Quantitative Decision Making: Mixed integer linear Programming for Supply Chain Network Design

Fakultät Wirtschaft & Management, Institut für Technologie und Management

31. July 2018

Agenda

1. Prescriptive Analytics
2. Supply Chain Network Design
3. SCDN with Spreadsheets

Predictive Analytics

- *Prescriptive analytics determines actions to take to make the future happen*
- *Prescriptive Analytics uses models to specify optimal behaviors and actions and **automatically** make the actions occur*
- *Prescriptive Analytics determines alternative courses of actions or decisions, given the current and projected situations and a set of objectives, requirements, and constraints*

Past

Present

Future

- What actions could be taken to increase sales?
- What trucks should take which route?
- What products should be offered together?

Source: Das et al (2013); Franks (2014), Davenport (2014)

Methods for Prescriptive Analytics

Linear Programming

- Modelling and determination of the optimal solution to a problem
- Assumption: objective function and constraints are linear

$$\min z = x_1 + x_2 + x_3 + \cdots + x_n$$

- For the objective function, the assumptions of proportionality and additivity applies:
- Proportionality: the decision variables can contribute differently to the result of the target function (decision variables can have different factors)
- Additivity: The decision variables are added exclusively

Integer Programming

- Modelling and determination of the optimal solution to a problem
- Assumption: Decision variables can only take integer values
- *Mixed integer programming* (MIP): Some decision variables are limited to integer values
- *Mixed integer linear programming* (MILP): MIP with linear objective function and constraints

- Due to the integer condition the number of possible solutions becomes finite, instead of the infinity with real numbers
- For linear problems, however, there are algorithms that limit the solutions to a few real numbers. In comparison, the calculation effort increases with integers

Methods for Prescriptive Analytics

Non-linear Programming

- Modelling and determination of the optimal solution to a problem
- Either objective function, constraints or both are non-linear

$$\min z = x_1^p + \frac{1}{x_2} + e^{x_3} + \cdots + \log(x_n)$$

- Differentiation into local and global optima: global is the best solution compared to all possible solutions, local is the best solution in a neighborhood (all solutions around the solution) of solutions.
- In linear problems every local optimum is a global optimum
- In non-linear problems, this does not apply and makes the determination of the solution more difficult.

Goal Programming

- Modelling and determination of the optimal solution to a problem
- Several objectives are pursued simultaneously

- Aspiration levels are set for the different objectives, from which target deviations are calculated (goal deviation).
- The minimization of the target deviation determines the target function.
- A target priority is assigned, which influences the importance of the targets in minimization.

Case: Coca Cola

- Coca Cola produces beverages (soft drinks, sports drinks, juices,...) worldwide with 700,000 employees and a turnover of $US 41.86 mil.
- Among the brands is Minute Maid, under which Orange juice is sold - an end product made from fruits with limited growth periods and variable taste, but should always taste the same.
- In an algorithm, the so-called Black Book, various data (satellite images, weather, expected yield, cost pressure, regional consumer preferences, data on the approximately 600 different flavors of an orange, as well as the sweetness and acidity of oranges) are combined to ensure a constant taste.
- The algorithm contains about 1 trillion decision

- The taste of orange juice is always the same
- With the algorithm, Coca Cola is able to plan up to 15 months ahead and to react quickly and flexibly to changes (e.g. to weather changes that significantly influence crop yields).

Case: Zalando

- Zalando is a European eCommerce retailer specialising in clothing with 12,000 employees and a turnover of €987 million.
- The picking process is a decisive factor in the length of goods at Zalando. This is usually carried out in two stages, whereby order-independent picking is carried out in the first stage from a complete assortment constructed like a rope ladder and in the second stage from the reduced quantity depending on specific customer orders.
- The travel time of the picker in the first stage depends on his route, but the route must be calculated very quickly, since Zalando receives several thousand orders per hour that are to be picked.
- Zalando uses an algorithm that estimates the travel time with a neural network to divide the orders and calculate the resulting route. The algorithm also specifies the stopping points of the order picker's trolley.

- The algorithm estimates the travel time with an average deviation of 32.25 seconds per picker's travel our.
- The travel time of order pickers could be reduced by 11% compared to random order distribution

Summary: Predictive Analytics

- Prescriptive analytics determines alternative courses of action or decision options for a given prevailing and projected situation and a set of goals, requirements and constraints
- Prescriptive Analytics describes the use of various optimization techniques
  - Note: in the literature, simulation is counted among other things as Prescriptive Analytics. In a simulation, various (previously defined) options for action are evaluated. However, the action is not predefined.

