

# NETPLAN V2.0.1 User Manual

Eduardo Ibanez

April 27, 2011

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Setting up</b>	<b>2</b>
2.1	Accessing Cplex . . . . .	2
2.2	Setting the default folder . . . . .	3
2.3	NETPLAN folder . . . . .	3
<b>3</b>	<b>Basic commands</b>	<b>4</b>
<b>4</b>	<b>Constructing a model</b>	<b>5</b>
4.1	parameters.csv . . . . .	5
4.2	node_List.csv . . . . .	6
4.3	arcs_List.csv . . . . .	7
4.4	trans_List.csv . . . . .	7
4.5	Simplification for data storage . . . . .	8
4.6	nodes_[Parameter].csv . . . . .	9
4.7	arcs_[Parameter].csv . . . . .	9
4.8	arcs_TransEnergy.csv . . . . .	10
4.9	events/CapacityLossN.csv . . . . .	10
<b>5</b>	<b>Modeling Approach</b>	<b>10</b>
5.1	Energy systems modeling . . . . .	11
5.2	Transportation systems modeling . . . . .	11
<b>6</b>	<b>General cost minimization formulation</b>	<b>12</b>
6.1	Nomenclature . . . . .	13
6.1.1	Decision Variables . . . . .	13
6.2	Sets and networks . . . . .	13
6.3	Time and time steps . . . . .	13
6.4	Energy Parameters . . . . .	14
6.5	Transportation Parameters . . . . .	14
6.6	Auxiliary parameters . . . . .	15
6.7	Other parameters and functions . . . . .	15
<b>7</b>	<b>General formulation</b>	<b>15</b>
7.1	Compact notation . . . . .	17
7.2	Multiobjective solver . . . . .	17

# 1 Introduction

This document serves as an introduction to the use of NETPLAN, including how to download, access and execute the software. It is intended to cover the basic ideas from a user point of view. For more information you may visit the NETPLAN website<sup>1</sup>.

Whenever possible, the connection between then model and the implementation for variables and parameters is included. As always, I appreciate your feedback and help to correct mistakes or to make the document easier to understand.

## 2 Setting up

Please download the code corresponding to latest version of NETPLAN from the NETPLAN website. Extract its content to a folder called `netplan` in your ISU “My documents”. We will use these files throughout this document.

You may download one of the example “data” folders or create your own with the help of this manual instructions.

### 2.1 Accesing Cplex

CPLEX resides on several ISU servers, which are listed in the following URL, under the Linux categories: <http://it.eng.iastate.edu/remote>. To access any of them, you will need a telnet facility. I suggest using PuTTY, which is a free implementation of Telnet and SSH for Win32 and Unix platforms, along with an xterm terminal emulator.<sup>2</sup>

The following steps are necessary to connect to those servers:

1. Download PuTTY from <http://clue.eng.iastate.edu/downloads.shtml>. Alternatively, the official PuTTY page is <http://www.chiark.greenend.org.uk/~sgtatham/putty/>. At this page, you will find some alternatives; I used (successfully) the installer, `putty-0.60-installer.exe`.
2. Run PuTTY and get the window shown in Fig. 1. Input the name of the server in the Host Name, e.g., `linux-8.ece.iastate.edu`.

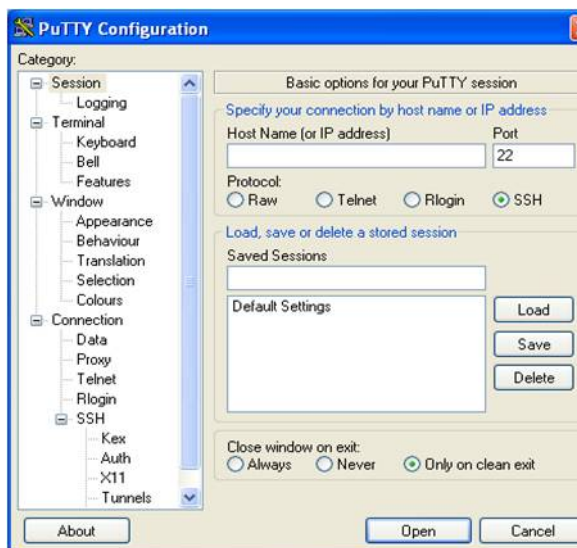


Figure 1: Running PuTTY

<sup>1</sup><http://github.com/eibanez/NETPLAN>

<sup>2</sup>This section has been adapted from Dr. James McCalley’s class notes

3. Use your ISU username and password to log in. You will find yourself on a UNIX terminal emulator.

## 2.2 Setting the default folder

Once you are logged in the server, use the following commands to access your **netplan** folder that you previously created:

```
cd YOUR_USERNAME
cd My\ Documents
cd netplan
```

## 2.3 NETPLAN folder

When you open the **netplan** folder you will find these folders and files:

- **data**: Folder that stores the model that is being tested
- **data/events**: Folder with the definition of events for resiliency analysis
- **manual**: This manual and the files used to produce it
- **src**: Folder with NETPLAN source code
- **src/nsga2**: Folder with NSGA-II implementation on a single machine
- **src/nsga2b**: Folder with parallelization code for NSGA-II (under development)
- **CHANGELOG**: Text file to keep track of latest changes
- **Makefile**: File used to simplify the compilation procedure

During the execution, two more folders called **prepdata** and **nsgadata** will be created. The first one is used to store auxiliary files used during the optimization and most of the results. The latter will contain the results from the NSGA-II multiobjective optimizations.

The source code in the **src** folder is divided into different modules. The main files (programs that you run) are the following:

- **preprocessor.cpp**: Takes information from **data** folder and creates MPS and temporary files
- **postprocessor.cpp**: Takes MPS files and solves problem, writes solution in files

The source code in the **src** folder is divided into different stages, as presented in Fig. 2. The main files (programs that you run) are the following:

- **preprocessor.cpp** (stage 1): Takes information from **data** folder and creates MPS and temporary files
- **postprocessor.cpp**: Takes MPS files and solves problem, writes solution in files (for checking purposes)
- **nsga2/main.cpp** (stage 2): Main file for the NSGA-II implementation, which takes the MPS and auxiliary files and solve the multiobjective problem
- **postnsga.cpp** (stage 3): Reads the individuals that form the Pareto front of solution and reports the solutions.

The rest of the files are considered libraries and contain auxiliary code for assist the programs above:

- **read.cpp**: Functions to read files (global parameters, networks, properties, etc.)
- **write.cpp**: Writing files (solutions, temporary files, etc.)

