Norwegian School of Economics

BAN402 Project 2

Decision Modelling in Business

Candidates: 28,139
Fall 2024

Contents

Notes on color coding in model formulations:

- Black: The general mathematical model
- Blue: Specific formulations and examples from case
- Purple: Supplementary explanations of the model to clarify our approach and/or why different parts of the model are needed.
- Orange: Modification made to previously formulated models

Part A

Task A1:

a)

We introduce a new binary variable:

$$z_k = \begin{cases} 1 & \text{if the total demand in the tour of truck } k \text{ is above } t, \\ 0 & \text{otherwise} \end{cases}, \qquad \begin{matrix} \forall k \in T \\ z_k \in \{0,1\} \end{matrix}$$

The overall estimated CO2 emissions can be estimated as the fixed total emission per truck (E) multiplied by the number of trucks (K), plus the excess emissions from a total demand above t ($\bar{E} \; - E$) for all trucks with a total demand above t.

Thus, the new objective function to be minimized can be expressed as:

$$\min total\_emissions = E * K + (\bar{E} - E) * \sum_{k \in T} z_k$$

To link z to the total demand, we add the following constraints:

$$\sum_{i \in S} b_i * y_{i,k} \geq t * z_k, \qquad \forall k \in T \tag{A.9a}$$

$$\sum_{i \in S} b_i * y_{i,k} - t \leq z_k * M, \qquad \forall k \in T \tag{A.9b}$$

Constraint (A.9a) ensures that z_k must take value zero if the total demand in a route for a given truck k is below the threshold t.

Constraint (A.9b) ensures that z_k must take value one if the total demand in a given trucks route is above the threshold. M is the maximum difference between total route demand and the threshold (the LHS of the constraint).

b)

We add a new binary variable to capture whether a tour includes more than three stressful roads:

$$a_k = \begin{cases} 1 & \text{if tour } k \text{ includes 3 or more stressfull roads,} \\ 0 & \text{otherwise} \end{cases}, \qquad \begin{matrix} \forall k \in T \\ a_k \in \{0,1\} \end{matrix}$$

We add a new constraint, stating that the number of tours with more than three stressful roads cannot exceed five:

$$\sum_{k \in T} a_k \leq 5 \tag{A.10}$$

Finally, we link a_k to the number of stressful roads on a given tour:

For easier interpretation we define the number of stressful roads on the tour of a given truck k as $StressCount$:

$$StressCount_k = \sum_{j \in N_i} x_{ijk}\,, \qquad \forall i \in S$$

If a given tour's $StressCount$ is 3 or more, a_k must equal one. Here M is a sufficiently large number not to limit the maximum difference between $StressCount$ and 3.

$$StressCount_k - 3 < M * a_k, \qquad \forall k \in T \tag{A.11a}$$

If a given tour's $StressCount$ is less than 3, a_k must equal zero as it is binary:

$$a_k \leq \frac{StressCount_k}{3}, \qquad \forall k \in T \tag{A.11b}$$

Task A2:

***(Note:** Regarding constraint numbering, we will ignore the constraints added in task A1 in the following. That is, when we for instance write A.9, we are referring to expression A.9 in the paper from (Wieczorrek et.al, 2023) and not our constraint linking z to total demand from task A1)*

We add the new set L with tuples of suppliers situated close to each other:

L: Set of tuples (j, k, h) where j, k and h are suppliers in S and close in proximity in the network.

We add a new parameter:

$u_{j,k,h}$: the reduction in delay cost improvement if a measure i is allocated to suppliers j, k and h.

We add a new binary variable γ to capture whether measure i is allocated to a given supplier-tuple (j, k, h):

$$\gamma_{i,j,k,h} = \begin{cases} 1 & \text{if measure } i \text{ is allocated to all suppliers in tuple } (j,k,h), \\ 0 & \text{otherwise} \end{cases}, \qquad \forall i \in M, \ \forall (j,k,h) \in L$$

We modify the objective function, so it decreases with amount $u_{j,k,h}$ for each true γ:

$$\max \sum_{i \in M} \sum_{j \in S} (\Delta_{ij} - C_i^M) x_{ij} - \sum_{(j,k,h) \in L} \sum_{i \in M} \gamma_{i,j,k,h} * u_{j,k,h} \tag{A.9}$$

To link γ to x we add the following constraints:

If measure i is not allocated to at least one of the suppliers in tuple (j, k, h), then γ must take value 0:

$$\gamma_{i,j,k,h} \leq x_{i,j} * x_{i,k} * x_{i,h}, \qquad \forall i \in M, \ \forall (j,k,h) \in L \tag{A.13a}$$

If measure i is allocated to all suppliers in tuple (j, k, h), then γ must take value 1:

$$\gamma_{i,j,k,h} > x_{i,j} + x_{i,k} + x_{i,h} - 3, \qquad \forall i \in M, \ \forall (j,k,h) \in L \tag{A.13b}$$

γ is binary:

$$\gamma_{i,j,k,h} \in \{0,1\}, \qquad \forall i \in M, \ \forall (j,k,h) \in L \tag{A.14}$$

Part B

Task B1:

Model formulation:

Sets:

T: Set of teams [GER, SCO, ..., CZE]

G: Set of groups [A,B,..., F]

D: Set of dates [12, ... ,26]

M: Set of matches [M1, M2, ..., M36]

V: Set of venues [BER, LEI, ..., MUN]

TM_t: Ordered set of matches for each team t, $\forall t \in T$. Ex: TM_{ESP}: $[M3,\ M16,\ M27]$; named *TeamMatches* in AMPL.

