**0**	**1**
17	**R1**	**electric_stove**	**2**	**0**	**0**	**0**	**0**	**0**	**2**
18	**R1**	**electric_stove**	**3**	**0**	**0**	**0**	**0**	**0**	**1**
19	**R1**	**electric_stove**	**4**	**0**	**0**	**0**	**0**	**0**	**1.5**
20	**R1**	**electric_stove**	**5**	**0**	**0**	**0**	**0**	**0**	**2**
21	**R1**	**electric_stove**	**6**	**0**	**0**	**0**	**0**	**0**	**3**
22	**R1**	**electric_stove**	**7**	**0**	**0**	**0**	**0**	**0**	**2**
23	**R1**	**electric_stove**	**8**	**0**	**0**	**0**	**0**	**0**	**3**
24	**R2**	**electric_stove**	**1**	**0**	**0**	**0**	**0**	**0**	**1**
25	**R2**	**electric_stove**	**2**	**0**	**0**	**0**	**0**	**0**	**1**
26	**R2**	**electric_stove**	**3**	**0**	**0**	**0**	**0**	**0**	**1**
27	**R2**	**electric_stove**	**4**	**0**	**0**	**0**	**0**	**0**	**1.5**
28	**R2**	**electric_stove**	**5**	**0**	**0**	**0**	**0**	**0**	**2**
29	**R2**	**electric_stove**	**6**	**0**	**0**	**0**	**0**	**0**	**2**
30	**R2**	**electric_stove**	**7**	**0**	**0**	**0**	**0**	**0**	**2.5**
31	**R2**	**electric_stove**	**8**	**0**	**0**	**0**	**0**	**0**	**2**

For the purposes of brevity, we omitted the majority of the `gasboiler` entries. However, these remain unchanged, apart from a `0` entry in the cook column to indicate that a `gasboiler` does not meet `cook` demand.

We added an `electric_stove` process for each of the timeslices, which meets the `cook` demand. This can be seen through the addition of a positive number in the `cook` column.

The process is very similar for the `Residential2050Consumption.csv` file, however, for this example, we often placed larger numbers to indicate higher demand in 2050. For the complete file see the link here INCLUDE LINK HERE

Next, we must edit the files within the `input` folder. For this, we must add the `cook` service demand to each of these files.

First, we will amend the `BaseYearExport.csv` and `BaseYearImport.csv` files. For this, we say that there is no import or export of the `cook` service demand. A brief example is outlined below for `BaseYearExport.csv`:

RegionName	Attribute	Time	electricity	gas	heat	CO2f	wind	solar	**cook**
Unit	•	Year	PJ	PJ	PJ	kt	PJ	PJ	**PJ**
R1	Exports	2010	0	0	0	0	0	0	**0**
…	…	…	…	…	…	…	…	…	**…**
R2	Exports	2100	0	0	0	0	0	0	**0**

The same is true for the `BaseYearImport.csv` file:

RegionName	Attribute	Time	electricity	gas	heat	CO2f	wind	solar	**cook**
Unit	•	Year	PJ	PJ	PJ	kt	PJ	PJ	**PJ**
R1	Imports	2010	0	0	0	0	0	0	**0**
…	…	…	…	…	…	…	…	…	**…**
R2	Imports	2100	0	0	0	0	0	0	**0**

Next, we must edit the `GlobalCommodities.csv` file. This is where we define the new commodity `cook`. It tells MUSE the commodity type, name, emissions factor of CO2 and heat rate, amongst other things.

The example used for this tutorial is below:

Commodity	CommodityType	CommodityName	CommodityEmissionFactor_CO2	HeatRate	Unit
Electricity	Energy	electricity	0	1	PJ
Gas	Energy	gas	56.1	1	PJ
Heat	Energy	heat	0	1	PJ
Wind	Energy	wind	0	1	PJ
CO2fuelcomsbustion	Environmental	CO2f	0	1	kt
Solar	Energy	solar	0	1	PJ
Cook	**Energy**	**cook**	**0**	**1**	**PJ**

Finally, the `Projections.csv` file must be changed. This is a large file which details the expected cost of the technology in the first year of the simulation. Due to its size, we will only show two rows of the new column `cook`.

RegionName	Attribute	Time	…	**cook**
Unit	•	Year	…	**MUS$2010/kt**
R1	CommodityPrice	2010	…	**100**
…	…	…	…	**…**
R2	CommodityPrice	2100	…	**100**

We set every price of cook to be `100MUS$2010/kt`

5.5.2 Addition of cooking technology

Next, we must add a technology to service this new demand. This is achieved through a similar process as the section in the *"adding a new technology"* section. However, we must be careful to specify the end-use of the technology as `cook`.

For this example, we will add two competing technologies to service the cooking demand: `electric_stove` and `gas_stove` to the `Technodata.csv` file in `/technodata/residential/Technodata.csv`.

Again for the interests of space, we have omitted the existing `gasboiler` and `heatpump` technologies. But we copy the `gasboiler` row for `R1` and paste it for the new `electric_stove` for both `R1` and `R2`. For `gas_stove` we copy and paste the data for `heatpump` from region `R1` for both `R1` and `R2`.

An important modification, however, is specifying the end-use for these new technologies to be `cook` and not `heat`.

ProcessName	RegionName	Time	Level	cap_par	...	Fuel	EndUse	Agent2	Agent1
Unit	•	Year	•	MUS$2010/PJ_a		•	•	Retrofit	New
gasboiler	R1	2020	fixed	3.8	...	gas	heat	1	0
...	...	...	...	...	...	...	...	...	...
electric_stove	**R1**	**2020**	**fixed**	**3.8**	**...**	**electricity**	**cook**	**1**	**0**
electric_stove	**R2**	**2020**	**fixed**	**3.8**	**...**	**electricity**	**cook**	**1**	**0**
gas_stove	**R1**	**2020**	**fixed**	**8.8667**	**...**	**gas**	**cook**	**1**	**0**
gas_stove	**R2**	**2020**	**fixed**	**8.8667**	**...**	**gas**	**cook**	**1**	**0**

As can be seen we have added two technologies, in the two regions with different `cap_par` costs. We specified their respective fules, and the enduse for both is `cook`. For the full file please see here INSERT LINK HERE.

We must also add the data for these new technologies to the following files:

- `CommIn.csv`
- `CommOut.csv`
- `ExistingCapacity.csv`

This is largely a similar process to the tutorial shown in *"adding a new technology"*. We must add the input to each of the technologies (gas and electricity for `gas_stove` and `electric_stove` respectively), outputs of `cook` for both and the existing capacity for each technology in each region.