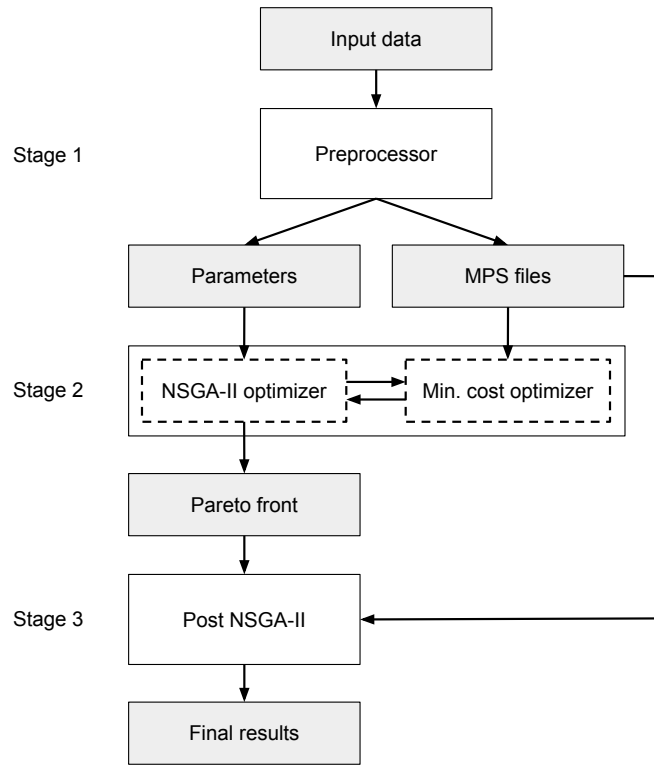


Figure 2: Software structure in NETPLAN

- `global.cpp`: Common global definitions
- `netscore.h`: Common tasks and misc.
- `node.cpp`: Declares a special class to store node information, along with functions to read, modify and write node data
- `arc.cpp`: Similar but for arcs
- `step.cpp`: Functions related with time and time steps
- `index.cpp`: Special variables and functions to store order of arcs and nodes (similar to the a vector index, hence the name)
- `solver.cpp`: Functions to solve a problem (includes Benders decompositions)

### 3 Basic commands

This is a list of the basic commands that are used to compile and execute the different functions in NETPLAN:

- `make`: Compiles the source code
- `./prep` (stage 1): Runs the preprocessor, taking the information in the `data` folder and creating the optimization model
- `./post`: Evaluates the minimum cost problem after it has been created
- `./nsga2` (stage 2): Runs the full multiobjective optimization

- `./postnsga` (stage 3): Evaluates the Pareto front solutions
- `make clean`: Eliminates the compiled programs (useful to clean the folder back to the original state)
- `make cleanmps`: Eliminates results and auxiliary files

## 4 Constructing a model

This will be a brief description of how to construct a functioning model. There is a small example in the website that might be helpful to visualize how it is done. References to the different formulation parameters are included.

All the data manipulation is done inside the `data` folder. The following subsections will explain the content of the different files, which use the CSV<sup>3</sup> format. It is recommended using plain-text editors (such as Notepad on Windows) to modify `parameters.csv`. Any other file can be edited with Notepad or Excel 2007 or newer.

### 4.1 `parameters.csv`

This is the most important file where the global parameters of the optimization are defined. It consists of two columns without a header. The first column is used for keywords and the second for values. The following is a list of the keywords, acceptable inputs or ranges and default values. The last few parameters are specific to NSGA-II.

Some values can be defined here and will be used as defaults. The philosophy is to take the most specific first and then try more relaxed definitions.

Parameters are listed as **keyword** [range] **default**

- **StepName** [letters] **None**: Definition of the key letters that will be used to describe time steps. It is required and it must have one or more letters. E.g., “ym” would indicate that we are using years and months.
- **StepLength** [letters and numbers] **None**: Definition of the maximum length of the time steps (also required). E.g., “y2m12” indicates that simulation is 2 years long and each year has 12 months.
- **StepHours** [number] **1**: Defines the length in hours the last member in the time step (“m” in the example). If it is used once all time steps will be the same. It can also be used multiple times to define different time lengths, e.g., to describe different segments on the load duration curve.
- **DefStep** [letters as in StepName] **None**: Defines what is the default time step level for arcs and nodes. E.g., “ym”
- **DefDiscount** [0–1] **0**: Default discount rate ( $r$  in the formulation) that will be applied to all nodes and arcs, unless otherwise specified.
- **DefInflation** [0–1] **0**: Similar to discount but this reflects price inflation.
- **DefDemandRate** [number] **0**: Used to define the rate at which demand grows globally. E.g., “0.02” means that demands grow 2% annually by default.
- **UseDCFlow** [true/false] **false**: Use DC power flow equations.
- **CodeDC** [two letters] **None**: Define two letter code to identify nodes that use DC power flow. E.g., “EL”.
- **UseBenders** [true/false] **false**: Use Benders decomposition to solve minimum cost problem.
- **OutputLevel** [0–2] **2**: Level of output on screen (0 for most information).

---

<sup>3</sup>Comma separated values

- **TransStep** [letters as in StepName] **None**: Default transportation step. E.g., “y” means that all transportation is represented on an annual basis.
- **TransInfra** [letters] —: The first letter represents a new transportation infrastructure. The rest are the different modes that can use that infrastructure. E.g., “rt” adds infrastructure railroad and indicated that t (trains) can use railroad. This command should be used as many time as transportation infrastructures considered.
- **TransComm** [letters] —: The first letter defines a new commodity. The rest define what modes can transport the current commodity. E.g., “1t” indicates that commodity 1 (coal type 1) can travel by train. This command should be repeated for each commodity.
- **TransCoal** [letters] **None**: The letters correspond to the commodity codes that represent coal, i.e., the members of  $\mathcal{K}_c$ . The command is to be used once.
- **AddObj** [letters and numbers] —: This command adds a new sustainability objective. E.g., “emCO2” for CO<sub>2</sub> emissions. It can be repeated.
- **AddMetric** [letters and numbers] —: Similar to the above but it creates a metric that is not minimized, just reported.
- **NumberEvents** [integer] **0**: Declares the number of events to be used for resiliency. E.g., “2”.
- **popsize** [multiple of 4] **20**: Population size for a generation (NSGA-II).
- **ngen** [number] **200**: Number of generations (NSGA-II).
- **pcross\_real** [0–1] **0.75**: Crossover probability for real variables (NSGA-II).
- **pmut\_real** [0–1] **0.4**: Mutation probability for real variables (NSGA-II).
- **eta\_c** [number] **7**: Distribution index for real variable SBX crossover (NSGA-II).
- **eta\_m** [number] **20**: Distribution index for real variable polynomial mutation (NSGA-II).
- **pcross\_bin** [0–1] **0.4**: Crossover probability for binary variables (NSGA-II).
- **pmut\_bin** [0–1] **0.5**: Mutation probability for binary variables (NSGA-II).
- **stages** [integer] **2**: Number of bits used to represent an investment (NSGA-II).
- **pstart** [0–1] **0.5**: Probability that each of the bits takes the value of 1 for the first randomly generation (NSGA-II).