Parameters:

$d_{v1,v2}$: Distance from venue v_1 to venue v_2, where $v_1, v_2 \in V$

$NumMatches_v$: The number of matches held at venue v in the original schedule, $\forall v \in V$

$$MD_{m,d} = \begin{cases} 1 & \text{if match } m \text{ is played on date } d \text{ in the original schedule,} \\ 0 & \text{otherwise} \end{cases}, \quad \begin{matrix} \forall m \in M \\ \forall d \in D \end{matrix}$$

Decision variables:

Because we want the same matchup and dates as in the original schedule, the decision-maker is faced with the decision of assigning matches to venues, such that the total distance travelled is minimized subject to a set of constraints.

We define the decision variable x as follows:

$$x_{m,v} = \begin{cases} 1 & \text{if match } m \text{ is played at venue } v, \\ 0 & \text{otherwise} \end{cases}$$

To calculate distance traveled by teams from our proposed schedule ($x_{m,v}$) we need a variable z that specifies whether a team uses each arc in the network connecting the venues:

$$z^1_{t,v_1,v_2} = \begin{cases} 1 & \text{if team } t \text{ plays their first match at venue } v_1 \text{ and second match at venue } v_2, \\ 0 & \text{otherwise} \end{cases}$$

$$z^2_{t,v_1,v_2} = \begin{cases} 1 & \text{if team } t \text{ plays their second match at venue } v_1 \text{ and third match at venue } v_2, \\ 0 & \text{otherwise} \end{cases}$$

To link z to our proposed schedule $x_{m,v}$, we furthermore need a variable TV, that links each team's ordered list of matches to a venue:

$$TV_{t,v}^{1} = \begin{cases} 1 & \text{if team } t \text{ plays their first match at venue } v, \\ 0 & \text{otherwise} \end{cases}$$

$$TV_{t,v}^{2} = \begin{cases} 1 & \text{if team } t \text{ plays their second match at venue } v, \\ 0 & \text{otherwise} \end{cases}$$

$$TV_{t,v}^{3} = \begin{cases} 1 & \text{if team } t \text{ plays their third match at venue } v, \\ 0 & \text{otherwise} \end{cases}$$

To ensure that the group-related constraints (6 and 12) are met, we finally need a variable y that connects our decision variable $x_{m,v}$ to groups:

$$y_{g,m,v} = \begin{cases} 1 & \text{if } x_{m,v} \text{ belongs to group } g, \\ 0 & \text{otherwise} \end{cases}$$

In order to count the distinct number of venues used for each group (constraint 11) we need an additional variable w that links each venue to groups:

$$w_{g,v} = \begin{cases} 1 & \text{if venue } v \text{ hosts at least one match for group } g, \\ 0 & \text{otherwise} \end{cases}$$

Objective function:

Our objective is to find an optimal schedule to minimize the total distance travelled:

$$\min total\_distance = \sum_{t \in T, v_1 \in V, v_2 \in V} (z_{t,v_1,v_2}^{1} + z_{t,v_1,v_2}^{2}) d_{v_1,v_2}$$

Subject to:

Linking z^i to TV:

$$z_{t,v_1,v_2}^{i} \leq TV_{t,v_1}^{i}, \qquad \begin{matrix} \forall i \in \{1,2\} \\ \forall v_1, v_2 \in V \\ \forall t \in T \end{matrix} \tag{1}$$

$$z_{t,v_1,v_2}^{i} \leq TV_{t,v_2}^{i+1}, \qquad \begin{matrix} \forall i \in \{1,2\} \\ \forall v_1, v_2 \in V \\ \forall t \in T \end{matrix} \tag{2}$$

$$z_{t,v_1,v_2}^{i} \geq TV_{t,v_1}^{i} + TV_{t,v_2}^{i+1} - 1, \qquad \begin{matrix} \forall i \in \{1,2\} \\ \forall v_1, v_2 \in V \\ \forall t \in T \end{matrix} \tag{3}$$

Constraint (1) ensures that z^i cannot be true, if team t does not play at venue v_1 in their first(second) match, for $i = 1(2)$.

Constraint (2) ensures z^i cannot be true, if team t does not play at venue v_2 in their second(third) match for $i = 1(2)$

Constraint (3) ensures that z^i must be true, if team t play at venue v_1 in their first(second) match and v_2 for their second (third) match, for $i = 1(2)$.

Example: $z^1_{ESP,Berlin,Hamburg}$ can only equal 1 if ESP plays both their first match in Berlin ($TV^1_{ESP,Berlin} = 1$) and second match in Hamburg ($TV^2_{ESP,Hamburg} = 1$). If either of them are zero, then $z^1_{ESP,Berlin,Hamburg}$ is zero.

Linking w to y:

For the number of matches hosted at a given venue for a given group to be positive, w must take value 1. The "big M" is set to 6 as each group plays 6 matches.

$$\sum_{m \in M} y_{g,m,v} \leq 6 * w_{g,w}, \qquad \begin{matrix} \forall g \in G \\ \forall v \in V \end{matrix} \tag{4}$$

If the number of matches hosted at a given venue for a given group is zero, w must take value zero:

$$\sum_{m \in M} y_{g,m,v} \geq w_{g,w}, \qquad \begin{matrix} \forall g \in G \\ \forall v \in V \end{matrix} \tag{5}$$

All matches must belong to one group:

$$\sum_{v \in V} \sum_{g \in G} y_{g,m,v} = 1, \qquad \forall m \in M \tag{6}$$

All matches must be assigned a venue:

$$\sum_{v \in V} x_{m,v} = 1, \qquad \forall m \in M \tag{7}$$

Each venue should host the same number of matches as in the original solution:

$$\sum_{m \in M} x_{m,v} = NumMatches_v, \qquad \forall v \in V \tag{8}$$

After a match is held at a venue, there is a cooldown of at least 2 days:

$$\sum_{m \in M} x_{m,v} * MD_{m,d} \leq 1 - \sum_{m \in M} x_{m,v} * MD_{m,d-1}, \qquad \begin{matrix} \forall v \in V, \\ \forall d \in D, \\ d \geq 14 \end{matrix} \tag{9}$$

$$\sum_{m \in M} x_{m,v} * MD_{m,d} \leq 1 - \sum_{m \in M} x_{m,v} * MD_{m,d-2}, \qquad \begin{matrix} \forall v \in V, \\ \forall d \in D, \\ d \geq 14 \end{matrix} \tag{10}$$