Due to the additional demand for gas and electricity brought on by the new `cook` demand, it is necessary to relax the growth constraints for `gassupply1` in the `technodata/gas/technodata.csv` file. For this example, we set this file as follows:

ProcessName	RegionName	Time	...	MaxCapacityAddition	MaxCapacityGrowth	TotalCapacityLimit	...	Agent1
Unit	•	Year	...	PJ	%	PJ	...	New
gassupply1	R1	2020	...	**100**	**5**	**500**	...	0
gassupply1	R2	2020	...	**100**	**5**	**120**	...	0

To prevent repetition of the *"adding a new technology"* section, we will leave the full files here INSERT LINK HERE.

Again, we run the simulation with our modified input files using the following command, in the relevant directory:

```
python -m pip muse settings.toml
```

Once this has run we are ready to visualise our results.

[2]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

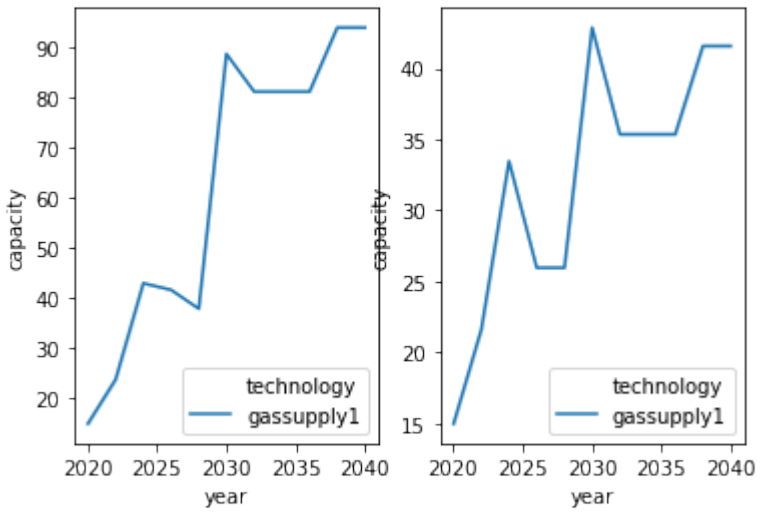
[3]:
```
mca_capacity = pd.read_csv("../tutorial-code/add-service-demand/Results/MCACapacity.
↪csv")
mca_capacity.head()
```

[3]:

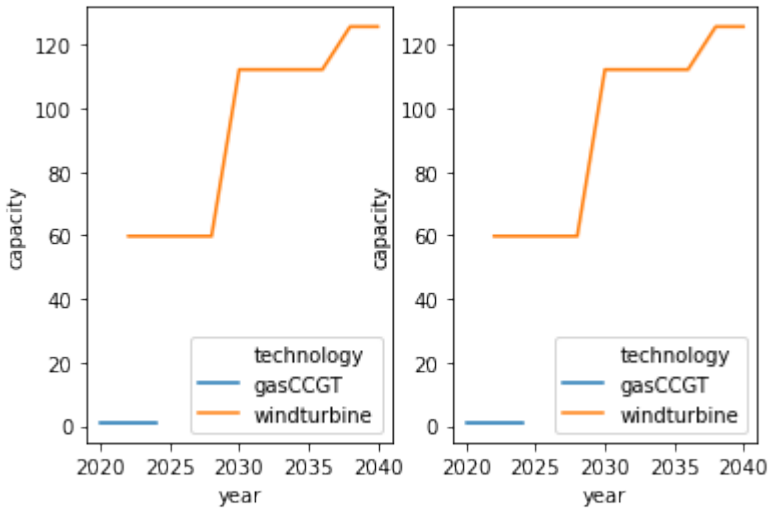
	technology	region	agent	type	sector	capacity	year
0	gas_stove	R1	A1	retrofit	residential	10.0	2020
1	gasboiler	R1	A1	retrofit	residential	10.0	2020
2	gas_stove	R2	A1	retrofit	residential	10.0	2020
3	gasboiler	R2	A1	retrofit	residential	10.0	2020
4	gas_stove	R1	A2	retrofit	residential	10.0	2020

[4]:
```
for name, sector in mca_capacity.groupby("sector"):
    print("{} sector:".format(name))
    fig, ax =plt.subplots(1,2)
    sns.lineplot(data=sector[sector.region=="R1"], x="year", y="capacity", hue=
↪"technology", ax=ax[0])
    sns.lineplot(data=sector[sector.region=="R2"], x="year", y="capacity", hue=
↪"technology", ax=ax[1])
    plt.show()
    plt.close()
```

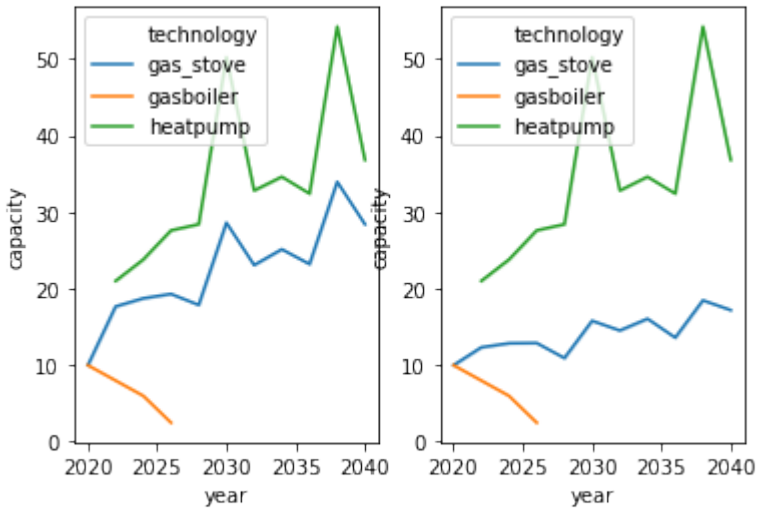
```
gas sector:
```




```
power sector:
```




```
residential sector:
```



We can see our new technology, the `gas_stove` is used over the `electric_stove`. Therefore, there is an increase in `gassupply1` to accommodate for this growth in demand. However, this is not enough to displace `windturbine` by `gasCCGT`.

5.5.3 Next steps

This brings us to the end of the user guide! Using the information explained in this tutorial, or following similar steps, you will be able to create complex scenarios of your choosing.

For the full code to generate the final results, see here INSERT LINK HERE.

CHAPTER

SIX

INPUT FILES

In this section we detail each of the files required to run MUSE. We include information based on how these files should be used, as well as the data that populates them.