## 4.2 node\_List.csv

This file defines the set of energy nodes,  $\mathcal{N}^E$ . Nodes are named the following way: “CMIAy1m1”, where the characters represent:

- First character is energy network: coal, electricity, etc.
- Second character is node type: Coal steam generator, storage, transmission
- Third and fourth are location: Iowa, Northwest Power Pool Area, etc.
- The rest refer to time step: month 1 of year 1, year 2, etc.

In this file we define the first four characters, i.e. network, type and location (e.g., “CMIA”). NETPLAN takes care of adding the time information. The first line of the file is reserved for the header. Starting on the second line, each line is occupied by the definition of a node. Of course, each node code needs to be unique. The first lines would go like the following (note that anything after the % is considered a comment and it’s omitted).

```

code
% Coal production nodes
CPAL
CPAZ
CPAR
CPCO
...

```

### 4.3 arcs\_List.csv

An arc is simply determined by the origin and destination nodes with a separation character, in the following fashion: “ETIAy1m1\_ETMNY1m1”. In this example we would be representing the electric transmission from Iowa to Minnesota for the first month of the first year.

Similar to the previous file, the first line is reserved for the header. After that, each line defines the four character code for origin and destination node in the first and second columns, respectively. For example:

```

from,to
% Coal production to transportation node,
CPAL,2TAL
CPAZ,4TAZ
CPAR,3TAR
...

```

### 4.4 trans\_List.csv

This file introduces the transportation variables,  $f_{(i,j,k,m)}(t)$ : flow of transportation arc from node  $i$  to node  $j$  for commodity  $k$  using transportation mode  $m$  during time step  $t$  are coded like this: “ttMONEy1\_1TMONEy1”:

- In the first half, the first two letters are the same and represent the transportation mode (train in this case)
- The next two characters represent the origin (here is Missouri)
- Destination code is defined in the next two characters (Nebraska in this case)
- The remaining characters before the underscore (.) represent time step (year 1)
- The second half begins with the commodity character (1 represents coal)
- The next character is always a “T” to represent transportation.
- The remaining characters are repeated from the first part (origin, destination, time).

NETPLAN uses the definition of commodities, infrastructures and modes from the global parameters to simplify the definition of the transportation network. This file starts with a line reserved for the header and then each line is composed by four columns:

- Origin two character code
- Destination two character code
- Distance in miles
- Characters to represent fleets allowed on that connection (if left blank, all fleets are used for that link)

This is an example:

```

from,to,mileage,fleet
AL,FL,591.5,
AL,GA,182,t
AL,MS,207.5,k
...

```

The transportation network is then coded into arcs and nodes, compatible with the energy network definition. In the example above a node called “1TMONEy1” would be created. It is this node that should be used to enter the transportation load for commodity “1” between “MO” and “NE”.

## 4.5 Simplification for data storage

Before we continue, it is worth describing how data is stored in the rest of the files. This system is aimed at reducing the number of entries so that new cases can be modified quickly.

Let’s say for example that we want to establish the efficiency of all transmission lines to 98%, for all locations and all time steps. In that case we could assign a value of 0.98 for efficiency using the following entry.

```
EL_EL    0.98
```

Using this simplified method we reduce the number of necessary data entries (and thus the chances of making errors) and it is much easier to modify later. The code I am preparing is capable of extending the simple above to all the arcs were it is required.

Now assume that we want to say that the wind generation units in Texas have a cost of 0.2, for all time steps. The necessary entry would be:

```
EWTX_ETTX    0.2
```

When dealing with time, I have been using the following scheme:

	const	y1	y2	...	y1m1	y1m2	...	y2m1	...
ET_ET	0.98	0.95		...		0.87	...		...

The above is a table of possible values of efficiencies for all electric transmission lines for different times. The way is setup allows minimizing the number of entries with respect to time steps. Note that there are empty cells in the table above. The steps to follow when you read the table are:

- Check if there is a value available in the smallest time step.
  - If such a value exists, take it. Example: 0.87 for y1m2.
  - If it doesn’t, continue. Example: y1m1
- Check the following time step that includes the previous one:
  - If there is a value, take it. Example: 0.95 for y1
  - If there isn’t, continue. Example: y2
- Continue until reaching the constant value.

In the case above, some values would be the following:

- y1m1 = 0.95
- y1m2 = 0.87
- y2m1 = 0.98



## 4.6 nodes\_[Parameter].csv

All the remaining files that follow this nomenclature are used to define the rest of parameters in the model for nodes. This list presents these parameters, where the **keyword** is part of the file name (e.g., `nodes\_Demand.csv`):

- **Step** [characters in StepName] **DefStep**: Time step for a particular node
- **Demand** [number/X] **0**: Fixed demand at the node ( $d_j^E(t)$  or  $d_{(i,j,k)}^T(t)$ ). If “X” is used, demand equations 3 and 6 are not enforced.
- **DemandPower** [number/X] **X**: Definition of Demand in terms of power. If “X” is used, demand defaults to the previous parameter
- **DemandRate** [number] **DefDemandRate**: This parameter can be used to simulate a fixed annual increase in the demand.
- **CostUD** [number/X] **X**: If different than “X”, a variable representing unserved demand is introduced and the cost is a penalty for not serving that demand.
- **DiscountRate** [number] **DefDiscountRate**: Discount rate to be applied to **CostUD**.
- **InflationRate** [number] **DefInflationRate**: Inflation rate to be applied to **CostUD**.
- **PeakPower** [number/X] **X**: If different than “X”, constraints are created so that meeting peak power is enforced.
- **PeakPowerRate** [number] **0**: Annual rate to be applied to peak power.

## 4.7 arcs\_[Parameter].csv

Similarly, the remaining parameters for arcs are defined here. This is a list, including the **keyword** that needs to be used in the file name.

- **OpCost** [number] **0**: Operational cost for an arc.
- **InvCost** [number/X] **X**: If different than “X”, investment is allowed on the arcs, using this investment cost.
- **DiscountRate** [number] **DefDiscountRate**: Discount rate to be applied to operational and investment costs.
- **InflationRate** [number] **DefInflationRate**: Inflation rate to be applied to operational and investment costs.
- **Distance** [number/X] **X**: If different than “X”, it becomes a multiplier of costs. Thus, the input of costs is in per mile terms.
- **Eff** [number] **1**: Arc efficiency
- **OpMin** [number] **0**: Minimum operational flow.
- **OpMax** [number/Inf] : Capacity at the arc due to existing infrastructure at time 0. No capacity is enforced if defined as “Inf”.
- **InvMin** [number] **0**: Minimum investment on the arc.
- **InvMax** [number/Inf] **Inf**: Maximum investment on the arc.
- **InvStart** [time step] **First time step**: Defines when the first investment is allowed.
- **LifeSpan** [time step/X] **X**: If different than “X” investments have a predetermined life span.