Supply Chain Network Design

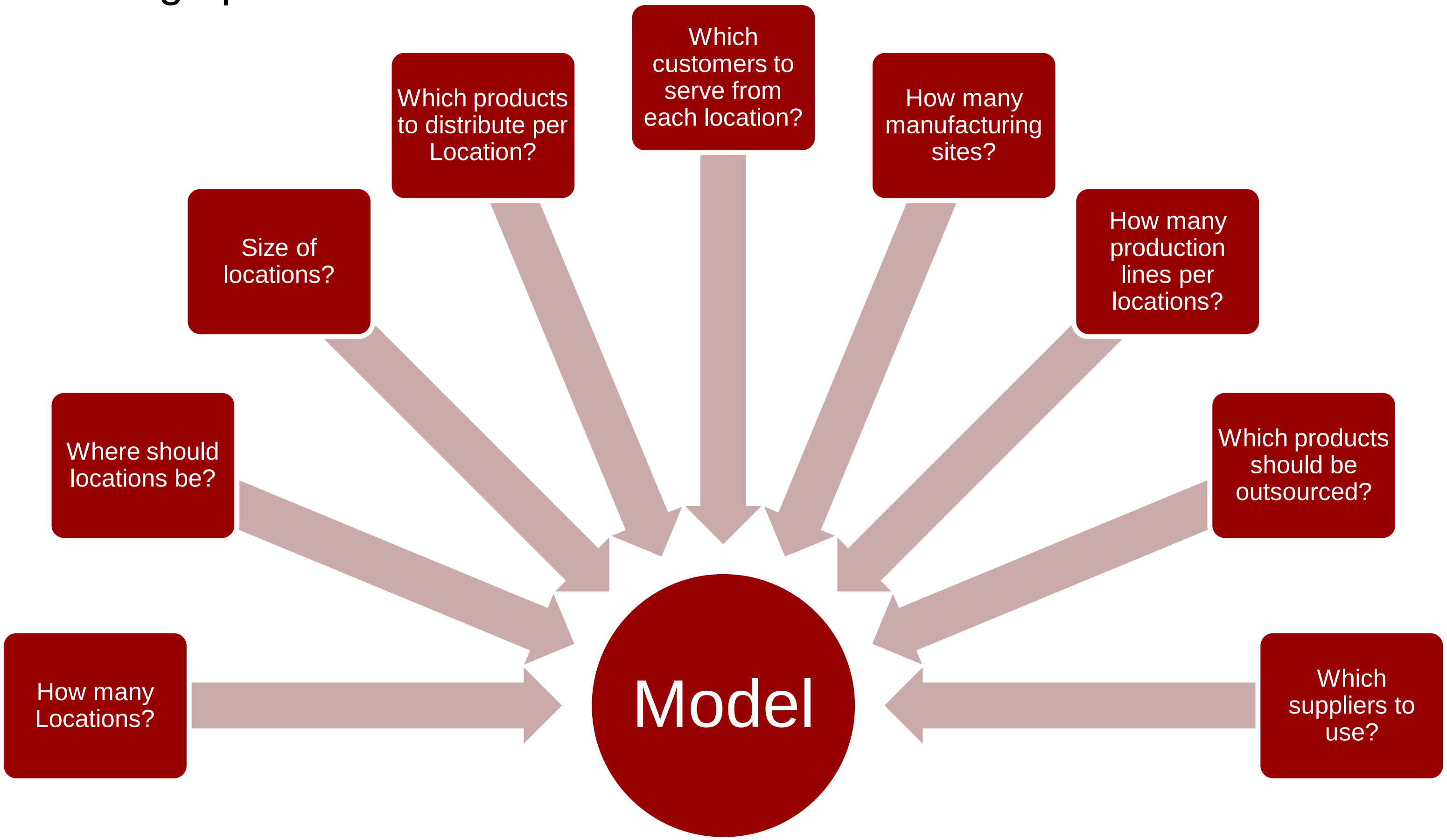
> *Supply Chain Network Design is the discipline used to determine the optimal location and size of facilities and the flow through the facilities .*

- SCDN is sometimes referred to as
  - Network modeling
  - Network optimization
  - Location optimization
  - Facility Location Problem
  - …
- Steps
  - Build mathematical Model
  - Solve this model by using optimization techniques
  - Analyze results to pick the best model

Source: Watson et al. (2012)

Different Network Design Problems I/II

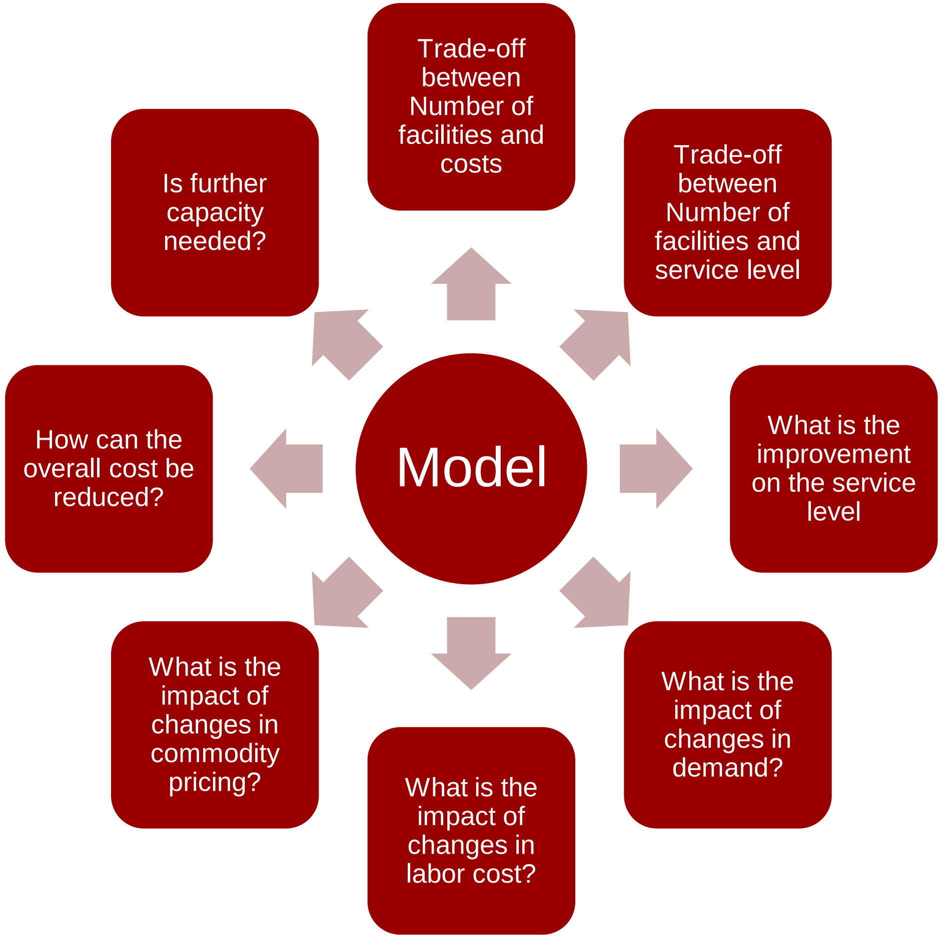
- The Model is dependent on the specific problem that an organization is facing when setting up its network



Source: Watson et al. (2012)

Different Network Design Problems II/II

- From a model, an organization can learn different things to evaluate its network