6.1 TOML primer

The full specification for TOML files can be found here. A TOML file is separated into sections, with each section except the topmost introduced by a name in square brackets. Sections can hold key-value pairs, e.g. a name associated with a value. For instance:

```
general_string_attribute = "x"

[some_section]
section_attribute = 12

[some_section.subsection]
subsetion_attribute = true
```

TOML is quite flexible in how one can define sections and attributes. The following three examples are equivalent:

```
[sectors.residential.production]
name = "match"
costing = "prices"
```

```
[sectors.residential]
production = {"name": "match", "costing": "prices"}
```

```
[sectors.residential]
production.name = "match"
production.costing = "prices"
```

Additionally, TOML files can contain tabular data, specified row-by-row using double square bracket. For instance, below we define a table with two rows and a single *column* called *some_table_of_data* (though column is not quite the right term, TOML tables are made more flexible than most tabular formats. Rather, each row can be considered a dictionary).

```
[[some_table_of_data]]
a_key = "a value"

[[some_table_of_data]]
a_key = "another value"
```

As MUSE requires a number of data file, paths to files can be formated in a flexible manner. Paths can be formatted with shorthands for specific directories and are defined with curly-brackets. For example:

```
projection = '{path}/inputs/projection.csv'
timeslices_path = '{cwd}/technodata/timeslices.csv'
consumption_path = '{muse_sectors}/technodata/timeslices.csv'
```

path refers to the directory where the TOML file is located

cwd refers to the directory from which the muse simulation is launched

muse_sectors refers to the directory where default sectoral data is located

6.2 Simulation settings

This section details the TOML input for MUSE. The format for TOML files is described in this *previous section*. Here, however, we focus on sections and attributes that are specific to MUSE.

The TOML file can be read using `read_settings()`. The resulting data is used to construt the market clearing algorithm directly in the `MCA's factory function`.

6.2.1 Main section

This is the topmost section. It contains settings relevant to the simulation as a whole.

```
time_framework = [2020, 2025, 2030, 2035, 2040, 2045, 2050]
regions = ["USA"]
interpolation_mode = 'Active'
log_level = 'info'

equilibrium_variable = 'demand'
maximum_iterations = 100
tolerance = 0.1
tolerance_unmet_demand = -0.1
```

time_framework Required. List of years for which the simulation will run.

region Subset of regions to consider. If not given, defaults to all regions found in the simulation data.

interpolation_mode interpolation when reading the initial market. One of *linear*, *nearest*, *zero*, *slinear*, *quadratic*, *cubic*. Defaults to *linear*.

log_level: verbosity of the output.

equilibirum_variable whether equilibrium of *demand* or *prices* should be sought. Defaults to *demand*.

maximum_iterations Maximum number of iterations when searching for equilibrium. Defaults to 3.

tolerance Tolerance criteria when checking for equilibrium. Defaults to 0.1.

tolerance_unmet_demand Criteria checking whether the demand has been met. Defaults to -0.1.

excluded_commodities List of commodities excluded from the equilibrium considerations. Defaults to the list *["CO2f", "CO2r", "CO2c", "CO2s", "CH4", "N2O", "f-gases"]*.

plugins Path or list of paths to extra python plugins, i.e. files with registered functions such as `register_output_quantity()`.

6.2.2 Carbon market

This section containts the settings related to the modelling of the carbon market. If omitted, it defaults to not including the carbon market in the simulation.

Example

```
[carbon_budget_control]
budget = []
```

budget Yearly budget. There should be one item for each year the simulation will run. In other words, if given and not empty, this is a list with the same length as *time_framework* from the main section. If not given or an empty list, then the carbon market feature is disabled. Defaults to an empty list.

method Method used to equilibrate the carbon market. Defaults to a simple iterative scheme. [INSERT OPTIONS HERE]

commodities Commodities that make up the carbon market. Defaults to an empty list.

control_undershoot Whether to control carbon budget undershoots. Defaults to True.

control_overshoot Whether to control carbon budget overshoots. Defaults to True.

method_options: Additional options for the specific carbon method.

6.2.3 Global input files

Defines the paths specific simulation data files. The paths can be formatted as explained in the *TOML primer*.

```
[global_input_files]
projections = '{path}/inputs/Projections.csv'
regions = '{path}/inputs/Regions.csv'
global_commodities = '{path}/inputs/MUSEGlobalCommodities.csv'
```

projections: Path to a csv file giving initial market projection. See *Initial Market Projection*.

regions: Path to a csv file describing the regions. See *Regional data*.

global_commodities: Path to a csv file describing the comodities in the simulation. See *Commodity Description*.

6.2.4 Timeslices

Time-slices represent a sub-year disaggregation of commodity demand. Generally, timeslices are expected to introduce several levels, e.g. season, day, or hour. The simplest is to show the TOML for the default timeslice:

```
[timeslices]
winter.weekday.night = 396
winter.weekday.morning = 396
winter.weekday.afternoon = 264
winter.weekday.early-peak = 66
winter.weekday.late-peak = 66
winter.weekday.evening = 396
winter.weekend.night = 156
winter.weekend.morning = 156
winter.weekend.afternoon = 156
winter.weekend.evening = 156
spring-autumn.weekday.night = 792
```

(continues on next page)

(continued from previous page)

```
spring-autumn.weekday.morning = 792
spring-autumn.weekday.afternoon = 528
spring-autumn.weekday.early-peak = 132
spring-autumn.weekday.late-peak = 132
spring-autumn.weekday.evening = 792
spring-autumn.weekend.night = 300
spring-autumn.weekend.morning = 300
spring-autumn.weekend.afternoon = 300
spring-autumn.weekend.evening = 300
summer.weekday.night = 396
summer.weekday.morning  = 396
summer.weekday.afternoon = 264
summer.weekday.early-peak = 66
summer.weekday.late-peak = 66
summer.weekday.evening = 396
summer.weekend.night = 150
summer.weekend.morning = 150
summer.weekend.afternoon = 150
summer.weekend.evening = 150
level_names = ["month", "day", "hour"]
```

This input introduces three levels, via `level_names`: `month`, `day`, `hours`. Other simulations may want fewer or more levels. The `month` level is split into three points of data, `winter`, `spring-autumn`, `summer`. Then `day` splits out weekdays from weekends, and so on. Each line indicates the number of hours for the relevant slice. It should be noted that the slices are not a cartesian products of each levels. For instance, there no `peak` periods during weekends. All that matters is that the relative weights (i.e. the number of hours) are consistent and sum up to a year.