- **Suscep** [number/X] **X**: If different than “X”, the arc is endowed with susceptance and included in the DC power flow equations.
- **CapacityFactor** [number] **0**: If bigger than 0, the arc is taken into account to meet peak demand at the destination node.
- **OpSUST** [number] **0**: Coefficient to add the current arc to the calculation of a sustainability objective or metric. Substitute “SUST” by the appropriate code used in the parameter file.

#### 4.8 arcs\_TransEnergy.csv

This file defines the fuel consumption of fleet,  $fuelCons_{(i,j,m)}(t)$ , which creates extra demand on the energy side. The first row is reserved as a header. After that, one must write blocks of rows corresponding to the fuel consumption. Thus, a transportation link can create demand in multiple energy nodes.

The first column is matched to the first portion of the transportation arc. The second is used to define the four character code of the energy node to add demand. In the example below, the train (represented by the character **t**) creates a demand of 170.5 MMBtu per million ton-mile transported in nodes “DT01” and “DT02”, that is diesel nodes at location “01” and “02”.

```
from,to,const
tt0102,DT01,170.5 % MMBtu/MMton-mile
tt0102,DT02,170.5
kk0102,DT01,1678.5
kk0102,DT02,1678.5
```

#### 4.9 events/CapacityLossN.csv

The files in the **events** folder are used to define the events to calculate resiliency metrics. Substitute “N” by the appropriate event number (from 1 up to the maximum defined in the global parameters) in the file name.

Within these files we define the fraction of capacity that is available due to a contingency. If not explicitly defined, NETPLAN will use the default of 1 and assume that the capacity was not lost. For example, the entry below indicates that pulverized coal generators are available at 70% during the second year.

```
from,to,y2
EC,,0.7
```

The software automatically identifies which are the operational years affected for each event in order to perform the minimum number of calculations that are necessary.

## 5 Modeling Approach

The energy system is comprised of (but not limited to) electricity, natural gas, liquid fuels, nuclear, biomass, hydroelectric, wind, solar, and geothermal resources. Modeling of national freight and passenger transportation focuses on state-to-state travel; we consider both infrastructures (rail, highways, locks/dams, roads, ports, airports) and fleets (trains, barges, trucks, personal vehicles, airplanes, etc.), and there may be different kinds of fleets for each mode (e.g., diesel trains and electric trains or conventional and plug-in hybrid electric).

Fig. 3 captures the scope of the modeling effort. The transportation and energy systems interact mainly at two different stages: operation and investment. At the operational level each system needs to satisfy its demand with the existing capacity. However, operation of the two systems, and ultimately investment, are interdependent; while the transportation sector demands energy in the form of fuel, the energy sector requires the movement of raw bulk energy sources (e.g. coal or natural gas for thermal power plants). At the same time, the cost of meeting those reciprocal demands has an impact on final prices for energy and transportation. The ever-growing public need for energy and transportation creates the necessity to invest in

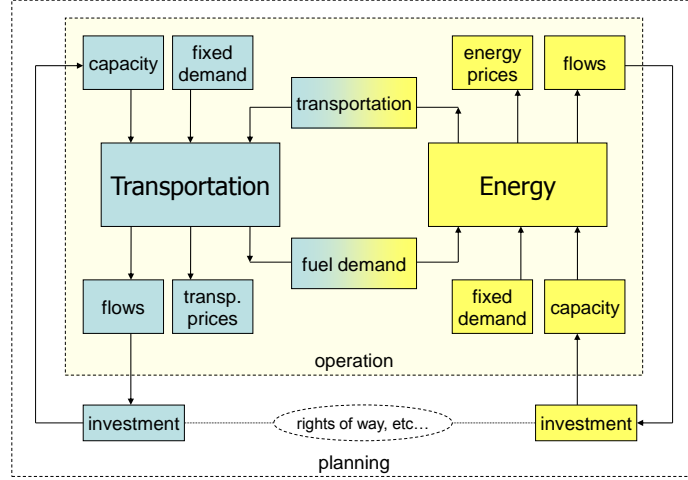


Figure 3: Proposed model that integrates the energy and transportation systems at two levels: operation and planning

new capacity. Given the potential for increased coupling between energy and transportation, it is apparent that better designs of both can be achieved if these designs are performed together.

Our modeling approach consists of two levels. The lower level performs a cost minimization on a 40 year infrastructure investment plan, constrained by lower and upper limits on technology investments, and evaluates resiliency and sustainability metrics. The higher level uses an evolutionary program to perform a multiobjective search for the Pareto optimal front in the space of cost, resiliency, and sustainability metrics, by manipulating the technology investment lower and upper limits. The remainder of this section, together with Section 3, is devoted to describing the lower level. The higher level is described in Section 4.

## 5.1 Energy systems modeling

A generalized network flow transportation model [1–3] is used to model energy systems, where commodity flow is energy, and transportation paths are AC and DC electric transmission, gas pipelines (for natural gas and/or hydrogen), and liquid fuel pipelines (for petroleum-based fuels, biofuels such as ethanol or biodiesel, and anhydrous ammonia). Energy transport by rail, barge, and truck is included in the freight transport model.

Each source node, specified with location, is connected to a fictitious source node that supplies all energy. Arcs emanating from each source are characterized by maximum extraction rate and extraction cost. Petroleum, coal, natural gas, and uranium have finite capacities, while renewables have infinite capacities. All sources have finite maximum extraction rates. Conversion and transportation are endowed with: capacity, efficiency, operational cost, investment cost, lifetime, component sustainability metrics (e.g.,  $\text{CO}_2$ ), and component resiliency (e.g., reliability).

## 5.2 Transportation systems modeling

The freight transport system is modeled as a multicommodity flow network where the flows are in the units of tons of each major commodity. A commodity is major if its transportation requirements comprise at least 2% of the nation’s total freight ton-miles. Data available to make this determination [4] indicates this criterion includes 23 commodities that comprise 90% of total ton-miles (e.g., the top eight, comprising 55%, are in descending order: coal, cereal grains, foodstuffs, gasoline and aviation fuel, chemicals, gravel, wood products, and base metals).

There are two fundamental differences between this formulation and that of the energy formulation. Whereas the energy formulation must restrict energy flows of specific forms to particular networks (for example, natural gas or hydrogen cannot move through electric lines or liquid fuel lines), commodities may be transported over any of the transport modes (rail, barge, truck). Also whereas energy movement requires only infrastructure (electric lines, liquid fuel pipelines, gas pipelines), commodity movement requires infrastructure (rail, locks/dams, roads, ports) and fleet (trains, barges, trucks), and there may be different kinds of fleets for each mode (e.g., diesel trains or electric trains).

To accommodate these differences, the transportation formulation is comprised of two multicommodity flows [5], one embedded inside the other. Commodities flow through the network formed by the different types of fleet available. At the same time, the units in those fleets travel along the network formed by the different infrastructures. An easy way to visualize this is captured in Fig. 4, where the flow from node A to node B is divided according to the types of infrastructures first and then into the different types of available fleets.