Source: Watson et al. (2012)

Frequency of model solving

Historically models have been used every several years

- But the speed is increasing

Increasing frequency

- Networks with fast changing products or high demand volatility tend to use it more frequently
- Major events often trigger SCND, such as mergers and acquisitions
- More frequently, the same model (or an updated version of it) is run on a regular basis to adjust the supply chain

Source: Watson et al. (2012)

The impact of a location

Transportation cost

the location determines distances, access to infrastructure and transport carriers

Risk

the locations and its number determines the vulnerability to disruptions and critical external events

Emission

With distance usually the emissions increase

Service Level

the location determines the time to customer

Local Resources

the locations determines access to labor, skill, materials, and utilities

Taxes

the locations determines taxes on operations, profits and shipping products from and to markets

Source: Watson et al. (2012)

Four elements of a SCND Model

- Objective
  - The goal of an optimization and the criteria to compare solutions
- Constraints
  - The rules of a legitimate solution
- Decision Variables
  - What the optimization is allowed to choose from.
- Data

Source: Watson et al. (2012)

Aspects of Data in Network Design

Time of data collection vs. time of collection

- Data collection is time-consuming and the effort increases with the precision of the data
- If too much time is used for data collection, it may take too long to act.

The amount of data

- The necessary data for the decision problem must be available as precisely as possible
- Collection of irrelevant data should be avoided

Time reference of the data used

- Previous year's data is not a perfect prediction of the future but forms the basis for baseline models
- Estimated data on future events and values are likely to be inaccurate and should be treated as follows

The model will primarily contain quantifiable data

- However, non-quantifiable data should be included in the subsequent evaluation of the model

Source: Watson et al. (2012)

Modelling Elements

Non-quantifiable Data

Some data collected cannot be quantified and can only be taken into account in models to a limited extent. Aber…

...this data can still be used!

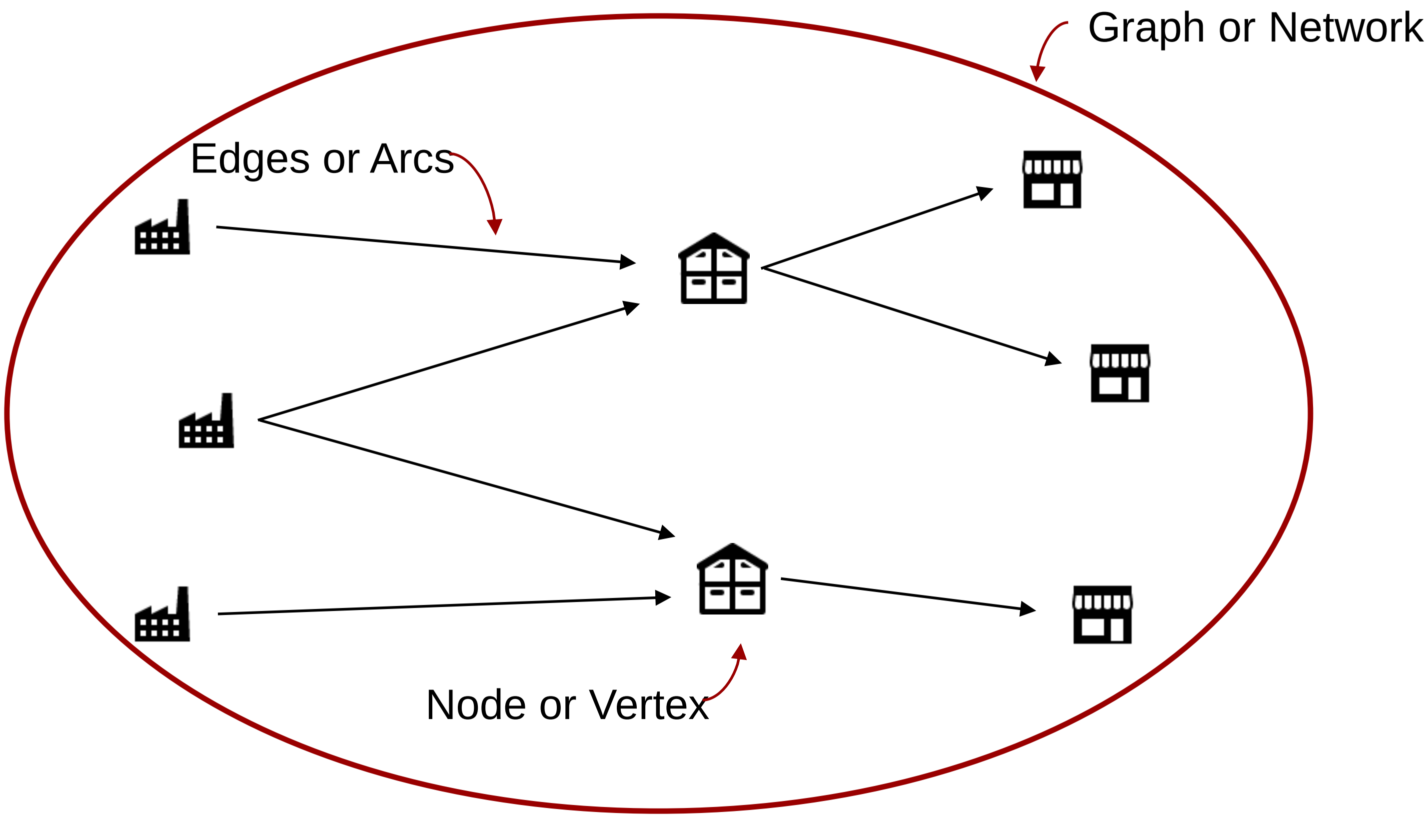
Usually several **scenarios** are developed, which are solved by the model.

The range of possible solutions can now be evaluated with the non-quantifiable data.