The input above defines the finest times slice in the code. In order to define rougher timeslices we can introduce items in each levels that represent aggregates at that level. By default, we have the following:

```
[timeslices.aggregates]
all-day = ["night", "morning", "afternoon", "early-peak", "late-peak", "evening"]
all-week = ["weekday", "weekend"]
all-year = ["winter", "summer", "spring-autumn"]
```

Here, `all-day` aggregates the full day. However, one could potentially create aggregates such as:

```
[timeslices.aggregates]
daylight = ["morning", "afternoon", "early-peak", "late-peak"]
nightlife = ["evening", "night"]
```

Once the finest timeslice and its aggregates are given, it is possible for each sector to define the timeslice simply by refering to the slices it will use at each level.

```
[sectors.some_sector.timeslice_levels]
day = ["daylight", "nightlife"]
month = ["all-year"]
```

Above, `sectors.some_sector.timeslice_levels.week` defaults its value in the finest timeslice. Indeed, if the subsection `sectors.some_sector.timeslice_levels` is not given, then the sector will default to using the finest timeslices.

Similarly, it is possible to specify a timeslice for the mca by adding an *mca.timeslice_levels* section. However, be aware that if the MCA uses a rougher timeslice framework, the market will be expressed within it. Hence information from sectors with a finer timeslice framework will be lost.

6.2.5 Standard sectors

Sectors are declared in the TOML file by adding a subsection to the *sectors* section:

```
[sectors.residential]
type = 'default'
[sectors.power]
type = 'default'
```

Above, we've added two sectors, residential and power. The name of the subsection is only used for identification. In other words, it should be chosen to be meaningful to the user, since it will not affect the model itself.

Sectors are defined in `Sector`.

A sector accepts a number of attributes and subsections.

type Defines the kind of sector this is. *Standard* sectors are those with type "default". This value corresponds to the name with which a sector class is registerd with MUSE, via `register_sector()`. [INSERT OTHER OPTIONS HERE]

priority An integer denoting which sectors runs when. Lower values imply the sector will run earlier. If two sectors share the same priority. Later sectors can depend on earlier sectors for the their input. If two sectors share the same priority, then their order is not defined. Indeed, it should indicate that they can run in parallel. For simplicity, the keyword also accepts standard values:

- "preset": 0
- "demand": 10
- "conversion": 20
- "supply": 30
- "last": 100

Defaults to "last".

interpolation Interpolation method user when filling in missing values. Available interpolation methods depend on the underlying scipy method's kind attribute.

investment_production In its simplest form, this is the name of a method to compute the production from a sector, as used when splitting the demand across agents. In other words, this is the computation of the production which affects future investments. In it's more general form, *production* can be a subsection of its own, with a "name" attribute. For instance:

```
[sectors.residential.production]
name = "match"
costing = "prices"
```

MUSE provides two methods in `muse.production`:

- **share: the production is the maximum production for the existing capacity and** the technology's utilization factor. See `muse.production.maximum_production()`.
- **match: production and demand are matched according to a given cost metric. The** cost metric defaults to "prices". It can be modified by using the general form given above, with a "costing" attribute. The latter can be "prices", "gross_margin", or "lcoe". See `muse.production.demand_matched_production()`.

production can also refer to any custom production method registered with MUSE via `muse.production.register_production()`.

Defaults to "share".

dispatch_production The name of the production method used to compute the sector's output, as returned to the muse market clearing algorithm. In other words, this is computation of the production method which will affect other sectors.

It has the same format and options as the *production* attribute above.

demand_share A method used to split the MCA demand into seperate parts to be serviced by specific agents. There is currently only one option, "new_and_retro", corresponding to *new* and *retro* agents.

interactions Defines interactions between agents. These interactions take place right before new investments are computed. The interactions can be anything. They are expected to modify the agents and their assets. MUSE provides a default set of interactions that have *new* agents pass on their assets to the corresponding *retro* agent, and the *retro* agents pass on the make-up of their assets to the corresponding *new* agents.

interactions are specified as a *TOML array*, e.g. with double brackets. Each sector can specify an arbitrary number of interactaction, simply by adding an extra interaction row.

There are two orthogonal concepts to interactions:

- a *net* defines the set of agents that interact. A set can contain any number of agents, whether zero, two, or all agents in a sector. See `muse.interactions.register_interaction_net()`.
- an *interaction* defines how the net actually interacts. See `muse.interactions.register_agent_interaction()`.

In practice, we always consider sequences of nets (i.e. more than one net) that interact using the same interaction function.

Hence, the input looks something like the following:

```
[[sectors.commercial.interactions]]
net = 'new_to_retro'
interaction = 'transfer'
```

"new_to_retro" is a function that figures out all "new/retro" pairs of agents. Whereas "transfer" is a function that performs the transfer of assets and information between each pair.

Furthermore, it is possible to pass parameters to either the net of the interaction as follows:

```
[[sectors.commercial.interactions]]
net = {"name": "some_net", "param": "some value"}
interaction = {"name": "some_interaction", "param": "some other value"}
```

The parameters will depend on the net and interaction functions. Neither "new_to_retro" nor "transfer" take any arguments at this point. MUSE interaction facilities are defined in `muse.interactions`.

output Outputs are made up of several components. MUSE is designed to allow users to mix-and-match both how and what to save.

output is specified as a TOML array, e.g. with double brackets. Each sector can specify an arbitrary number of outputs, simply by adding an extra output row.

A single row looks like this:

```
[[sectors.commercial.outputs]]
filename = '{cwd}/Results/{Sector}/{Quantity}/{year}{suffix}'
quantity = "capacity"
sink = 'csv'
overwrite = true
```

The following attributes are available:

- **quantity: Name of the quantity to save. Currently, only *capacity* exists,** refering to `muse.outputs.capacity()`. However, users can customize and create further output quantities by registering with MUSE via `muse.outputs.register_output_quantity()`. See `muse.outputs` for more details.
- **sink: the sink is the place (disk, cloud, database, etc...) and format with which** the computed quantity is saved. Currently only sinks that save to files are implemented. The filename can specified via *filename*, as given below. The following sinks are available: "csv", "netcfd", "excel". However, more sinks can be added by interested users, and registered with MUSE via `muse.outputs.register_output_sink()`. See `muse.outputs` for more details.
- **filename: defines the format of the file where to save the data. There are several** standard values that are automatically substituted:
  - cwd: current working directory, where MUSE was started
  - path: directory where the TOML file resides
  - sector: name of the current sector (.e.g. "commercial" above)
  - Sector: capitalized name of the current sector
  - quantity: name of the quantity to save (as given by the quantity attribute)
  - Quantity: capitablized name of the quantity to save
  - year: current year
  - suffix: standard suffix/file extension of the sink