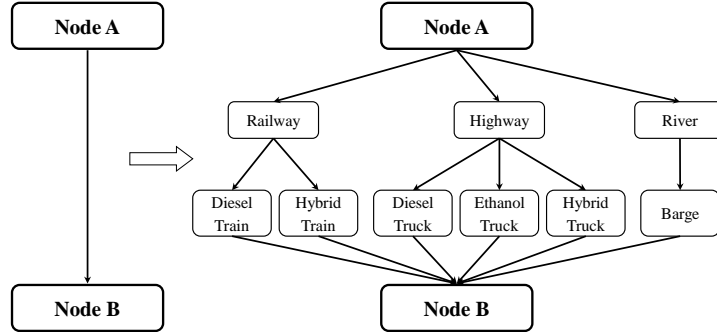


Figure 4: Decomposition of transportation arc in two steps: infrastructure and fleet.

## 6 General cost minimization formulation

The optimization problem associated with this model can be conceptually described by (1),

$$\begin{aligned}
 &\mathbf{min} \quad CostOp + CostInv \\
 &\mathbf{subject\ to:} \\
 &\quad Meet\ energy\ demand, \\
 &\quad Meet\ transportation\ demand, \\
 &\quad Capacity\ constraints \\
 &\quad Power\ flow\ constraints\ on\ electric\ transmission
 \end{aligned} \tag{1}$$

The objective is to minimize the combined energy and a transportation cost with constraints of meeting demands on energy, and freight transport while following the capacity constraints. The operational characteristics of the model provide additional constraints on the system, such as the inclusion of power flow relations for electric transmission lines (we used the so-called “DC” power flow approximation for this purpose).

This section contains a rigorous description of the formulation needed to achieve the characteristics described above. The explanation of the formulation is preceded by an introduction to the nomenclature used.

## 6.1 Nomenclature

### 6.1.1 Decision Variables

$e_{(i,j)}(\mathbf{t})$ : Operational flow of energy arc from node  $i$  to node  $j$ , for time step  $\mathbf{t}$  (MWh)

$eInv_{(i,j)}(\mathbf{t})$ : Capacity investment on energy arc from node  $i$  to node  $j$ , for time step  $\mathbf{t}$ . (MW)

$f_{(i,j,k,m)}(\mathbf{t})$ : Operational flow of transportation arc from node  $i$  to node  $j$  for commodity  $k$  using transportation mode  $m$  during time step  $\mathbf{t}$  (ton)

$fleetInv_{(i,j,m)}(\mathbf{t})$ : Fleet  $m$  capacity investment for transportation arc from node  $i$  to node  $j$  during time step  $\mathbf{t}$  (ton/hour)

$infInv_{(i,j,l)}(\mathbf{t})$ : Infrastructure  $l$  capacity investment for transportation arc from node  $i$  to node  $j$  for time step  $\mathbf{t}$  (ton/hour)

$\theta_i(\mathbf{t})$ : Phase angle at node  $i$ , used to model DC power flow (radians)

## 6.2 Sets and networks

$\mathcal{N}^E$ : Set of energy nodes

$\mathcal{N}_d^E \subset \mathcal{N}^E$ : Subset of energy nodes where demand equations are enforced

$\mathcal{N}_p^E \subset \mathcal{N}^E$ : Subset of energy nodes where peak demand equations are enforced

$\mathcal{A}^E$ : Set of energy arcs

$\mathcal{A}_{DC}^E \subset \mathcal{A}^E$ : Set of AC electric transmission arcs, which satisfy DC power flow equations

$\mathcal{N}^T$ : Set of transportation nodes

$\mathcal{A}^T$ : Set of transportation arcs

$\mathcal{A}_j^T \subset \mathcal{A}^T$ : Subset of transportation arcs that create an energy demand at energy node  $j$

$\mathcal{K}$ : Set of commodities

$\mathcal{K}_e \subset \mathcal{K}$ : Subset of commodities used by the energy system, e.g. coal

$\mathcal{M}$ : Set of fleet or transportation modes

$\mathcal{M}_l \subset \mathcal{M}$ : Subset of fleet that can use transportation infrastructure  $l$

$\mathcal{M}_j \subset \mathcal{M}$ : Subset of fleet that require energy from energy node  $j$

$n_{(i,k)}^E \in \mathcal{N}^E$ : Energy node that corresponds to the geographic location  $i$  and commodity  $k$  in the transportation system

## 6.3 Time and time steps

$\mathcal{T} = [1, \dots, T_1] \times [0, \dots, T_2] \times \dots \times [0, \dots, T_s]$ : Definition of simulation time domain, which is divided in  $s$  time steps. Time step  $i$  is divided into  $T_i$  segments. For instance, simulation time could be divided in “years”, “months” and “days” with the appropriate number of divisions within each step

$\mathbf{t} = [t_1, t_2, \dots, t_s] \in \mathcal{T}$ : Time instance in the simulation domain

$\mathbf{t}_1$ : Value of the top level step for time  $\mathbf{t}$ , i.e.,  $t_1$

$\mathcal{T}_{\mathbf{t}}$ : Time that spans between the beginning of the simulation and time  $\mathbf{t}$

$\Delta(\mathbf{t})$ : Length of time step  $\mathbf{t}$  (h)

## 6.4 Energy Parameters

- $\eta_{(i,j)}(\mathbf{t})$ : Efficiency of arc  $(i, j)$  during time  $\mathbf{t}$  (unitless)
- $lbe_{(i,j)}(\mathbf{t})$ : Lower bound for flow in arc  $(i, j)$  for time  $\mathbf{t}$  (MWh)
- $ube_{(i,j)}(\mathbf{t})$ : Upper bound for flow in arc  $(i, j)$  during time  $\mathbf{t}$  due to the initial existing infrastructure (MW)
- $lbeInv_{(i,j)}(\mathbf{t})$ : Minimum allowed capacity increase in arc  $(i, j)$  at time  $\mathbf{t}$  (MW)
- $ubeInv_{(i,j)}(\mathbf{t})$ : Maximum allowed capacity increase in arc  $(i, j)$  at time  $\mathbf{t}$  (MW)
- $eLife_{(i,j)}$ : Expected lifetime for investments in arc  $(i, j)$  at time  $\mathbf{t}$  (year)
- $costOp_{(i,j)}(\mathbf{t})$ : Operational cost for flow in arc  $(i, j)$  during time  $\mathbf{t}$  (\$/MWh)
- $costInv_{(i,j)}(\mathbf{t})$ : Investment cost for capacity increase in arc  $(i, j)$  in network  $l$ , at time  $\mathbf{t}$  (\$/MW)
- $heatContent_k(\mathbf{t})$ : Heat content of commodity  $k$ , at time  $\mathbf{t}$  (MWh/ton)
- $d_j^E(\mathbf{t})$ : Fixed energy demand at node  $j$  during time  $\mathbf{t}$  (MWh)
- $peakD_j^E(\mathbf{t})$ : Peak demand at electric node  $j$  during time  $\mathbf{t}$  (MWh)
- $cf_{(i,j)}(\mathbf{t})$ : Capacity factor of generation between nodes  $i$  and  $j$  at time  $\mathbf{t}$  (unitless)
- $b_{(i,j)}$ : Susceptance of transmission line between nodes  $i$  and  $j$  (p.u.)