(**Note:** In order to implement this logic in AMPL, the days 12 and 13 were added to set D)

Number of different venues used for the matches of each group must be at least four:

$$\sum_{v \in V} w_{g,v} \geq 4, \qquad \forall\, g \in G \tag{11}$$

Each venue must not host more than two matches of the same group:

$$\sum_{m \in M} y_{g,m,v} \leq 2, \qquad \forall\, g \in G, \qquad \forall v \in V \tag{12}$$

All venues should have hosted at least one match by 18 June:

$$\sum_{d=14}^{18} \sum_{m \in M} x_{m,v} * MD_{m,d} \geq 1\,, \qquad \forall v \in V \tag{13}$$

The last match held at each venue cannot be before 24 June:

$$\sum_{d=24}^{26} \sum_{m \in M} x_{m,v} * MD_{m,d} \geq 1\,, \qquad \forall v \in V \tag{14}$$

The first games for *mainseed teams* should be held at the same venue as the original solution:

$$x_{M1,MUN} + x_{M3,BER} + x_{M5,GEL} + x_{M8,DUS} + x_{M9,FRK} + x_{M12,LEI} = 6 \tag{15}$$

Finally, the constraints that each pair of teams must play on the same date as the original solution and play every other team in its group is ensured by keeping the matches and dates from the original solution.

Solution:

Solving the model in AMPL gives the following optimal schedule:

	14.6	15.6	16.6	17.6	18.6	19.6	20.6	21.6	22.6	23.6	24.6	25.6	26.6
Berlin		**M3** (ESP-CRO)					**M16** (ESP-ITA)						**M33** (SVK-ROU)
Leipzig					**M12** (POR-CZE)			**M21** (SVK-UKR)			**M28** (CRO-ITA)		
Hamburg		**M4** (ITA-ALB)				**M15** (CRO-ALB)			**M24** (GEO-CZE)				**M35** (GEO-POR)
Dortmund		**M2** (HUN-SUI)			**M11** (TUR-GEO)				**M23** (TUR-POR)			**M32** (FRA-POL)	
Gelsen-kirchen			**M5** (SRB-ENG)					**M19** (POL-AUT)				**M31** (NED-AUT)	
Dusseldorf				**M8** (AUT-FRA)				**M20** (NED-FRA)			**M27** (ALB-ESP)		
Cologne			**M7** (POL-NED)			**M13** (SCO-SUI)				**M25** (SUI-GER)			**M36** (CZE-TUR)
Frankfurt				**M9** (BEL-SVK)			**M17** (DEN-ENG)			**M26** (SCO-HUN)			**M34** (UKR-BEL)
Stuttgart			**M6** (SVN-DEN)			**M14** (GER-HUN)			**M22** (BEL-ROU)			**M30** (DEN-SRB)	
Munich	**M1** (GER-SCO)			**M10** (ROU-UKR)			**M18** (SVN-SRB)					**M29** (ENG-SVN)	

Groups:	A	B	C	D	E	F

Germany plays in matches M1, M14 and M25. These matches are played in Munich, Stuttgart and Cologne, respectively.

The total distance travelled by each team is given in the table below. By switching to the proposed schedule, the league could have reduced the distance traveled by 3,204 km and saved an estimated total of 986,832 kg CO2.

Task B1: Candidates: 28,139

Group	Team	Abbrevation	Distance travelled (km)	CO2 emissions aprox. (kg)
A	Germany *	GER	600	184,800
	Scotland	SCO	809	249,172
	Hungary	HUN	654	201,432
	Switzerland	SUI	99	30,492
B	Spain *	ESP	564	173,712
	Croatia	CRO	689	212,212
	Italy	ITA	470	144,760
	Albania	ALB	401	123,508
C	Slovenia	SVN	222	68,376
	Denmark	DEN	404	124,432
	Serbia	SRB	889	273,812
	England *	ENG	693	213,444
D	Poland	POL	143	44,044
	Netherlands	NED	120	36,960
	Austria	AUT	60	18,480
	France *	FRA	73	22,484
E	Belgium *	BEL	404	124,432
	Slovakia	SVK	570	175,560
	Romania	ROU	848	261,184
	Ukraine	UKR	801	246,708
F	Turkey	TUR	99	30,492
	Georgia	GEO	353	108,724
	Portugal *	POR	780	240,240
	Czech Republic	CZE	824	253,792
		Total:	11,569	3,563,252

Task B2:

a)

To accommodate the new constraint for which matches and venues are fixed, we replace constraint (15) in the model from B1 with the following:

$$x_{M1,MUN} + x_{M14,STU} + x_{M25,FRK} = 3 \tag{15*}$$

To ensure that a match is held in every venue for every date that venue was used in the original solution, we add the following parameter:

$$VenUsed_{v,d} = \begin{cases} 1 & \text{if there is a match at venue } v \text{ on date } d \text{ in the original schedule,} \\ 0 & \text{otherwise} \end{cases}$$

We also add a constraint stating that if a $VenUsed$ equals 1, then the sum of matches in that venue on that date must also equal one

$$\sum_{m \in M} x_{m,v} * MD_{m,d} = VenUsed_{v,d}, \qquad \forall v \in V, \qquad \forall d \in D \tag{16}$$

Note that with this new constraint (16) added, constraint (8) from the model in B1 becomes redundant.