  Defaults to *{cwd}/{default_output_dir}/{Sector}/{Quantity}/{year}{suffix}*.
- **overwrite: If *False* MUSE will issue an error and abort, instead of** overwriting an existing file. Defaults to *False*. This prevents important output files from being overwritten.

technodata Path to a csv file containing the characterization of the technologies involved in the sector, e.g. lifetime, capital costs, etc... See *Techno-data*.

timeslice_levels Slices to consider in a level. If absent, defaults to the finest timeslices. See *Timeslices*

commodities_in Path to a csv file describing the inputs of each technology involved in the sector. See *General features*.

commodities_out Path to a csv file describing the outputs of each technology involved in the sector. See *Output Commodities*.

existing_capacity Path to a csv file describing the initial capacity of the sector. See *Existing Sectoral Capacity*.

agents Path to a csv file describing the agents in the sector. See *Agents*.

6.2.6 Preset sectors

The commodity production, commodity consumption and product prices of preset sectors are determined exogeneously. They are know from the start of the simulation and are not affected by the simulation.

Preset sectors are defined in `PresetSector`.

The three components, production, consumption, and prices, can be set independantly and not all three need to be set. Production and consumption default to zero, and prices default to leaving things unchanged.

The following defines a standard preset sector where consumption is defined as a function of macro-economic data, i.e. population and gdp.

```
[sectors.commercial_presets]
type = 'presets'
priority = 'presets'
timeslice_shares_path = '{path}/technodata/TimesliceShareCommercial.csv'
macrodrivers_path = '{path}/technodata/Macrodrivers.csv'
regression_path = '{path}/technodata/regressionparameters.csv'
timeslices_levels = {'day': ['all-day']}
forecast = [0, 5]
```

The following attributes are accepted:

type: See the attribute in the standard mode, *type*. *Preset* sectors are those with type "presets".

priority See the attribute in the standard mode, *priority*.

timeslices_levels: See the attribute in the standard mode, *Timeslices*.

consumption_path: CSV output files, one per year. This attribute can include wild cards, i.e. '*', which can match anything. For instance: *consumption_path = "{cwd}/Consumtion*.csv"* will match any csv file starting with "Consumption" in the current working directory. The file names must include the year for which it defines the consumption, e.g. *Consumption2015.csv*.

The CSV format should follow the following format:

Table 1: Consumption

	RegionName	ProcessName	TimeSlice	electricity	diesel	algae
0	USA	fluorescent light	1	1.9	0	0
1	USA	fluorescent light	2	1.8	0	0

The index column as well as "RegionName", "ProcessName", and "TimeSlice" must be present. Further columns are reserved for commodities. "TimeSlice" refers to the index of the timeslice.

supply_path: CSV file, one per year, indicating the amount of a commodities produced. It follows the same format as *consumption_path*.

supply_path: CSV file, one per year, indicating the amount of a commodities produced. It follows the same format as *consumption_path*.

prices_path: CSV file indicating the amount of a commodities produced. The format of the CSV files follows that of *Initial Market Projection*.

demand_path: Incompatible with *consumption_path* or *macrodrivers_path*. A CSV file containing the consumption in the same format as *Initial Market Projection*.

macrodrivers_path: Incompatible with *consumption_path* or *demand_path*. Path to a CSV file giving the profile of the macrodrivers. Also requires *regression_path*.

regression_path: Incompatible with *consumption_path* or *demand_path*. Path to a CSV file giving the regression parameters with respect to the macrodrivers. Also requires *macrodrivers_path*.

timeslice_shares_path Optional csv file giving shares per timeslice. Requires *macrodrivers_path*.

filters: Optional dictionary of entries by which to filter the consumption. Requires *macrodrivers_path*. For instance,

```
filters.region = ["USA", "ASEA"]
filters.commodity = ["algae", "fluorescent light"]
```

6.2.7 Legacy Sectors

Legacy sectors wrap sectors developed for a previous version of MUSE to the open-source version.

Preset sectors are defined in `PresetSector`.

The can be defined in the TOML file as follows:

```
[global_input_files]
macrodrivers = '{path}/input/Macrodrivers.csv'
regions = '{path}/input/Regions.csv'
global_commodities = '{path}/input/MUSEGlobalCommodities.csv'

[sectors.Industry]
type = 'legacy'
priority = 'demand'
agregation_level = 'month'
excess = 0

userdata_path = '{muse_sectors}/Industry'
technodata_path = '{muse_sectors}/Industry'
timeslices_path = '{muse_sectors}/Industry/TimeslicesIndustry.csv'
output_path = '{path}/output'
```

For historical reasons, the three *global_input_files* above are required. The sector itself can use the following attributes.

type: See the attribute in the standard mode, *type*. *Legacy* sectors are those with type "legacy".

priority See the attribute in the standard mode, *priority*.

agregation_level: Information relevant to the sector's timeslice.

excess: Excess factor used to model early obsolescence.

timeslices_path: Path to a timeslice *time_slices*.

userdata_path: Path to a directory with sector-specific data files.

technodata_path: Path to a technodata CSV file. See. *Techno-data*.

output_path: Path to a diretory where the sector will write output files.

6.3 Input Files

6.3.1 Initial Market Projection

MUSE needs an initial projection of the market prices for each period of the simulation.

- The price trajectory is needed if the MCA works in *equilibrium* mode as an initial trajectory for the base year of the simulation. The market will override the calculated prices obtained from each commodity equilibrium for all the future periods following the base year
- Similarly, if the market works in a *carbon budget* mode, the prices are used as a starting point. The only difference from the previous case is given by the fact that the MCA will be calculating an additional global market price for carbon dioxide (and additional pollutants if required)
- If the MCA works in an *exogenous* mode, it will use the initial market projection as the projection for the the base year and all the future periods of the simulation

The forward price trajectory should follow the structure reported in the table below.