Corporate Strategy	Risks	Interruption costs
willingness to change	Public Relation	Competitors
unions	Tax relief	Relationship with supply chain partners

Source: Watson et al. (2012)

Nomenclature



- Node – A Point (Factory, DC or Store)
- Edge – A link between points (Physical flow of goods)
- Network – a collection of Nodes and Edges

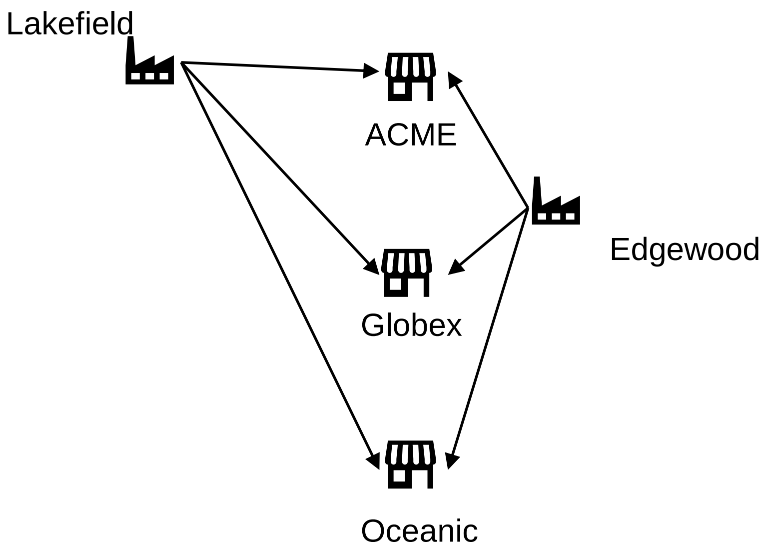
Problem 6-1

Example - Solution is presented

- BAER Hardware manufactures screwdrivers in its plants in Lakefield and Edgewood. Lakefield can produce 300 pieces per week and Edgewood 500. They sell the products to ACME Corp, Globex, and Oceanic. They have a demand of 200, 180 and 360 screwdrivers per week. The transportation costs per screewdriver are given in table 1. Which plant should produce how many screwdrivers for which customer, if costs should be minimized?

from \ to	ACME	Globex	Oceanic
Lakefield	0,25	0,45	0,70
Edgewood	0,35	0,20	0,50

Table 1



A simple network transportation model I/V

Data (Structure)

- $Plants\ i\ \in\ I$ There are a number of I Plants, each having an index i. E.g.: {i=1: Lakefield; i=2: Edgewood}
- $Customers\ j\ \in J$ There are a number of J Customers, each having an index j. E.g.: {j=1: ACME; j=2: Globex; j=3: Oceanic}

Data (attributes of the structure)

- $Demand\ D_j$ $\forall j \in J$ The demand of each customer j in J. E.g. $D_1 = 300$
- $Supply\ S_i$ $\forall i \in I$ The supply of each plant i in I. E.g. $S_1 = 200$
- $Cost\ c_{ij}$ $\forall j, i$ The costs of transporting goods from a plant i to a customer j. E.g. $c_{11} = 0{,}25$

Decision variables

- x_{ij} The flows form every plant i to every customer j. E.g. x_{11} is the flow from Lakefield to ACME

A simple network transportation model II/V

Constraints

- $x_{ij} \geq 0$ $\forall i, j$ Every flow has to be greater or equal then zero. There is no negative flow.
- $\sum_j x_{ij} \geq D_j$ $\forall j \in J$ The sum of all flows to a customer j has to be greater or equal the demand of j. This has to hold for every customer.
- $\sum_i x_{ij} \leq S_i$ $\forall i \in I$ The sum of all flows from a plant i has to be smaller or equal the supply of i. This has to hold for every plant.

Objective

- $\min z = f(x) = \sum_i \sum_j x_{ij} * c_{ij}$ Minimize the total transportations cost which are the sum of all flows of screwdrivers shipped per link multiplied with the links' transportation costs

Nomenclature

Values
D – Demand
S – Supply / Capacity
c – cost
x – flow of goods

Indices
j – Receiver (e.g. Customer)
i – Sender (e.g. plants)

A simple network transportation model III/V

The complete model (in mathematical formulation)

$$\min z = f(x) = \sum_i \sum_j x_{ij} * c_{ij}$$

The objective function

"Subject to" or "such that"

$$s.t.$$

$$\sum_i x_{ij} \geq D_j \qquad \forall j \in J$$

$$\sum_j x_{ij} \leq S_i \qquad \forall i \in I$$

$$x_{ij} \geq 0 \qquad \forall i, j$$

The constraints

Nomenclature

Values

D – Demand

S – Supply / Capacity

c – cost

x – flow of goods

Indices

j – Receiver (e.g. Customer)

i – Sender (e.g. plants)

A simple network transportation model IV/V

- Some programs for linear programming use a matrix based notation to present the model. In this notation, every constraint has its row and every decision variable its column.
- However, in larger problems with many constraints and decision variables but only few decision variables per constraints, this creates a *sparse matrix* – mostly containing zeros. Thus, this formulation becomes somewhat unpractical.

$$\min z = f(x) = \sum_i \sum_j x_{ij} * c_{ij}$$

$s.t.$

$$\sum_i x_{ij} \geq D_j \qquad \forall j \in J$$

$$\sum_j x_{ij} \leq S_i \qquad \forall i \in I$$

$$x_{ij} \geq 0 \qquad \forall i, j$$

The complete model (matrix notation)

Z=	c11*x11 +	c12*x12 +	c13*x13 +	c21*x21 +	c22*x22 +	c23*x23		
s.t.								
	x11 +			x21			>=	D1
		x12 +			x22		>=	D2
			x13 +			x23	>=	D3
	x11 +	x12 +	x13				<=	S1
				x21 +	x22 +	X23	<=	S2
	x11						>=	0
		x12					>=	0
			x13				>=	0
				x21			>=	0
					x22		>=	0
						x23	>=	0

A simple network transportation model IV/IV

The final model with data (with all data entered)