Solving the new model in AMPL gives the following optimal schedule:

	14.6	15.6	16.6	17.6	18.6	19.6	20.6	21.6	22.6	23.6	24.6	25.6	26.6
Berlin		M3 (ESP-CRO)						M20 (NED-FRA)				M31 (NED-AUT)	
Leipzig					M12 (POR-CZE)			M19 (POL-AUT)			M27 (ALB-ESP)		
Hamburg			M7 (POL-NED)			M15 (CRO-ALB)			M23 (TUR-POR)				M35 (GEO-POR)
Dortmund		M4 (ITA-ALB)			M11 (TUR-GEO)				M24 (GEO-CZE)			M32 (FRA-POL)	
Gelsen-kirchen			M6 (SVN-DEN)				M16 (ESP-ITA)						M36 (CZE-TUR)
Dusseldorf				M9 (BEL-SVK)				M21 (SVK-UKR)			M28 (CRO-ITA)		
Cologne		M2 (HUN-SUI)				M13 (SCO-SUI)			M22 (BEL-ROU)			M29 (ENG-SVN)	
Frankfurt				M10 (ROU-UKR)			M17 (DEN-ENG)			M25 (SUI-GER)			M33 (SVK-ROU)
Stuttgart			M5 (SRB-ENG)			M14 (GER-HUN)				M26 (SCO-HUN)			M34 (UKR-BEL)
Munich	M1 (GER-SCO)			M8 (AUT-FRA)			M18 (SVN-SRB)					M30 (DEN-SRB)	

Groups: A B C D E F

The new total distance is 13,503 km, a reduction of 1,270 km compared to the original schedule.

The new estimated total CO2-emissions are 4,158,924 kg, a reduction of 391,160 kg compared to the original schedule.

It is worth noting that although the distance traveled - and by extension CO2 emissions - is still lower than in the original schedule, the objective function value is worse than the solution from B1. This is to be expected as the model was severely restricted.

b)

In our solution to B2 **Slovenia** is the team that travels the longest distance, at **1282 km** and **Italy** travels the shortest distance at **93 km**. A difference of **1,189 km**.

Task B3:

a)

The Euro 2024 tournament has two objectives that they could want to minimize: let's define them as *unfairness* (measured by the maximum difference between team's traveling distances) and *distance* (measured by the total traveling distance in kilometers across all teams).

To minimize *unfairness,* we can modify the model from B2 in the following way:

First, we add three variables:

d_t: the total distance travelled by team t for a given schedule:

$$d_t = \sum_{v_1 \in V, v_2 \in V} (z^1_{t,v_1,v_2} + z^2_{t,v_1,v_2}) d_{v_1,v_2}, \qquad \forall t \in T$$

$MaxD$: the maximum distance travelled by any team for a given schedule

$MinD$: the minimum distance travelled by any team for a given schedule

To assign values to $MaxD$ and $MinD$, we add the following constraints:

$$d_t \leq MaxD, \qquad \forall t \in T \tag{17}$$

$$d_t \geq MinD \geq 0, \qquad \forall t \in T \tag{18}$$

Finally, we change the objective function to:

$$\min uf = MaxD - MinD$$

Solving the model gives the following *distance* and *unfairness* in the optimal solution:

Distance: 14,360 km

Unfairness: 696 km

In this new solution **Scotland** is the team that travels the longest distance, at **991 km** and **Switzerland** travels the shortest distance at **295 km**. A difference of **696 km**.

b)

To see if there is a match schedule that keeps the optimal *distance* from 2a and well as the optimal *unfairness* from 3a, we resolve the model to minimize *unfairness* subject to the additional constraint that the *distance* cannot exceed 13,503 km:

$$\sum_{t \in T} d_t \leq 13{,}503 \qquad (19)$$

Solving the model gives the following *distance* and *unfairness* in the optimal solution:

Distance: 13,503

Unfairness: 795

Hence, it is not possible to maintain optimal *unfairness* and *distance* simultaneously. (as 795 > 696)

However, by first finding the optimal *distance*, fixing it, and then minimizing *unfairness* it is worth noting that we were able to reduce *unfairness* from **1,189** km to **795** km without sacrificing *distance*. Thus, this schedule is strictly better than the one from B2.

To allow for a 5% and 10% larger travelled distance we modify constraint (16) as follows:

$$\sum_{t \in T} d_t \leq 13{,}503 * ScaleFactor \qquad (19^{*})$$

Solving the model for different $ScaleFactors$ yields:

ScaleFactor:	1	1.05	1.1
Distance:	13,503	13,816	14,456
Unfairness:	795	734	696

Notice that the solution with $ScaleFactor = 1.1$ is strictly worse than our solution from B3a (total distance has increased with 96km). However, as the algorithm only seeks to minimize *unfairness*, it neglects extra distance as long as it is below the threshold set in constraint 19. This motivates an alternative objective function formulation that takes both *unfairness* and *distance* into account simultaneously (See Appendix).

Part C

Task C1:

Model formulation:

Sets:

R: Set of refineries [R1, R2]

T: Set of days [0, ... ,10]

I: Set of crude oils [CrA, CrB]

B: Set of components [distilA, distilB, ISO, POL]

P: Set of products [premium, regular, distilF, super]

D: Set of depots [D1, D2]

K: Set of markets [K1, K2, ... K9]

$K_e \subset K$: Subset of extreme markets [ES1, ES2, EN1, EN2]

$K_e^S \subset K_e$: Subset of extreme markets in the south [ES1, ES2]

$K_e^N \subset K_e$: Subset of extreme markets in the north [EN1, EN2]

Parameters:

$C_{i,t}^{CRU}$: Cost of purchasing one unit of crude oil i on day t.

$a_{r,i,b}$: Amount of component b obtained from processing one unit of crude oil i at the refinery r.

$MaxProc_r$: The maximum processing capacity of crude oil per day at refinery r

$C_{r,i}^{DIS}$: Cost of processing one unit of crude oil i at refinery r

C^{INVI}: Cost of storing one unit of any crude oil per day at any refinery.

C^{TRA1}: Cost of transporting one unit of any component from a refinery to the hub

$Q_{b,p}$: Quantum (units) of component b required to produce product p

C_p^{PRO}: Cost of producing one unit of product p

C^{INVB}: Cost of storing one unit of any component per day at the hub.