Table 2: Initial market projections

RegionName	Attribute	Time	com1	com2	com3
Unit	•	Year	MUS$2010/PJ	MUS$2010/PJ	MUS$2010/PJ
region1	CommodityPrice	2010	20	1.9583	2
region1	CommodityPrice	2015	20	1.9583	2
region1	CommodityPrice	2020	20.38518042	1.996014941	2.038518042
region1	CommodityPrice	2025	20.77777903	2.034456234	2.077777903
region1	CommodityPrice	2030	21.17793872	2.073637869	2.117793872
region1	CommodityPrice	2035	21.58580508	2.113574105	2.158580508
region1	CommodityPrice	2040	22.00152655	2.154279472	2.200152655
region1	CommodityPrice	2045	22.42525441	2.195768786	2.242525441
region1	CommodityPrice	2050	22.85714286	2.238057143	2.285714286

RegionName represents the region ID and needs to be consistent across all the data inputs

Attribute defines the attribute type. In this case it refers to the CommodityPrice; it is relevant only for internal use

Time corresponds to the time periods of the simulation; the simulated time framework in the example goes from 2010 through to 2050 with a 5-year time step

com1, . . . , comN Any further columns represent the commodities modelled, as defined in the global commodities the row Unit reports the unit in which the technology consumption is defined; it is for the user internal reference only. The names *comX* should be replaced with the names of the commodities.

6.3.2 Regional data

MUSE requires the definition of the methodology used for investment and dispatch and alias demand matching. The methodology has to be defined by region and subregion, meant as a geographical subdivision in a region. Currently, the methodology definition is important for the legacy sectors only.

Below the generic structure of the input commodity file for the electric heater is shown:

Table 3: Methodology used in investment and demand matching

SectorName	RegionName	Subregion	sMethodologyPlanning	sMethodologyDispatch
Agriculture	region1	region1	NPV	DCF
Bioenergy	region1	region1	NPV	DCF
Industry	region1	region1	NPV	DCF
Residential	region1	region1	EAC	EAC
Commercial	region1	region1	EAC	EAC
Transport	region1	region1	LCOE	LCOE
Power	region1	region1	LCOE	LCOE
Refinery	region1	region1	LCOE	LCOE
Supply	region1	region1	LCOE	LCOE

SectorName represents the sector_ID and needs to be consistent across the data input files

RegionName represents the region ID and needs to be consistent across all the data inputs

Subregion represents the subregion ID and needs to be consistent across all the data inputs

sMethodologyPlanning reports the cost quantity used for making investments in new technologies in each sector (e.g. NPV stands for net present value, EAC stands for equivalent annual costs, LCOE stands for levelised cost of energy)

sMethodologyDispatch reports the cost quantity used for the demand matching using existing technologies in each sector (e.g. DCF stands for discounted cash flow, EAC stands for equivalent annual cost, LCOE stands for levelised cost of energy)

6.3.3 Commodity Description

MUSE handles a configurable number and type of commodities which are primarily used to represent energy, services, pollutants/emissions. The commodities for the simulation as a whole are defined in a csv file with the following structure.

Table 4: Global commodities

Commodity	Commodity-Type	Commodity-Name	CommodityEmissionFactor_CO2	HeatRate	Unit
Coal	Energy	hardcoal	94.6	29	PJ
Agricultural-residues	Energy	agrires	112	15.4	PJ

Commodity represents the extended name of a commodity

CommodityType defines the type of a commodity (i.e. energy, material or environmental)

CommodityName is the internal name used for a commodity inside the model.

CommodityEmissionFactor_CO2 is CO2 emission per unit of commodity flow

HeatRate represents the lower heating value of an energy commodity

Unit is the unit used as a basis for all the input data. More specifically the model allows a totally flexible way of defining the commodities. CommodityName is currently the only column used internally as it defines the names of commodities and needs to be used consistently across all the input data files. The remaining columns of the file are only relevant for the user internal reference for the original sets of assumptions used.

6.3.4 Techno-data

The techno-data includes the techno-economic characteristics of each technology such as capital, fixed and variable cost, lifetime, utilisation factor. The techno-data should follow the structure reported in the table. The column order is not important and additional input data can alsobe read in this format. In the table, the electric boiler used in households is taken as an example for a generic region, region1.

Table 5: Techno-data

ProcessName	RegionName	Time	Level	cap_par	cap_exp	fix_par	...
resBoilerElectric	region1	2010	fixed	3.81	1.00	0.38	...
resBoilerElectric	region1	2030	fixed	3.81	1.00	0.38	...

ProcessName represents the technology ID and needs to be consistent across all the data inputs

RegionName represents the region ID and needs to be consistent across all the data inputs

Time represents the period of the simulation to which the value applies; it needs to contain at least the base year of the simulation

Level characterises either a fixed or a flexible input type

cap_par, cap_exp are used in the capital cost estimation. Capital costs are calculated as:

$$\text{CAPEX} = \text{cap\_par} * (\text{Capacity})^{\text{cap\_exp}}$$

where the parameter cap_par is estimated at a selected reference size (i.e. Capref), such as:

$$\text{cap\_par} = \left(\frac{\text{CAPEXref}}{\text{Capref}}\right)^{\text{cap\_exp}}$$

Capref is decided by the modeller before filling the input data files.

This allows the model to take into account economies of scale. ie. As *Capacity* increases, the price of the technology decreases.

fix_par, fix_exp

are used in the fixed cost estimation. Fixed costs are calculated as:

$$\text{FOM} = \text{fix\_par} * (\text{Capacity})^{\text{fix\_exp}}$$

where the parameter fix_par is estimated at a selected reference size (i.e. Capref), such as:

$$\text{fix\_par} = \left(\frac{\text{FOMref}}{\text{Capref}}\right)^{\text{fix\_exp}}$$

Capref is decided by the modeller before filling the input data files.

var_par, var_exp are used in the variable costs estimation. These variable costs are capacity dependent Variable costs are calculated as:

$$\text{VAREX} = \text{cap\_par} * (\text{Capacity})^{\text{cap\_exp}}$$

where the parameter var_par is estimated at a selected reference size (i.e. Capref), such as:

$$\text{var\_par} = \left(\frac{\text{VARref}}{\text{Capref}}\right)^{\text{var\_exp}}$$

Capref is decided by the modeller before filling the input data files.

MaxCapacityAddition represents the maximum addition of installed capacity per technology, region, year.

MaxCapacityGrowth represents the maximum growth in capacity as a fraction of the installed capacity per technology, region and year.

TotalCapacityLimit represents the total capacity limit per technology, region and year.

TechnicalLife represents the number of years that a technology operates before it is decommissioned.

UtilizationFactor is the number of operating hours of a process over the maximum number of hours in a year.

ScalingSize represents the minimum size of a technology to be installed.

efficiency is calculated as the ratio between the total output commodities and the input commodities.

AvailabiliyYear defines the starting year of a technology; for example the value equals 1 when a technology would be available or 0 when a technology would not be available.