## 6.5 Transportation Parameters

- $ubFleet_{(i,j,m)}(\mathbf{t})$ : Upper bound for total transportation flow for fleet  $m$  in arc  $(i, j)$  during time  $\mathbf{t}$  due to the initial existing fleet (ton/h)
- $lbFleetInv_{(i,j,m)}(\mathbf{t})$ : Minimum allowed capacity increase in arc  $(i, j)$  for fleet  $m$  at time  $\mathbf{t}$  (ton/h)
- $ubFleetInv_{(i,j,m)}(\mathbf{t})$ : Maximum allowed capacity increase in arc  $(i, j)$  for fleet  $m$  at time  $\mathbf{t}$  (ton/h)
- $fleetLife_{(i,j,m)}$ : Expected lifetime of fleet  $m$  in arc  $(i, j)$  (years)
- $ubInf_{(i,j,l)}(\mathbf{t})$ : Upper bound for total transportation flow across infrastructure  $l$  in arc  $(i, j)$  during time  $\mathbf{t}$  due to the initial existing infrastructure (ton/h)
- $lbInfInv_{(i,j,l)}(\mathbf{t})$ : Minimum allowed capacity increase in arc  $(i, j)$  for transportation infrastructure  $l$  at time  $\mathbf{t}$  (ton/h)
- $ubInfInv_{(i,j,l)}(\mathbf{t})$ : Maximum allowed capacity increase in arc  $(i, j)$  for transportation infrastructure  $l$  at time  $\mathbf{t}$  (ton/h)
- $infLife_{(i,j,k,m)}$ : Expected lifetime for infrastructure investment  $l$  in arc  $(i, j)$  (years)
- $costOp_{(i,j,k,m)}(\mathbf{t})$ : Operational cost for transportation flow in arc  $(i, j)$  in fleet  $m$ , commodity  $k$  and time  $\mathbf{t}$  (\$/ton)
- $costFleetInv_{(i,j,m)}(\mathbf{t})$ : Investment cost for capacity increase in fleet  $m$  for arc  $(i, j)$  and time  $\mathbf{t}$  (\$ h/ton)
- $costInfInv_{(i,j,l)}(\mathbf{t})$ : Investment cost for capacity increase in transportation infrastructure  $l$  for arc  $(i, j)$  and time  $\mathbf{t}$  (\$ h/ton)
- $fuelCons_{(i,j,m)}(\mathbf{t})$ : Fuel consumption for transportation mode  $m$  for transportation arc  $(i, j)$  during time step  $\mathbf{t}$  (MWh/ton)
- $d_{(i,j,k)}^T(\mathbf{t})$ : Fixed transportation demand of commodity  $k$  for arc  $(i, j)$  during time  $\mathbf{t}$  (ton)

## 6.6 Auxiliary parameters

These parameters are calculated as a combination of decision variables and predetermined parameters.

$CostOp^E$ : Operational cost from the energy system (\$)

$CostInv^E$ : Cost due to the investment upgrades on the energy system (\$)

$CostOp^T$ : Operational cost for transportation system (\$)

$CostFleetInv^T$ : Investment cost on transportation fleet (\$)

$CostInfInv^T$ : Cost on transportation infrastructure (\$)

$eCap_{(i,j)}(\mathbf{t})$ : Capacity for energy arc from node  $i$  to node  $j$ , for time step  $\mathbf{t}$ . (MW)

$fleetCap_{(i,j,m)}(\mathbf{t})$ : Fleet  $m$  capacity for transportation arc from node  $i$  to node  $j$  during time step  $\mathbf{t}$  (ton/h)

$infCap_{(i,j,l)}(\mathbf{t})$ : Infrastructure  $l$  capacity for transportation arc from node  $i$  to node  $j$  for time step  $\mathbf{t}$  (ton/h)

$d_j^{ET}(\mathbf{t})$ : Energy demand at node  $j$  during time  $\mathbf{t}$  due to the transportation of commodities (MWh)

## 6.7 Other parameters and functions

$r$ : Discount rate

$P_E$ : Power base for the electric system

$I[test]$ : Indicator function that equals 1 if  $test$  is true and 0 otherwise

## 7 General formulation

The following formulation formally incorporates the modeling capabilities that have been described in previous chapters and allow the joint optimization of the transportation and energy networks.

The objective to be minimized (2) is the total cost of operating and investment of the combined system. We can distinguish five components in the objective: operational cost of the energy system (16), investment cost of the energy system (17), operational cost of the transportation system (18), and investment cost of the transportation system divided into fleet investment (19) and infrastructure investment (20).

The energy sector is represented by a set of nodes  $\mathcal{N}^E$  and arcs  $\mathcal{A}^E$ . Each node represents an energy subnetwork in a geographic location. For example, for a particular location, there might be three energy nodes; one for electricity, one for gas, and one for petroleum. Energy arcs link these various nodes, both within the same subnetwork (representing transmission lines, natural gas pipelines, or petroleum pipelines) or between different networks, where conversion takes place (e.g., power plants).

$$\min \quad \{CostOp^E + CostInv^E + CostOp^T + CostFleetInv^T + CostInfInv^T\} \quad (2)$$

**subject to:**

Meet energy demand at the appropriate nodes

$$\sum_i \eta_{(i,j)}(\mathbf{t}) e_{(i,j)}(\mathbf{t}) - \sum_i e_{(j,i)}(\mathbf{t}) = d_j^E(\mathbf{t}) + d_j^{ET}(\mathbf{t}) \quad j \in \mathcal{N}_d^E \quad (3)$$

DC power flow equations

$$e_{(i,j)}(\mathbf{t}) - e_{(j,i)}(\mathbf{t}) = b_{(i,j)} \left( \theta_i(\mathbf{t}) - \theta_j(\mathbf{t}) \right) P_E \Delta(\mathbf{t}), \quad (i,j) \in \mathcal{A}_{DC}^E \quad (4)$$

Generation capacity must cover peak demand at electric nodes

$$\sum_i cf_{(i,j)}(\mathbf{t}) eCap_{(i,j)}(\mathbf{t}) \geq peakD_j^E(\mathbf{t}), \quad j \in \mathcal{N}_p^E \quad (5)$$

Transportation demand for non-energy commodities

$$\sum_m f_{(i,j,k,m)}(\mathbf{t}) = d_{(i,j,k)}^T(\mathbf{t}), \quad k \in \mathcal{K} \setminus \mathcal{K}_e \quad (6)$$