C_d^{TRA2}: Cost of transporting one unit of any product from the hub to depot d

$C_{d,k}^{TRA3}$: Cost of transporting one unit of any product from depot d to market k

C_{d}^{INVP}: Cost of storing one unit of any product at depot d per day

$\delta_{p,k,t}$: Maximum demand for product p in market k on day t

S_p: Selling price of one unit of product p in all markets.

$CExtreme_d$: Fixed cost incurred if a positive quantity is sent from depot d to extreme market k on a given day

$Izero_{p,d}$: Initial inventory of product p at depot d on day zero

$Ifinal_{p,d}^{prod}$: Final inventory of product p at depot d on day 10

$Ifinal_{b}^{comp}$: Final inventory of component b at the hub on day 10

Decision variables:

$IPUR_{i,r,t}$: # of units of crude oil i to purchase at refinery r on day t

$IPRC_{i,r,t}$: # of units of crude oil i to process at refinery r on day t

$BPRO_{b,p,t}$: # of units of component b used in production of product p at the hub on day t

$BshipH_{b,r,t}$: Total number of components b shipped from refinery r to the hub on day t:

$PshipD_{p,d,t}$: # of product p to ship to depot d on day t:

$PshipK_{p,d,k,t}$: # of units of product p to ship from depot d to market k on day t

$INVI_{i,r,t}$: Units of crude oil i, in storage at refinery r at the end of day t:

$INVB_{b,t}$: Units of component b, in storage at the hub at the end of day t:

$INVP_{p,d,t}$: Units of product p in storage at depot d at the end of day t:

$z_{d,k,t}$: A binary variable that takes value 1 if depot d ships to extreme market k on day t, and 0 otherwise.

Objective function:

We want to maximize total profit over the planning horizon

$$\max \pi$$

We assume that all products shipped to markets will be sold immediately, until the market's demand is satisfied. Thus, our revenue becomes the selling price * total amount of products shipped to markets.

For the regular markets, we get a revenue for all products shipped until and including day 9:

$$\sum_{t \in T: t>0 \text{ } and \text{ } t \leq |T|-2} \sum_{k \in K \setminus \{K_e\}} \sum_{d \in D} \sum_{p \in P} S_p * PshipK_{p,d,k,t}$$

For the extreme markets, we get a revenue for all products shipped until and including day 8:

$$\sum_{t \in T: t>0 \text{ } and \text{ } t \leq |T|-3} \sum_{k \in K_e} \sum_{d \in D} \sum_{p \in P} S_p * PshipK_{p,d,k,t}$$

To obtain the total profit we must subtract all costs:

We incur a ***purchasing cost*** equal to the total units of crude oil purchased multiplied by its purchase price:

$$\sum_{t \in T: t>0} \sum_{i \in I} \sum_{r \in R} IPUR_{i,r,t} * C_{i,t}^{CRU}$$

We incur a ***processing cost*** equal to the total amount of crude oil processed times its processing cost:

$$\sum_{t \in T: \, t>0} \sum_{i \in I} \sum_{r \in R} PRC_{i,r,t} * C_{i,r}^{DIS}$$

We incur a ***production cost*** equals to the total number of products produced times its production price:

$$\sum_{t \in T: \, t>0} \sum_{d \in D} \sum_{p \in P} PshipD_{p,d,t} * C_p^{PRO}$$

We incur a ***transport cost from refinery to hub*** on all components shipped from refineries to the hub each day:

$$C^{TRA1} * \sum_{t \in T: \, t>0} \sum_{r \in R} \sum_{b \in B} BshipH_{b,r,t}$$

We incur a ***transport cost from hub to depot*** multiplied by the total number of products sent from the hub to depots:

$$\sum_{t \in T:\, t>0} \sum_{p \in P} \sum_{d \in D} PshipD_{p,d,t} * C_d^{TRA2}$$

We incur **a *transport cost from depot to markets*** multiplied by the total number of products sent from the depots to markets:

$$\sum_{t \in T:\, t>0} \sum_{d \in D} \sum_{k \in K} \sum_{p \in P} PshipK_{p,d,k,t} * C_{d,k}^{TRA3}$$

We incur a ***storage cost at the refinery*** on all crude oils that are in storage at the end of each day:

$$C^{INVI} * \sum_{t \in T:\, t>0} \sum_{i \in I} \sum_{r \in R} INVI_{i,r,t}$$

We incur a ***storage cost at the hub*** on all components that are in storage at the end of each day:

$$C^{INVB} * \sum_{t \in T:\, t>0} \sum_{b \in B} INVB_{b,t}$$

We incur a **storage cost at the depots** on all products that are in storage at each depot at the end of each day:

$$\sum_{t \in T:\, t>0} \sum_{p \in P} \sum_{d \in D} INVP_{p,d,t} * C_d^{INVP}$$

We incur an extra **fixed cost** for each depot for each day if any products are shipped to an extreme market on that day:

$$\sum_{t \in T:\, t>0} \sum_{k \in K} \sum_{d \in D} CExtreme_d * z_{d,k,t}$$

In summary, the objective function to be maximized can be expressed as:

$$\sum_{t\in T: t>0 \ and \ t\leq |T|-2} \sum_{k\in K\setminus\{K_e\}} \sum_{d\in D} \sum_{p\in P} S_p * PshipK_{p,d,k,t} \quad \text{(Revenue regular)}$$

$$+ \sum_{t\in T: t>0 \ and \ t\leq |T|-3} \sum_{k\in K_e} \sum_{d\in D} \sum_{p\in P} S_p * PshipK_{p,d,k,t} \quad \text{(Revenue extreme)}$$

$$- \sum_{t\in T: t>0} \sum_{i\in I} \sum_{r\in R} IPUR_{i,r,t} * C_{i,t}^{CRU} \quad \text{(Purchasing cost)}$$