Type defines the type of a technology.

Fuel defines the fuel used by a technology.

EndUse defines the end use of a technology.

InterestRate is the technology interest rate.

Agent_0, . . . , Agent_N represent the allocation of the initial capacity to the each agent.

The input data has to be provided for the base year. Additional years within the time framework of the overall simulation can be defined. In this case, MUSE would interpolate the values between the provided periods and assume a constant value afterwards.

6.3.5 Time-slices

Note: This input file is only for legacy sectors. For anything else, please see *Simulation settings*.

Time-slices represent a sub-year disaggregation of commodity demand. They are fully flexible in number and names as to serve the specific representation of the commodity demand, supply, and supply cost profile in each energy sector. Each time slice is independent in terms of the number of represent hours, as long as it is meaningful for the users and their data inputs. 1 is the minimum number of time-slice as this would correspond to a full year. The time-slice definition of a sector affects the commodity price profile and the supply cost profile.

The csv file for the time-slice definition would report the length (in hours) of each time slice as characteristic to the selected sector to represent diurnal, weekly and seasonal variation of energy commodities, demand and supply, as shown in the table for 30 time-slices.

Table 6: Time-slices

AgLevel	SN	Month	Day	Hour	RepresentHours
Hour	1	Winter	Weekday	Night	396
Hour	2	Winter	Weekday	Morning	396
Hour	3	Winter	Weekday	Afternoon	264
Hour	4	Winter	Weekday	EarlyPeak	66
Hour	5	Winter	Weekday	LatePeak	66
Hour	6	Winter	Weekday	Evening	396
Hour	7	Winter	Weekend	Night	156
Hour	8	Winter	Weekend	Morning	156
Hour	9	Winter	Weekend	Afternoon	156
Hour	10	Winter	Weekend	Evening	156
Hour	11	SpringAutumn	Weekday	Night	792
Hour	12	SpringAutumn	Weekday	Morning	792
Hour	13	SpringAutumn	Weekday	Afternoon	528
Hour	14	SpringAutumn	Weekday	EarlyPeak	132
Hour	15	SpringAutumn	Weekday	LatePeak	132
Hour	16	SpringAutumn	Weekday	Evening	792
Hour	17	SpringAutumn	Weekend	Night	300
Hour	18	SpringAutumn	Weekend	Morning	300
Hour	19	SpringAutumn	Weekend	Afternoon	300
Hour	20	SpringAutumn	Weekend	Evening	300
Hour	21	Summer	Weekday	Night	396
Hour	22	Summer	Weekday	Morning	396
Hour	23	Summer	Weekday	Afternoon	264
Hour	24	Summer	Weekday	EarlyPeak	66
Hour	25	Summer	Weekday	LatePeak	66
Hour	26	Summer	Weekday	Evening	396
Hour	27	Summer	Weekend	Night	150
Hour	28	Summer	Weekend	Morning	150

continues on next page

Table 6 – continued from previous page

AgLevel	SN	Month	Day	Hour	RepresentHours
Hour	29	Summer	Weekend	Afternoon	150
Hour	30	Summer	Weekend	Evening	150

It reports the aggregation level of the sector time-slices (AgLevel), slice number (SN), seasonal time slices (Month), weekly time slices (Day), hourly profile (Hour), the amount of hours associated to each time slice (RepresentHours).

6.3.6 Input Commodities

Input commodities are the commodities consumed (also called consumables in MUSE) by each technology. They are defined in a csv file which describes the commodity inputs to each technology, calculated per unit of technology activity. See *below* for a description.

6.3.7 Output Commodities

Output commodities are the commodities produced (also called products in MUSE) by each technology. They are defined in a csv file which describes the commodity outputs from each technology, defined per unit of technology activity. Emissions, such as CO2 (produced from fuel combustion and reactions), CH4, N2O, F-gases, can also be accounted for in this file. See *below* for a description.

6.3.8 General features

To illustrate the data required for a generic technology in MUSE, the *electric boiler technology* is used as an example. The commodity flow for the electric boiler, capable to cover space heating and water heating energy service demands.

Fig. 1: The table below shows the basic data requirements for a typical technology, the electric boiler.

Technology: Electric Boiler ProcessName: resBoilerElectric	*Values*	*Units (input commodity unit/process activity unit)*
electricity	1.0	GWh/PJ
Output commodity	*Values*	*Unit (output commodity unit/process activity unit)*
Space cooling	0	PJ/PJ
Space heating	0.80	PJ/PJ
Water heating	0.20	PJ/PJ
Appliances	0	PJ/PJ
Lighting	0	PJ/PJ
Cooking	0	PJ/PJ
CO_2 from reaction	0	kt/PJ
CO_2 from combustion	0	kt/PJ
CO_2 captured	0	kt/PJ
CO_2 stored	0	kt/PJ
CH_4	0	kt/PJ
N_2O	0	kt/PJ
F-gases	0	kt/PJ

Below it is shown the generic structure of the input commodity file for the electric heater.

Table 7: Commodities used as consumables - Input commodities

ProcessName	RegionName	Time	Level	electricity
Unit	•	Year	•	GWh/PJ
resBoilerElectric	region1	2010	fixed	300
resBoilerElectric	region1	2030	fixed	290

ProcessName represents the technology ID and needs to be consistent across all the data inputs.

RegionName represents the region ID and needs to be consistent across all the data inputs.

Time represents the period of the simulation to which the value applies; it needs to contain at least the base year of the simulation.

Level characterises either a fixed or a flexible input type the following columns should contain the list of commodities the row.

Unit reports the unit in which the technology consumption is defined; it is for the user internal reference only.

The same structure for the csv file would also apply for the output commodity file. The input data has to be provided for the base year. Additional years within the time framework of the overall simulation can be defined. In this case, MUSE would interpolate the values between the provided periods and assume a constant value afterwards.

6.3.9 Existing Sectoral Capacity

For each technology, the decommissioning profile should be given to MUSE.

The csv file which provides the installed capacity in base year and the decommissioning profile in the future periods for each technology in a sector, in each region, should follow the structure reported in the table.

Table 8: Existing capacity of technologies: the residential boiler example

ProcessName	RegionName	Unit	2010	2020	2030	2040	2050
resBoilerElectric	region1	PJ/y	5	0.5	0	0	0
resBoilerElectric	region2	PJ/y	39	3.5	1	0.3	0

ProcessName represents the technology ID and needs to be consistent across all the data inputs.

RegionName represents the region ID and needs to be consistent across all the data inputs.

Unit reports the unit of the technology capacity; it is for the user internal reference only.