Z=	0,25*x11 +	0,45*x12 +	0,70*x13 +	0,35*x21 +	0,20*x22 +	0,50*x23		
s.t.								
	x11 +			x21			>=	200
		x12 +			x22		>=	180
			x13 +			x23	>=	360
	x11 +	x12 +	x13				<=	300
				x21 +	x22 +	X23	<=	500
	x11						>=	0
		x12					>=	0
			x13				>=	0
				x21			>=	0
					x22		>=	0
						x23	>=	0

Summary: Network Optimization

- In network optimization there are several problems to solve (where should locations be set up) and knowledge learned (costs if demand changes)
- A network model contains an objective, constraints, decision variables and data
- The math programming formulation keeps the model abstract and aggregates the constraints but does not give overview of the model. Whereas the matrix notation helps to understand the model but becomes hard to handle if complexity of the model increases

Setting the model up in a spreadsheet

- Notice: this way is meant to keep overview of the model in the spreadsheet. An aggregated form is possible, but harder to set up and control.
- Notice: the set up is very close to matrix notation, but there are slight difference
  - There will be an additional row for the decision variables, which the software will later change
  - Every other decision variable in the model will be changed to a 1 and later multiplied by the decision variables to evaluate the constraint

Z=	0,25	0,45	0,70	0,35	0,20	0,50			
$\bar{x}$	x11	x12	x13	x21	x22	x23			
s.t.									
	1			1			$\bar{x} * \bar{v}$		
	$\bar{v}$	1			1			>=	180
			1			1		>=	360
	1	1	1					<=	300
				1	1	1		<=	500
	1							>=	0
		1						>=	0
			1					>=	0
				1				>=	0
					1			>=	0
						1		>=	0

Multiplication of two vectors: decision variables ($\bar{x}$) and values of the constraint ($\bar{v}$).

The matrix notation in a spreadsheet software

Example - Solution is presented

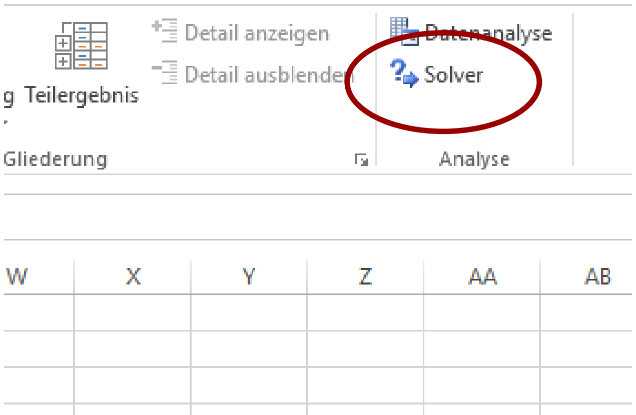
=SUMMENPRODUKT(B4:G4;B8:G8)

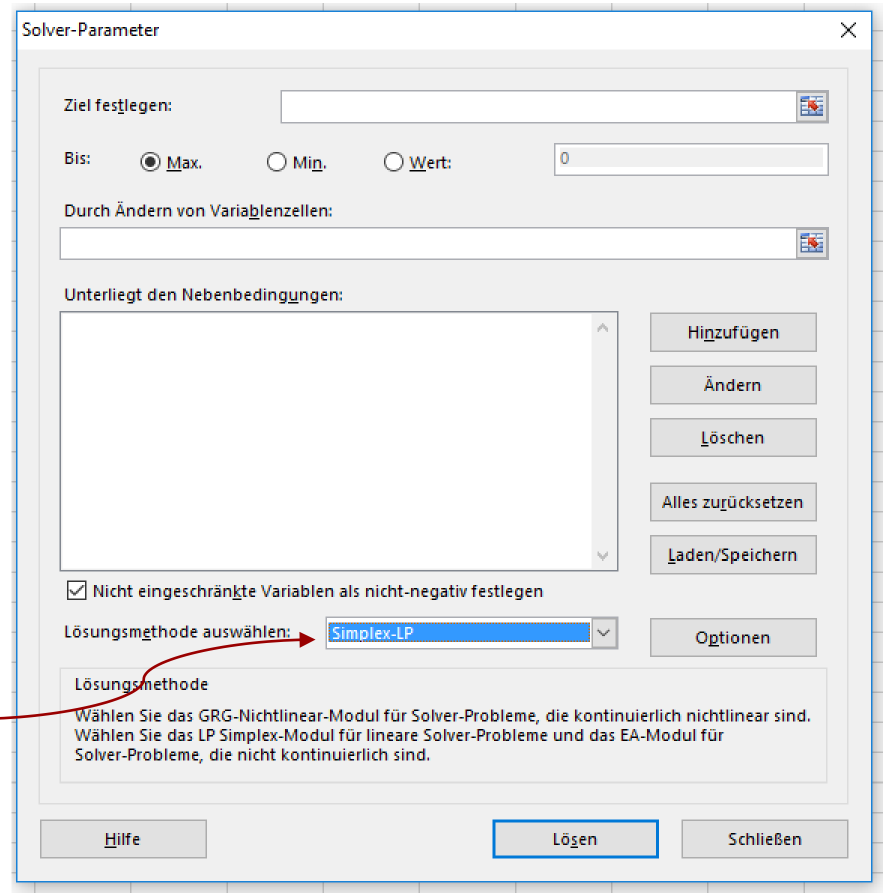
=SUMMENPRODUKT(B12:G12;B8:G8)

	A	B	C	D	E	F	G	H		
1	**Objective**									
2	min z =	0								
3	*Objective data (costs)*									
4		0,25	0,45	0,7	0,35	0,2	0,5			
5										
6	**Decision Variables**									
7		x11	x12	x13	x21	x22	x23			
8										
9										
10	**Constraints**									
11	*Demand*									
12		1			1			0	>=	200
13			1			1		0	>=	180
14				1			1	0	>=	360
15										
16	*Supply*									
17		1	1	1				0	<=	300
18					1	1	1	0	<=	500
19										
20	*Non-negativi*	1						0	>=	0
21			1					0	>=	0
22				1				0	>=	0
23					1			0	>=	0
24						1		0	>=	0
25							1	0	>=	0
26										
27										