$$- \sum_{t\in T: t>0} \sum_{i\in I} \sum_{r\in R} PRC_{i,r,t} * C_{i,r}^{DIS} \quad \text{(Processing cost)}$$

$$- \sum_{t\in T: t>0} \sum_{d\in D} \sum_{p\in P} PshipD_{p,d,t} * C_p^{PRO} \quad \text{(Production cost)}$$

$$- C^{TRA1} * \sum_{t\in T: t>0} \sum_{r\in R} \sum_{b\in B} BshipH_{b,r,t} \quad \text{(Transport to hub)}$$

$$- \sum_{t\in T: t>0} \sum_{p\in P} \sum_{d\in D} PshipD_{p,d,t} * C_d^{TRA2} \quad \text{(Transport to depot)}$$

$$- \sum_{t\in T: t>0} \sum_{d\in D} \sum_{k\in K} \sum_{p\in P} PshipK_{p,d,k,t} * C_{d,k}^{TRA3} \quad \text{(Transport to markets)}$$

$$- C^{INVI} * \sum_{t\in T: t>0} \sum_{i\in I} \sum_{r\in R} INVI_{i,r,t} \quad \text{(Storage at refinery)}$$

$$- C^{INVB} * \sum_{t\in T: t>0} \sum_{b\in B} INVB_{b,t} \quad \text{(Storage at hub)}$$

$$- \sum_{t\in T: t>0} \sum_{p\in P} \sum_{d\in D} INVP_{p,d,t} * C_d^{INVP} \quad \text{(Storage at depots)}$$

$$- \sum_{t\in T: t>0} \sum_{k\in K} \sum_{d\in D} CExtreme_d * z_{d,k,t} \quad \text{(Fixed cost)}$$

$$= \pi \quad \text{(Total profit over planning horizon)}$$

Subject to:

Because you cannot store components at the refineries, the components shipped to the hub must equal the total number of components derived from processing crude oils at each refinery each day:

$$BshipH_{b,r,t} = \sum_{i \in I} IPRC_{i,r,t} * a_{r,i,b}\,, \qquad \begin{matrix} \forall b \in B \\ \forall r \in R \\ \forall t \in T \end{matrix} \tag{1}$$

Because you cannot store products at the hub, all produced products must be shipped on the same day. This is the case if the sum of components used for each product matches the number of that product shipped to all depots each day.

$$\sum_{b \in B} BPRO_{b,p,t} = \sum_{d \in D} PshipD_{p,d,t}\,, \qquad \begin{matrix} \forall p \in P \\ \forall t \in T \end{matrix} \tag{2}$$

For each product produced, you must use exactly the specified amount of each component that goes into each unit of that product:

$$BPRO_{b,p,t} = Q_{b,p} * \sum_{d \in D} PshipD_{p,d,t}\,, \qquad \begin{matrix} \forall b \in B \\ \forall p \in P \\ \forall t \in T \end{matrix} \tag{3}$$

Balance constraint for the final inventory of crude oil i at refinery r on day t:

$$INVI_{i,r,t} = INVI_{i,r,t-1} + IPUR_{i,r,t} - IPRC_{i,r,t}, \qquad \begin{matrix} \forall i \in I \\ \forall r \in R \\ \forall t \in T: t > 0 \end{matrix} \tag{4}$$

Balance constraint for the final inventory of component b at the hub on day t:

$$INVB_{b,t} = INVB_{b,t-1} + \sum_{r \in R} BshipH_{b,r,t-1} - \sum_{p \in P} BPRO_{b,p,t}\,, \qquad \begin{matrix} \forall b \in B \\ \forall t \in T: t > 0 \end{matrix} \tag{5}$$

Balance constraint for the final inventory of product p at depot d on day t:

$$INVP_{p,d,t} = INVP_{p,d,t-1} + PshipD_{p,d,t-1} - \sum_{k \in K} PshipK_{p,d,k,t}\,, \qquad \begin{matrix} \forall d \in D \\ \forall p \in P \\ \forall t \in T: t > 0 \end{matrix} \tag{6}$$

If the number of products shipped from a depot to an extreme market on a given day is zero, z must be zero:

$$\sum_{p \in P} PshipK_{p,d,k,t} \geq z_{d,k,t}, \qquad \begin{matrix} \forall d \in D \\ \forall k \in K_e \\ \forall t \in T \end{matrix} \tag{7}$$

For the number of products shipped to extreme markets to be positive, z must equal 1:

$$\sum_{p\in P} PshipK_{p,d,k,t} \leq M * z_{d,k,t}, \qquad M = \sum_{p\in P} \delta_{p,k,t+2}, \qquad \begin{array}{c}\forall d \in D\\ \forall k \in K_e\\ \forall t \in T: t \leq |T| - 3\end{array} \tag{8}$$

We set the upper bound M equal to the maximum possible value of $PshipK$. This is the sum of demand in extreme markets for all products in two days (ref constraint 18).

You cannot process more of crude oil i than what you have available: either from purchased quantity, or in inventory on a given day:

$$IPRC_{i,r,t} \leq IPUR_{i,r,t} + INVI_{i,r,t-1}, \qquad \begin{array}{c}\forall i \in I\\ \forall r \in R\\ \forall t \in T: t > 0\end{array} \tag{9}$$

Daily processing cannot exceed max capacity at any refinery for any day:

$$\sum_{i\in I} IPRC_{i,r,t} \leq MaxProc_r, \qquad \begin{array}{c}\forall r \in R\\ \forall t \in T\end{array} \tag{10}$$

On a given day, each refinery cannot ship more components than the maximum amount that it can produce from processing its available supply of crude oils:

$$BshipH_{b,r,t} \leq \sum_{i\in I}((INVI_{i,r,t-1} + IPUR_{i,r,t}) * a_{r,i,b}), \qquad \begin{array}{c}\forall r \in R\\ \forall b \in B\\ \forall t \in T: t > 0\end{array} \tag{11}$$