Transportation demand for energy commodities

$$\sum_m f_{(i,j,k,m)}(\mathbf{t}) = \text{heatContent}_k^{-1}(\mathbf{t}) e_{(n_{(i,k)}^E, n_{(j,k)}^E)}(\mathbf{t}), \quad k \in \mathcal{K}_e \quad (7)$$

Fleet upper bound for transportation flows

$$\sum_k f_{(i,j,k,m)}(\mathbf{t}) \leq \text{fleetCap}_{(i,j,m)}(\mathbf{t}) \Delta(\mathbf{t}) \quad (8)$$

Infrastructure upper bound for transportation flows

$$\sum_k \sum_{m \in \mathcal{M}_l} f_{(i,j,k,m)}(\mathbf{t}) \leq \text{infCap}_{(i,j,l)}(\mathbf{t}) \Delta(\mathbf{t}) \quad (9)$$

**Decision variables:**

$$\text{Energy flows:} \quad 0 \leq \text{lbe}_{(i,j)}(\mathbf{t}) \leq e_{(i,j)}(\mathbf{t}) \leq e\text{Cap}_{(i,j)}(\mathbf{t}) \Delta(\mathbf{t}) \quad (10)$$

$$\text{Energy capacity inv.:} \quad \text{lbeInv}_{(i,j)}(\mathbf{t}) \leq e\text{Inv}_{(i,j)}(\mathbf{t}) \leq \text{ubeInv}_{(i,j)}(\mathbf{t}) \quad (11)$$

$$\text{Transportation flows:} \quad f_{(i,j,k,m)}(\mathbf{t}) \geq 0 \quad (12)$$

$$\text{Fleet inv.:} \quad \text{lbFleetInv}_{(i,j,m)}(\mathbf{t}) \leq \text{fleetInv}_{(i,j,m)}(\mathbf{t}) \leq \text{ubFleetInv}_{(i,j,m)}(\mathbf{t}) \quad (13)$$

$$\text{Infrastructure inv.:} \quad \text{lbInfInv}_{(i,j,l)}(\mathbf{t}) \leq \text{infInv}_{(i,j,l)}(\mathbf{t}) \leq \text{ubInfInv}_{(i,j,l)}(\mathbf{t}) \quad (14)$$

$$\text{Phase angles:} \quad -\pi \leq \theta_i(\mathbf{t}) \leq \pi \quad (15)$$

**where:**

$$\text{CostOp}^E = \sum_{\mathbf{t}} \sum_{(i,j)} (1+r)^{-\mathbf{t}_1} \text{costOp}_{(i,j)}^E(\mathbf{t}) e_{(i,j)}(\mathbf{t}) \quad (16)$$

$$\text{CostInv}^E = \sum_{\mathbf{t}} \sum_{(i,j)} (1+r)^{-\mathbf{t}_1} \text{costInv}_{(i,j)}^E(\mathbf{t}) e\text{Inv}_{(i,j)}(\mathbf{t}) \quad (17)$$

$$\text{CostOp}^T = \sum_{\mathbf{t}} \sum_{(i,j,k,m)} (1+r)^{-\mathbf{t}_1} \text{costOp}_{(i,j,k,m)}^T(\mathbf{t}) f_{(i,j,k,m)}(\mathbf{t}) \quad (18)$$

$$\text{CostFleetInv}^T = \sum_{\mathbf{t}} \sum_{(i,j,m)} (1+r)^{-\mathbf{t}_1} \text{costFleetInv}_{(i,j,m)}(\mathbf{t}) \text{fleetInv}_{(i,j,m)}(\mathbf{t}) \quad (19)$$

$$\text{CostInfInv}^T = \sum_{\mathbf{t}} \sum_{(i,j,l)} (1+r)^{-\mathbf{t}_1} \text{costInfInv}_{(i,j,l)}(\mathbf{t}) \text{infInv}_{(i,j,l)}(\mathbf{t}) \quad (20)$$

$$e\text{Cap}_{(i,j)}(\mathbf{t}) = \text{ube}_{(i,j)}(\mathbf{t}) + \sum_{\tau \in \mathcal{T}_{\mathbf{t}}} e\text{Inv}_{(i,j)}(\tau) \mathbb{I}[\mathbf{t}_1 - \tau_1 \leq e\text{Life}_{(i,j)}] \quad (21)$$

$$\text{fleetCap}_{(i,j,m)}(\mathbf{t}) = \text{ubFleet}_{(i,j,m)}(\mathbf{t}) + \sum_{\tau \in \mathcal{T}_{\mathbf{t}}} \text{fleetInv}_{(i,j,m)}(\tau) \mathbb{I}[\mathbf{t}_1 - \tau_1 \leq \text{fleetLife}_{(i,j,m)}] \quad (22)$$

$$\text{infCap}_{(i,j,l)}(\mathbf{t}) = \text{ubInf}_{(i,j,l)}(\mathbf{t}) + \sum_{\tau \in \mathcal{T}_{\mathbf{t}}} \text{infInv}_{(i,j,l)}(\tau) \mathbb{I}[\mathbf{t}_1 - \tau_1 \leq \text{infLife}_{(i,j,k,m)}] \quad (23)$$

$$d_j^{ET}(\mathbf{t}) = \sum_{(a,b) \in \mathcal{A}_j^T} \sum_{m \in \mathcal{M}_j} \text{fuelCons}_{(a,b,m)}(\mathbf{t}) \sum_k f_{(a,b,k,m)}(\mathbf{t}) \quad (24)$$

The flows across these arcs,  $e_{(i,j)}$ , are part of the decision variables of the problem. These flows must be such that they satisfy the demand for energy (3), part of which is due to the fuel required to perform the movement of commodities in the transportation system (24). Energy flows are also required to meet lower and upper bound constraints (10). Upper bounds are determined by the capacity of the existing energy infrastructure,  $e\text{Cap}_{(i,j)}(\mathbf{t})$ , and are a combination (21) of the initial capacity,  $\text{ube}_{(i,j)}(\mathbf{t})$ , and investment on upgrades,  $e\text{Inv}_{(i,j)}(\mathbf{t})$ . The former are given parameters while the latter are also decision variables with



their corresponding lower and upper bounds (11). Energy flows must also satisfy other constraints, such as DC power flow equations for electric transmission (4). Equation (5) ensures that generation for each electric region is able to cover the peak demand.

The transportation system is also formed by a set of nodes  $\mathcal{N}^T$  and a set of arcs  $\mathcal{A}^T$ , although the nomenclature is slightly different. Here each node represents a unique geographic location, while the transportation flows (12) are referred to as  $f_{(i,j,k,m)}$ , where  $(i,j)$  represents the origin and destination nodes,  $k$  the commodity that is being transported and  $m$  the mode of transportation used.