Setting up the solver I/II

Example - Solution is presented



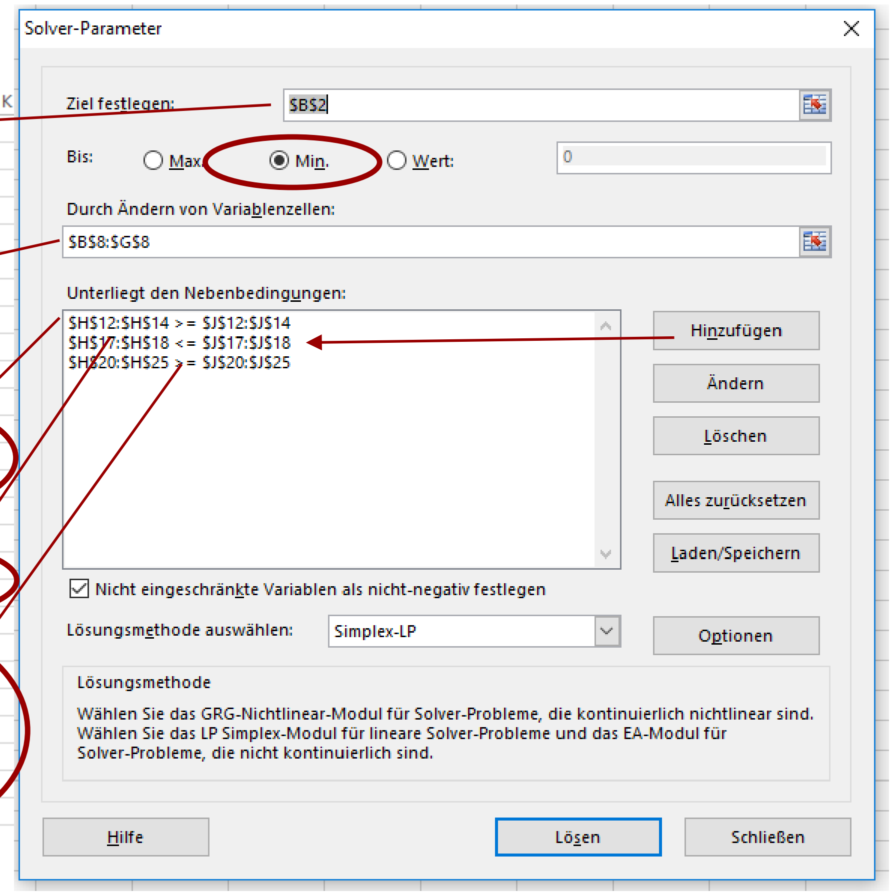


Please use simplex except stated otherwise

Setting up the solver II/II

Example - Solution is presented

	A	B	C	D	E	F	G	H	I	J	K
1	**Objective**										
2	min z =	0									
3	*Objective data (costs)*										
4		0,25	0,45	0,7	0,35	0,2	0,5				
5											
6	**Decision Variables**										
7		x11	x12	x13	x21	x22	x23				
8											
9											
10	**Constraints**										
11	*Demand*										
12		1			1			0	>=	200	
13			1			1		0	>=	180	
14				1			1	0	>=	360	
15											
16	*Supply*										
17		1	1	1				0	<=	300	
18					1	1	1	0	<=	500	
19											
20	*Non-negativi*	1						0	>=	0	
21			1					0	>=	0	
22				1				0	>=	0	
23					1			0	>=	0	
24						1		0	>=	0	
25							1	0	>=	0	
26											