On a given day, the hub cannot use more components in production than its available supply of components:

$$\sum_{p\in P} BPRO_{b,p,t} \leq INVB_{b,t-1} + \sum_{r\in R} BshipH_{b,r,t-1}, \qquad \begin{array}{c}\forall b \in B\\ \forall t \in T: t > 0\end{array} \tag{12}$$

On a given day, the hub cannot ship more of a given product than its maximum production of that product from its available supply of components:

$$\sum_{d\in D} PshipD_{p,d,t} \leq \sum_{b\in B}(INVB_{b,t-1} + \sum_{r\in R} BshipH_{b,r,t-1}) * Q_{b,p}, \qquad \begin{array}{c}\forall p \in P\\ \forall t \in T: t > 0\end{array} \tag{13}$$

On a given day, each depot cannot ship more of a given product to markets, than its available supply of that product:

$$\sum_{k\in K} PshipK_{p,d,k,t} \leq INVP_{p,d,t-1} + PshipD_{p,d,t-1}, \qquad \begin{array}{c}\forall p \in P\\ \forall d \in D\\ \forall t \in T: t > 0\end{array} \tag{14}$$

The final inventory of product p at depot d should be enough to anticipate future demand:

$$INVP_{p,d,t} \geq Ifinal_{p,d}^{prod}, \qquad \begin{matrix} t = 10 \\ \forall p \in P \\ \forall d \in D \end{matrix} \tag{15}$$

The final inventory of component b at the hub on day 10 must meet requirements:

$$INVB_{b,t} \geq Ifinal_{b}^{comp}, \qquad \begin{matrix} t = 10 \\ \forall b \in B \end{matrix} \tag{16}$$

The total number of products shipped to market k on a given day cannot exceed the market demand when it arrives:

$$\sum_{d \in D} PshipK_{p,d,k,t-1} \leq \delta_{p,k,t}, \qquad \begin{matrix} \forall p \in P \\ \forall k \in K \setminus K_e \\ \forall t \in T: t > 0 \end{matrix} \tag{17}$$

$$\sum_{d \in D} PshipK_{p,d,k,t-1} \leq \delta_{p,k,t+1}, \qquad \begin{matrix} \forall p \in P \\ \forall k \in K_e \\ \forall t \in T: t > 0 \text{ and } \mathrm{t} \leq |\mathrm{T}| - 2 \end{matrix} \tag{18}$$

(The notation $\forall k \in K \backslash K_e$ can be interpreted as forall k in set K excluding subset K_e)

There should be at least one shipment to each extreme market during the planning horizon:

$$\sum_{t \in T} \sum_{d \in D} z_{d,k,t} \geq 1, \qquad \forall k \in K_e \tag{19}$$

You cannot ship to both extreme north and extreme south on the same day:

$$\left(\sum_{k \in K_e^S} \sum_{d \in D} z_{d,k,t} \right) * \left(\sum_{k \in K_e^N} \sum_{d \in D} z_{d,k,t} \right) = 0, \qquad \forall t \in T \tag{20}$$

The initial balance of $INVI$ and $INVB$ is zero:

$$INVI_{i,r,t} = 0, \qquad \begin{matrix} t = 0 \\ \forall i \in I \\ \forall r \in R \end{matrix} \tag{21}$$

$$INVB_{b,t} = 0, \qquad \begin{matrix} t = 0 \\ \forall b \in B \end{matrix} \tag{22}$$

The initial balance of $INVP$ should be equal to the given parameter:

$$INVP_{p,d,t} = Izero_{p,d}, \qquad \begin{matrix} t = 0 \\ \forall p \in P \\ \forall d \in D \end{matrix} \tag{23}$$

Production and processing cannot start before day one:

$$BPRO_{b,p,t} = 0, \qquad \begin{matrix} t = 0 \\ \forall b \in B \\ \forall p \in P \end{matrix} \tag{24}$$

$$IPRC_{i,r,t} = 0, \quad \begin{matrix} t = 0 \\ \forall r \in R \\ \forall i \in I \end{matrix} \tag{25}$$

Non-negativity for all variables:

$$\begin{gathered} IPUR_{i,r,t}, IPRC_{i,r,t}, BPRO_{b,p,t}, \\ PshipK_{p,d,k,t}, BshipH_{b,t}, PshipD_{p,d,t}, INVI_{i,r,t}, INVB_{b,t}, INVP_{p,d,t} \geq 0 \end{gathered} \tag{26}$$

z is a binary variable:

$$z_{d,k,t} \in \{0,1\}, \qquad \begin{matrix} d \in D \\ \forall k \in K_e \\ \forall t \in T \end{matrix} \tag{27}$$

a)

Solving the above model in AMPL the optimal profit is 2,626,280

There are two shipments to extreme markets, ES1 and ES2. They start on day 3 and day 4, respectively.

There are two shipments to extreme markets EN1 and EN2. They start on day 5 and day 6, respectively.

All shipments to extreme markets are from depot 1.

b)

Adding up for all time periods and all markets, we get the following unsatisfied demand in the optimal solution:

Product:	**Total unsatisfied demand**
Premium	142
Regular	232.82
distilF	108
Super	99

c)

In the optimal solution, we get the following final inventory values for components at the hub, and the following slack variables for constraint (16):

Components:	**Final inventory:**	**Slack:**
distilA	1320.65	1220.65
distilB	1787.75	1687.75
ISO	400	0
POL	400	0

This means that constraint 16 is satisfied for all components, but only the final amount of components ISO and POL are binding in the optimal solution.

By extension, it is as expected that the slack variables for ISO and POL are zero, whilst the slack variables for $distilA$ and $distilB$ is the difference between the final inventory and the required amount, specified by $Ifinal_b^{comp}$.