These flows must satisfy the transportation demand, which is predetermined (6) for all commodities except for those that are energy related (e.g., coal, uranium). In that case transportation demands depend on the use of that commodity in the energy system (7). Transportation flows are constrained (8, 9) by the capacity (22) of the available fleet,  $fleetCap_{(i,j,m)}(\mathbf{t})$ , and the capacity (23) transportation infrastructure  $infCap_{(i,j,l)}(\mathbf{t})$ . Both can be increased by their respective investments,  $fleetInv_{(i,j,m)}(\mathbf{t})$  and  $infInv_{(i,j,m)}(\mathbf{t})$ , which have upper and lower bound constraints (13, 14).

## 7.1 Compact notation

A more compact version of the previous formulation can be produced using vectorial and matrix notation. First of all, the vector **capInv** includes all capacity investment variables, while the operational variables are grouped in vectors called **flows<sub>t</sub>**. The subscript  $t$  in this section represents the top level time step which a variable belongs, i.e.,  $\mathbf{t}_1$  following the notation from the previous section.

When using this notation we can reduce the model to the expressions (25).

$$\begin{aligned}
& \min \quad [\text{CostInv}^\top \quad \text{CostOp}_1^\top \quad \text{CostOp}_2^\top \quad \dots] \begin{bmatrix} \text{capInv} \\ \text{flows}_1 \\ \text{flows}_2 \\ \vdots \end{bmatrix} \\
& \text{subject to:} \\
& \begin{bmatrix} -\mathbf{I} & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{A}_1 & \mathbf{0} & \dots \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \dots \\ \mathbf{L}_1 & \mathbf{I} & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_2 & \dots \\ \mathbf{0} & \mathbf{0} & -\mathbf{I} & \dots \\ \mathbf{L}_2 & \mathbf{0} & \mathbf{I} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \text{capInv} \\ \text{flows}_1 \\ \text{flows}_2 \\ \vdots \end{bmatrix} \leq \begin{bmatrix} -\text{lbInv} \\ \text{ubInv} \\ \mathbf{d}_1 \\ -\text{lb}_1 \\ \text{ub}_1 \\ \mathbf{d}_2 \\ -\text{lb}_2 \\ \text{ub}_2 \\ \vdots \end{bmatrix} \tag{25}
\end{aligned}$$

The objective value is (2), which is developed in expressions (16-20). The first two rows in the constraints are directly related to the minimum and maximum investment allowed (11, 13, 14).

Following that there are three rows that are repeated for each year. The first two correspond to operational constraints, such as demand balances (3, 6, 7) or DC power flow constraints (4). Such expressions are summarized in the matrix  $\mathbf{A}_t$  and the vector  $\mathbf{d}_t$ . The next two lines correspond to the minimum and maximum operational flows from (8–10, 21–23). The matrix  $\mathbf{L}_t$  accounts for the fact that only investments done prior to the operational year can account for available capacity.

## 7.2 Multiobjective solver

NETPLAN, an energy and transportation investment multiobjective evolutionary algorithm is proposed as an approach to efficiently solve the problem described above and produce an approximation to the Pareto front. It has been conceived in a modular fashion, to allow the parallel development of each one of its components. At the highest level two levels can be distinguished, as depicted in Fig. 5.

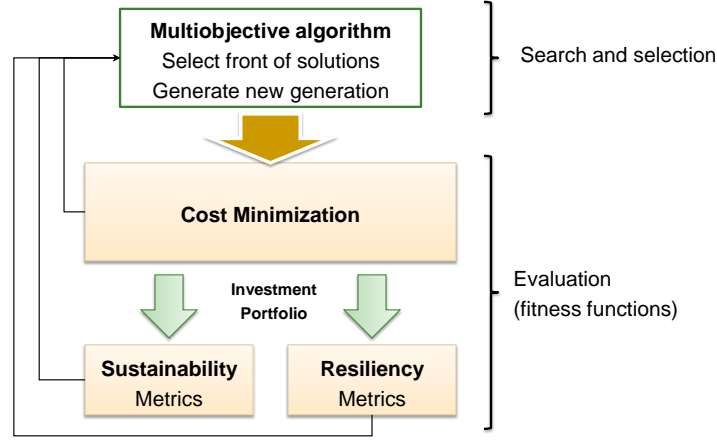


Figure 5: NETPLAN multiobjective approach

The high-level optimization is directed by the NSGA-II algorithm [6]. Each individual within a given population is characterized by a minimum level of investment that is to be enforced. Thus, a vector of minimum values for *capInv* is created for each cost minimization evaluation.

This vector is then passed on to the low-level optimization to be evaluated. This section captures the modeling of the energy and transportation systems, as explained above. First of all, the minimum objective cost problem is solved with the minimum enforced investment constraints incorporated. This process produces the investment portfolio for the individual solutions as well as the flows representing how the system is operated. Other economic factors, such as energy and transportation prices or economic opportunities, can also be extracted.

Once the portfolio is determined, metrics for sustainability and resiliency are calculated. At this stage, those metrics are simple functions based on investments and flows but more complex scenario-based calculations could be implemented using this modular approach.

The objective values are then recovered by the high-level optimization in order to produce the next generation in search for the Pareto front. This loop could also be used to guide the multiobjective optimizer towards a richer variety of solutions. For example, from the dual cost minimization problems one could identify more economic alternatives or, from the evaluation of resiliency, one could identify the weakest links in the system and reinforce them in future solutions. However, including these features would require major modifications in the multiobjective optimizer.

At this point, the use of minimum investment vectors facilitates shifting from the minimum cost solutions to others with better performance in terms of sustainability or resiliency. This approach is very simple to implement but, if needed, could be complemented with other methods such as subsidies for clean alternatives, taxes for polluting technologies (e.g., carbon tax) or preventing certain technologies to be used in favor of others.

## References

- [1] A. Quelhas, E. Gil, J. McCalley, and S. Ryan, "A multiperiod generalized network ow model of the U.S. integrated energy system: Part I - Model description," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 829–836, May 2007.
- [2] A. Quelhas and J. McCalley, "A multiperiod generalized network ow model of the U.S. integrated energy system: Part II - Simulation results," *IEEE Trans. Power Syst.*, vol. 22, no. 3, pp. 837–844, May 2007.
- [3] E. Gil and J. McCalley, "A U.S. energy system model for disruption analysis: Evaluating the effects of 2005 hurricanes," *IEEE Trans. Power Syst.*, vol. PP, no. 99, Jan. 2011.

- [4] U.S. Department of Transportation, Research and Innovative Technology Administration, Bureau of Transportation Statistics. (1997) “Commodity flow survey United States.” [Online]. Available: [http://www.bts.gov/publications/commodity\\_flow\\_survey/](http://www.bts.gov/publications/commodity_flow_survey/)
- [5] M. Bazaraa, J. Jarvis, and H. Sherali, *Linear Programming and Network Flows*, 2nd ed. New York: Wiley, 1990, ch. 12, pp. 587–597.
- [6] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002.