Example - Solution is presented

Problem 6-1 – The solved model

The minimum is 274

200 Screwdrivers go from Lakefield to ACME, 0 from Lakefield to Globex, …

Due to minimization, all customers only get as much as they need

The supply of Edgewood is completely used, Lakefield would have supply left

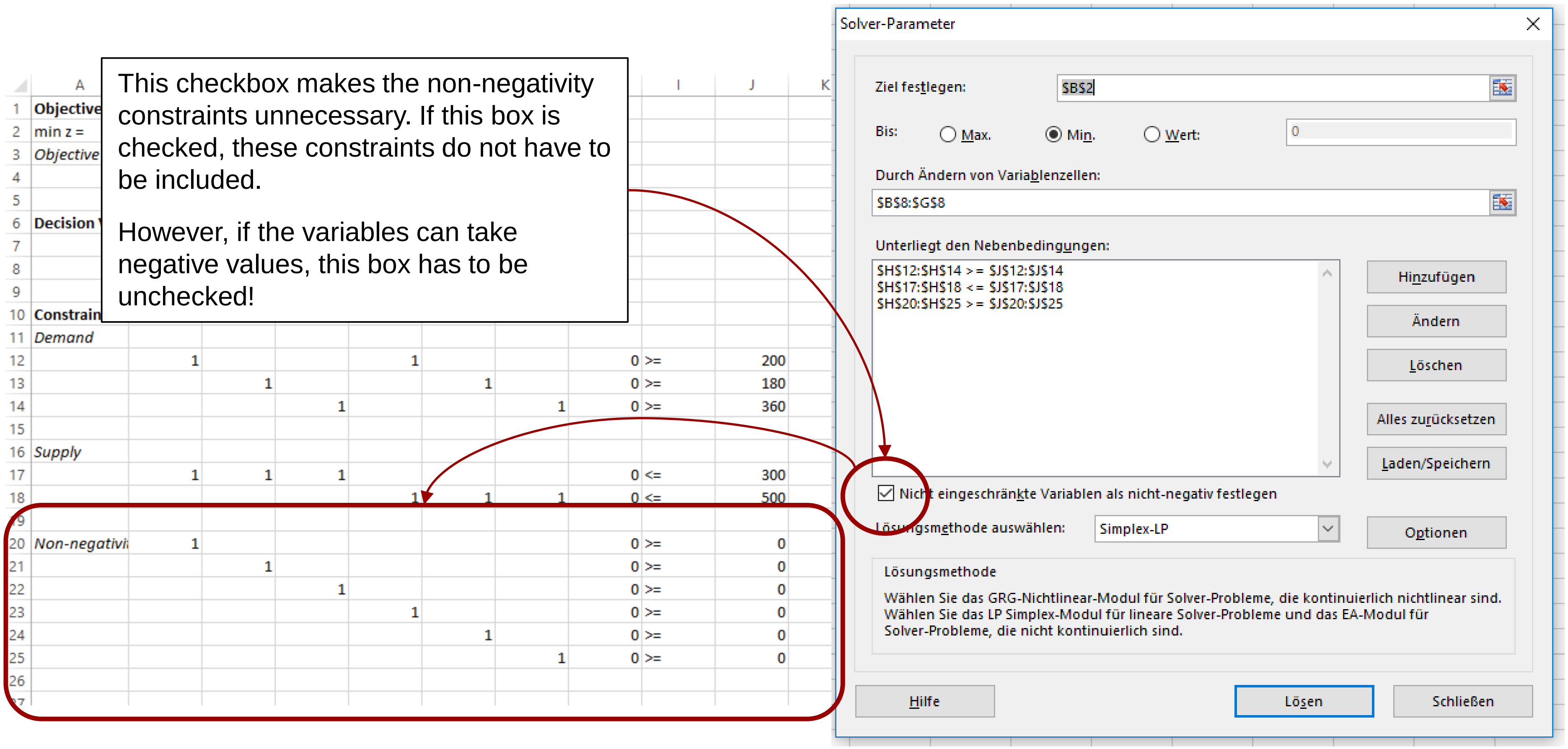
	A	B	C	D	E	F	G	H	I	J
1	**Objective**									
2	min z =	274								
3	*Objective data (costs)*									
4		0,25	0,45	0,7	0,35	0,2	0,5			
5										
6	**Decision Variables**									
7		x11	x12	x13	x21	x22	x23			
8		200	0	40	0	180	320			
9										
10	**Constraints**									
11	*Demand*									
12		1			1			200	>=	200
13			1			1		180	>=	180
14				1			1	360	>=	360
15										
16	*Supply*									
17		1	1	1				240	<=	300
18					1	1	1	500	<=	500
19										
20	*Non-negativi*	1						200	>=	0
21			1					0	>=	0
22				1				40	>=	0
23					1			0	>=	0
24						1		180	>=	0
25							1	320	>=	0
26										
27										

Remarks

Example - Solution is presented

This checkbox makes the non-negativity constraints unnecessary. If this box is checked, these constraints do not have to be included.

However, if the variables can take negative values, this box has to be unchecked!



Problem 6-2

- BEAR Electronics offers a after sales service for his products distributing spare parts. The Distribution Centers are located in Alma and Bridgeport and can handle 750 and 920 orders a month respectively. The both DCs are using different Service providers leading to different distribution costs, given in table 1. They are supposed to serve the five customer areas of Rycroft, Scottsville, Thorn Hill, Valleyside, and Witmer. The approximated demand of these areas is given in table 2. The distances are given in table 3.

	Any area
Alma	0,15
Bridgeport	0,22

Table 1: € / km

Rycroft	190
Scottsville	210
Thorn Hill	165
Valleyside	310
Witmer	225

Table 2: Demand

	Rycroft	Scottsville	Thorn Hill	Valleyside	Witmer
Alma	75	102	15	90	53
Bridgeport	68	72	83	64	12

Table 2: Distance in km

Problem 6-2 – Solution

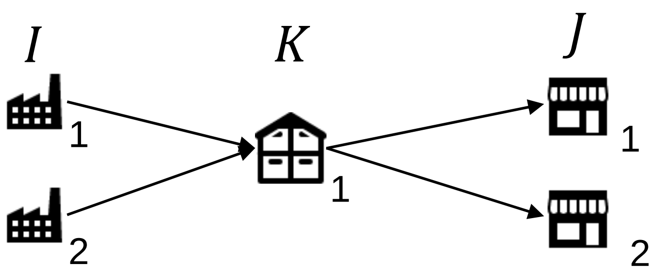
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	**Objective**													
2	*min z =*	10568,25												
3														
4	**Data**													
5	*Distances*													
6		75	102	15	90	53	68	72	83	64	12			
7														
8	*Costs*													
9		0,15	0,15	0,15	0,15	0,15	0,22	0,22	0,22	0,22	0,22			
10														
11	**Decision Variables**													
12		xAR	xAS	xAT	xAV	xAW	xBR	xBS	xBT	xBV	xBW			
13		190	85	165	310	0	0	125	0	0	225			
14														
15	**Constraints**													
16	*Demand*													
17		1					1					190	>=	190
18			1					1				210	>=	210
19				1					1			165	>=	165
20					1					1		310	>=	310
21						1					1	225	>=	225
22	*Supply*													
23		1	1	1	1	1						750	<=	750
24							1	1	1	1	1	350	<=	920
25														

Conservation of flow

transshipment of the flow of goods

- The conservation of flow is a nessecary condition for every distribution point in a network model that is located between source and sink
- *The basic idea*: Everything that goes in, has to go out again
  - A distribution center (k) between production and customer
  - A production plant between suppliers and customers

Distribution



$$\underbrace{\sum_i x_{ik}}_{\text{In}} = \underbrace{\sum_j x_{kj}}_{\text{out}} \qquad \forall k$$