2010,. . . , 2050 represent the simulated periods.

6.3.10 Agents

In MUSE, an agent-based formulation was originally introduced for the residential and commercial building sectors **:cite:`2019:sachs`**. Agents are defined using a CSV file, with one agent per row, using a somewhat historical format meant specifically for retrofit and new-capacity agent pairs. This CSV file can be read using `read_csv_agent_parameters()`. The data is also interpreted to some degree in the factory functions `create_retrofit_agent()` and `create_newcapa_agent()`.

For instance, we have the following CSV table:

Name	Type	AgentShare	RegionName	Objective1	SearchRule	DecisionMethod	...
A1	New	Agent5	ASEAN	EAC	all	epsilonCon	...
A4	New	Agent6	ASEAN	CapitalCosts	existing	weightedSum	...
A1	Retrofit	Agent1	ASEAN	efficiency	all	epsilonCon	...
A2	Retrofit	Agent2	ASEAN	Emissions	similar	weightedSum	...

For simplicity, not all columns are included in the example above. Though all column listed below are currently required.

The columns have the following meaning:

Name Name shared by a retrofit and new-capacity agent pair.

Type One of "New" or "Retrofit". "New" and "Retrofit" agents make up a pair with a given *name*. The demand is split into two, with one part coming from decommissioned assets, and the other coming from everything else. "Retrofit" agents invest only to make up for decommissioned assets. They are often limited in the technologies they can consider (by *SearchRule*). "New" agents invest on the rest of the demand, and can often consider more general sets of technologies.

AgentShare Name of the share of the existing capacity assigned to this agent. Only meaningful for retrofit agents. The actual share itself can be found in *Techno-data*.

RegionName Region where an agent operates.

Objective1 First objective that an agent will try and maximize or minimize during investment. This objective should be one registered with `@register_objective`. The following objectives are available with MUSE:

- `comfort`: Comfort provided by a given technology. Comfort does not change during the simulation. It is obtained straightforwardly from *Techno-data*.
- `efficiency`: Efficiency of the technologies. Efficiency does not change during the simulation. It is obtained straightforwardly from *Techno-data*.
- `fixed_costs`: The fixed maintenance costs incurred by a technology. The costs are a function of the capacity required to fulfil the current demand.
- `capital_costs`: The capital cost incurred by a technology. The capital cost does not change during the simulation. It is obtained as a function of parameters found in *Techno-data*.
- `emission_cost`: The costs associated for emissions for a technology. The costs is a function both of the amount produced (equated to the total demand in this case) and of the prices associated with each pollutant. Aliased to "emission" for simplicity.
- `fuel_consumption_cost`: Costs of the fuels for each technology, where each technology is used to fulfil the whole demand.
- `lifetime_levelized_cost_of_energy`: LCOE over the lifetime of a technology. Aliased to "LCOE" for simplicity.
- `net_present_value`: Present value of all the costs of installing and operating a technology, minus its revenues, of the course of its lifetime. Aliased to "NPV" for simplicity.
- `equivalent_annual_cost`: Annualized form of the net present value. Aliased to "EAC" for simplicity.

The weight associated with this objective can be changed using *ObjData1*. Whether the objective should be minimized or maximized depends on *Objsort1*. Multiple objectives are combined using the *DecisionMethod*

Objective2 Second objective. See *Objective1*.

Objective3: Third objective. See *Objective1*.

ObjData1 A weight associated with the *first objective*. Whether it is used will depend in large part on the *decision method*.

ObjData2 A weight associated with the *second objective*. See *ObjData1*.

ObjData3 A weight associated with the *third objective*. See *ObjData1*.

Objsort1 Whether to maximize (*True*) or minimize (*False*) the *first objective*.

Objsort2 Whether to maximize (*True*) or minimize (*False*) the *second objective*.

Objsort3 Whether to maximize (*True*) or minimize (*False*) the *third objective*.

SearchRule The search rule allows users to par down the search space of technologies to those an agent is likely to consider. The search rule is any function with a given signature, and registered with MUSE via `@register_filter`. The following search rules, defined in `filters`, are available with MUSE:

- `same_enduse`: Only allow technologies that provide the same enduse as the current set of technologies owned by the agent.
- `identity`: Allows all current technologies. E.g. disables filtering. Aliased to "all".
- `similar_technology`: Only allows technologies that have the same type as current crop of technologies in the agent, as determined by "tech_type" in *Techno-data*. Aliased to "similar".
- `same_fuels`: Only allows technologies that consume the same fuels as the current crop of technologies in the agent. Aliased to "fueltype".
- `currently_existing_tech`: Only allows technologies that the agent already owns. Aliased to "existing".

- `currently_referenced_tech`: Only allows technologies that are currently present in the market with non-zero capacity.
- `maturity`: Only allows technologies that have achieved a given market share.

The implementation allows for combining these filters. However, the CSV data format described here does not.

DecisionMethod Decision methods reduce multiple objectives into a single scalar objective per replacement technology. They allow combining several objectives into a single metric through which replacement technologies can be ranked.

Decision methods are any function which follow a given signature and are registered via the decorator `@register_decision`. The following decision methods are available with MUSE, as implemented in `decisions`:

- `mean`: Computes the average across several objectives.
- `weighted_sum`: Computes a weighted average across several objectives.
- `lexical_comparion`: Compares objectives using a binned lexical comparison operator. Aliased to "lexo".
- `retro_lexical_comparion`: A binned lexical comparison function where the bin size is adjusted to ensure the current crop of technologies are competitive. Aliased to "retro_lexo".
- `epsilon_constraints`: A comparison method which ensures that first selects technologies following constraints on objectives 2 and higher, before actually ranking them using objective 1. Aliased to "epsilon" ad "epsilon_con".
- `retro_epsilon_constraints`: A variation on epsilon constraints which ensures that the current crop of technologies are not deselected by the constraints. Aliased to "retro_epsilon".
- `single_objective`: A decision method to allow ranking via a single objective.

The functions allow for any number of objectives. However, the format described here allows only for three.

Quantity A factor used to determine the demand share of "New" agents.

MaturityThreshold Parameter for the search rule `maturity`.

6.3.11 Indices and tables

- genindex
- modindex
- search

6.4 Indices and tables

- genindex
- modindex
- search

CHAPTER

SEVEN

ADVANCED DEVELOPER GUIDE

7.1 Hooks

This is where we explain hooks.

7.2 Indices and tables

- genindex
- modindex
- search

CHAPTER

EIGHT

INDICES AND TABLES

- genindex
- modindex
- search