The reason for the extra final inventory of *distilA* and *distilB* may be understood by looking at parameters $a_{r,i,b}$and $Q_{b,p}$: For all refineries you get a slightly higher amount of *ISO* and *POL* than *distilA* and *distilB* from processing both *CrA* and *CrB*. However, the difference is small compared to the required component input for products, where all products but *distilF* require significantly more *ISO* and *POL* than *distilA* and *distilB*. Thus, to produce enough products to satisfy demand, you need to process a given number of crude oils to obtain enough *ISO* and *POL*, but also get more *distilA* and *distilB* than can be utilized as a byproduct. Because the algorithm still chooses to produce, we can assume that the contribution margin of each extra product sold exceeds the marginal storage cost of its associated byproducts.

Task C2:

The price increase is implemented by changing the relevant values of parameter $C_{i,t}^{CRU}$ in the data file. Resolving the model gives the following optimal profit: 2,608,010

The changes in inventories of crude oils at the end of each day at each refinery are given in the table below:

	t =	1	2	3	4	5	6	7	8	9	10
R1	CrA	0	0	0	0	972.17	314.06	0	0	0	0
	CrB	0	0	0	0	727.83	535.95	0	0	0	0
R2	CrA	0	0	0	0	0.00	0.00	0	0	0	0
	CrB	0	0	0	0	412.02	107.12	0	0	0	0

In the original optimal solution from task C1, there was no buildup of crude oils in either refinery, as there was no incentive to incur the extra storage cost.

However, with the new price increase on day 6 and later, the most economically sound thing to do is now to purchase extra units on day 5 for processing for all upcoming days where the storage cost is less than the difference in purchasing-price:

$$C^{INVI} * \#\ of\ days\ in\ storage < (C_6^{CRU} - C_5^{CRU})$$

$$2 * \#\ of\ days\ in\ storage < 5$$

$$\#\ of\ days\ in\ storage < 2.5$$

Following this logic, *refinery 1* should for instance purchase enough crude oils on day 5, to cover that day's processing, as well as storing the optimal crude oils amounts for processing on day 6 and 7, but not for day 8 and onwards.

Summing up the optimal amount of *CrA* processed at *refinery 1* on day 6 and 7 we get: $563.53 + 408.64 = 972.17$, which is the exact amount observed in the refinery's storage at the end of day 5 in the optimal solution.

Appendix:

A weighted objective function for the model from Task B3:

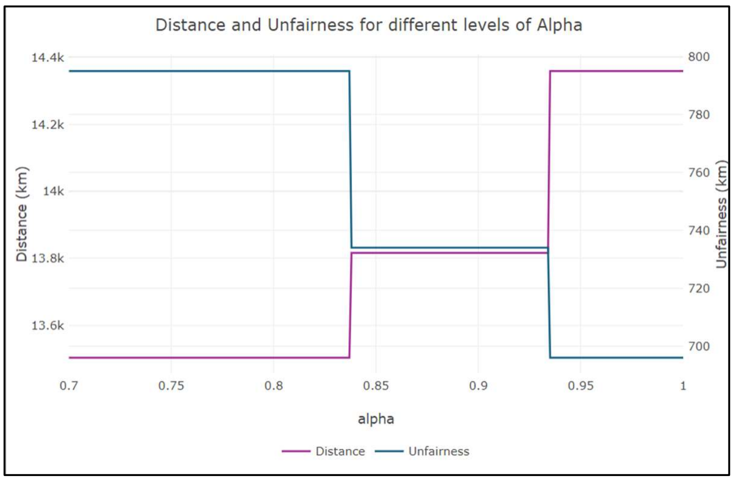
An alternative approach to consider both objectives at the same time, is to incorporate both *unfairness* and *distance* in the objective function, for example weighted by α and $(1-\alpha)$.

We modify the model from B3 by removing constraint (19) and change the objective function to:

$$\min O = \alpha\left(\sum_{t\in T} d_t\right) + (1-\alpha)(MaxD - MinD)$$

The next question is then: what should we use as a weight α?

Mathematically, there is no right answer as to how we should prioritize *distance* against *unfairness*. However, as we expect that there might be diminishing returns both with respect to minimizing *distance* and *unfairness* solving the model for different levels of α and plotting the two objectives might give us some useful insights:



By looking at the plot we see that choosing $\alpha >\approx 0.94$, significantly increases *distance*, only to achieve a comparatively small reduction in *unfairness*. Similarly, by choosing $\alpha <\approx 0.84$ you significantly increase *unfairness* only to slightly reduce *distance*. Thus, we argue that $\alpha \in [0.84, 0.94]$ yields the most balanced solution.

This approach might be even more helpful in linear programming problems with more feasible solutions to consider.

Maximum difference versus _variance_ as a measure of unfairness:

As a final reflection, we would like discuss *variance in distance travelled by teams* as an alternative measure of *unfairness*:

$$Var(d_t): \frac{1}{|T|}\sum_{t\in T}\left(d_t - \bar{d}\right)^2, \qquad \bar{d} = \frac{1}{|T|}\sum_{t\in T} d_t$$

The benefit of minimizing *variance* as opposed to *maximum difference in distances* can be illustrated with the following example:

Distance travelled by team *t* (km):

t =	*ESP*	*GER*	*ITA*	*SUI*	*CZH*	*SCO*
Schedule A	1	500	500	500	500	1000
Schedule B	2	2	998	2	997	1000

If you only consider the maximum differences between teams travelled distances, *Schedule B* is preferred to *Schedule A*:

	Schedule A	**Schedule B**
Total distance	3,001	3,001
Max(d) - min(d)	999	998
Variance (d)	99,800	297,805

However, in order to reduce the difference in distance travelled between Spain and Scotland by one kilometer, the difference in distance travelled between every other team in the league increased dramatically. This mechanism is captured in *variance*, but not *maximum difference* of distances.

Thus, under the assumption that the utility of reducing travel distances is linear, we argue that *variance* of d_t is a better measure of unfairness than the *max difference* in d_t.