Nomenclature

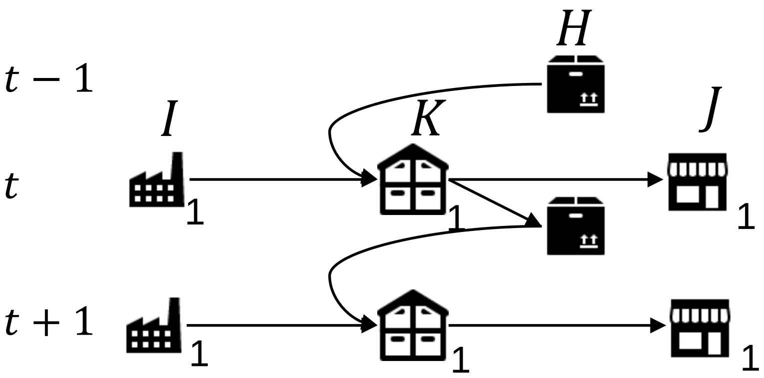
Values
D – Demand
S – Supply / Capacity
c – cost
x – flow of goods
H – Amount of goods held

Indices
j – Receiver (e.g. Customer)
i – Sender (e.g. plants)
k – hub (e.g. distribution center)
t – time period

Conservation of flow
Holding goods

- (Goods) flow conservation is a necessary condition for every distribution point in a network model that lies between source and sink.
  - Stock that was produced in one period (e.g. t=0), but consumed in another period (e.g. t=1) and remains in a warehouse ($h \in H$).

Holding goods over periods



$$\underbrace{\sum_i x_{ikt} + H_{k,t-1}}_{\text{in}} = \underbrace{\sum_j x_{kjt} + H_{k,t}}_{\text{out}} \quad \forall k$$

Nomenclature

Values
D – Demand
S – Supply / Capacity
c – cost
x – flow of goods
H – Amount of goods held

Indices
j – Receiver (e.g. Customer)
i – Sender (e.g. plants)
k – hub (e.g. distribution center)
t – time period

Conservation of flow – Example model

$$\min z = \sum_i \sum_k x_{ik} * c_{ik} + \sum_k \sum_j x_{kj} * c_{kj}$$

Minimize costs including transportation (c) times the flow on a lane (x) leaving i to any k and any k to any j

s.t.

$$\sum_j x_{kj} \geq D_j \qquad \forall j \in J$$

All demand (D) has to be satisfied

$$\sum_k x_{ik} \leq S_i \qquad \forall i \in I$$

The supply (S) of a flow (x) of a facility i to any k can not be exceeded.

$$\sum_i x_{ik} = \sum_j x_{kj} \qquad \forall k$$

For all facilities k, the sum of all flow (x) into k from any i has to be equal to the sum of all flow out to any j.

$$x_{ik} \geq 0 \qquad \forall i, k$$

All flows (x) from i to k have to be greather or equal to zero

$$x_{kj} \geq 0 \qquad \forall k, j$$

All flows (x) from k to j have to be greather or equal to zero

Nomenclature

Values

D – Demand
S – Supply / Capacity
c – cost
x – flow of goods
H – Amount of goods held

Indices

j – Receiver (e.g. Customer)
i – Sender (e.g. plants)
k – hub (e.g. distribution center)
t – time period

Practical use

- For the Matrix notation consider the following:
  - If $\sum_i x_{ik} = \sum_j x_{kj} \quad \forall k$
  - Then $\sum_i x_{ik} - \sum_j x_{kj} = 0 \quad \forall k$
- Therefore, in a matrix notation with supply nodes {1,2,3}, one distribution node {4} node and demand nodes {5,6,7}, the matrix would be set up as followed

Z=	0,25	0,45	0,70	0,35	0,20	0,50			
	x14	x24	x34	x45	x46	x47			
s.t.									
	1	1	1	-1	-1	-1		=	0

Problem 6-3

- BEAR Foods is supplying fresh water canisters to its customers Initech, Umbreall Corp and Universal Exports. Their Monthly demand is given in Table 1. The water canisters are manufactured in three plants in Burlington, Gallows and Harborview. The distribution is made via the DCs in Hazeldell and Hillside. The Capacity of the companies facilities are given in Table 2. The transportation costs per water canister are given in table 3.
- What is the minimum cost flow of goods?

Initech	15.000
Umbrella Co	18.500
Universal Ex	16.800

Table 1: Demand

Burlington	40.000
Gallows	20.000
Harborview	12.000
Hazeldell	30.000
Hillside	30.000

Table 2: Capacity

From \ To	id	4	5	6	7	8
Burlington	1	1,25	1,73			
Gallows	2	1,52	1,26			
Harborview	3	0,74	0,82			
Hazeldell	4	0		0,89	1,35	1,09
Hillside	5		0	1,28	0,96	1,17
Initech	6			0		
Umbrella Co	7				0	
Universal Ex	8					0

Table 3: Transportation costs

Problem 6-3 – Solution

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	**Objective**															
2	*min z*	106545														
3																
4	**Data**															
5	*Costs*															
6		1,25	1,73	1,52	1,26	0,74	0,82	0,89	1,35	1,09	1,28	0,96	1,17			
7																
8	**Decision Variables**															
9		x14	x15	x24	x25	x34	x35	x46	x47	x48	x56	x57	x58			
10		18300	0	0	20000	11700	300	15000	0	15000	0	18500	1800			
11																
12	**Constraints**															
13	*Capacity (Plants)*															
14		1	1											18300	<=	40.000
15				1	1									20000	<=	20.000
16						1	1							12000	<=	12.000
17	*Capacity (DC)*															
18		1		1		1								30000	<=	30.000
19			1		1		1							20300	<=	30.000
20	*Demand*															
21								1			1			15000	>=	15.000
22									1			1		18500	>=	18.500
23										1			1	16800	>=	16.800
24	*Conservation of flow*															
25		1		1		1		-1	-1	-1				0	=	0
26			1		1		1				-1	-1	-1	0	=	0

Due to consv. Of flow, the capacity constraint for one direction is enough

The sum of everything going in minus the sum of everything going out equals zero

Summary: SCDN with Spreadsheets

- Matrix notation provides the model layout with an overview, but also enlarges the space required for display
- For distribution points between source and sink, river conservation applies: everything that flows in must flow out again.

References

- Watson, M.; Lewis, S.; Cacioppo, P., Jayaraman, J. (2012): Supply Chain Network Design.
- Shapiro, J. F. (2007): Modeling the Supply Chain. 2nd Ed.