2019 User Group Meeting
SATURN 102: Part 1 - Matrix Estimation

November 2019

UNIVERSITY OF LEEDS

FVVB Ltd

ATKINS
Member of the SNC-Lavalin Group

Introduction to Matrix Estimation

Objective:

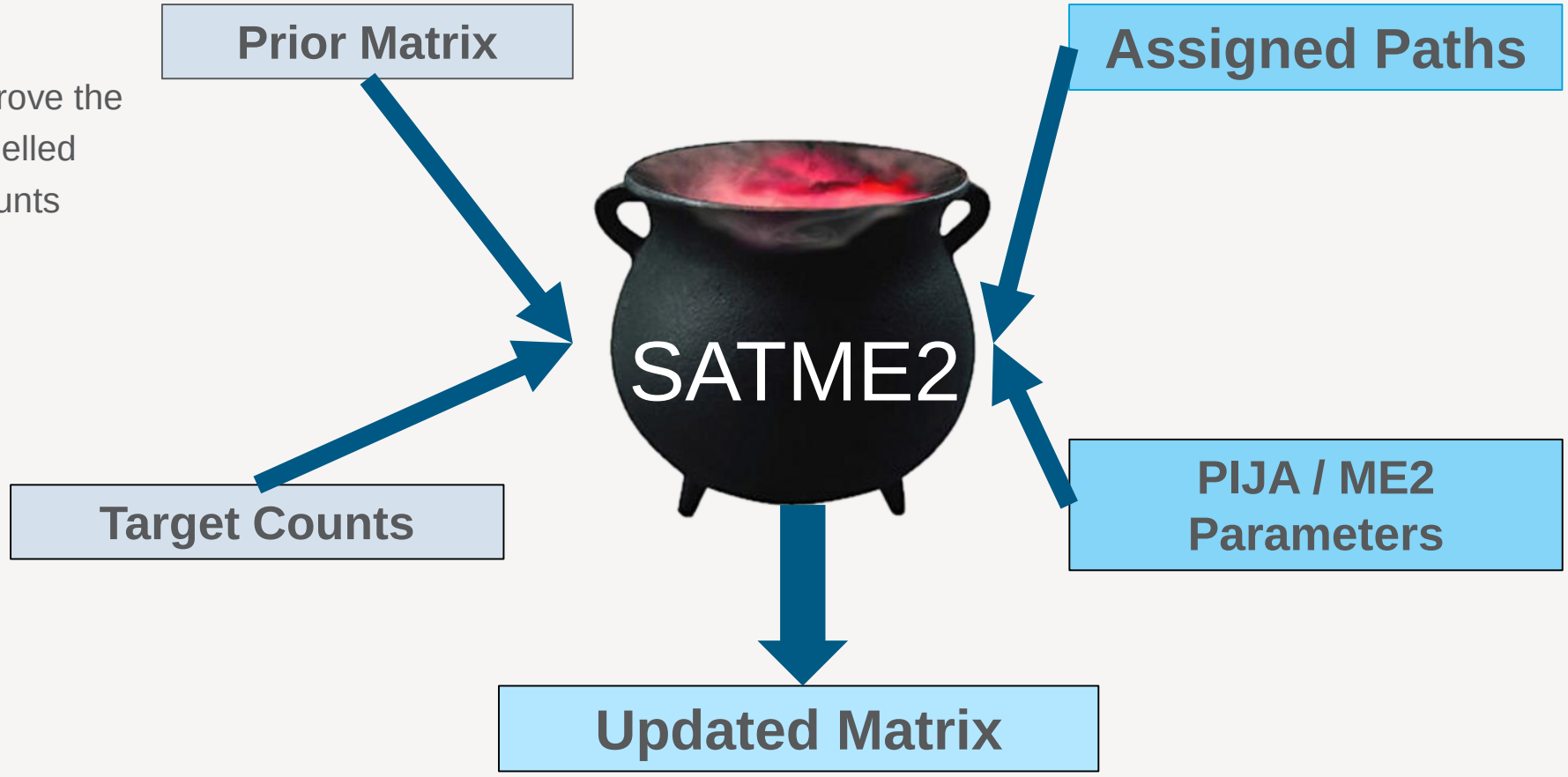
› Refine a prior matrix to improve the replication of assigned modelled flows to observed target counts

Existing Guidance:

› Prior Matrix
› Target Counts

Our focus:

› Assigned Paths
› PIJA / ME2 Parameters



Problems with Matrix Estimation

In practice:

- Historically, mis-used and/or abused
- ‘Garbage in = Garbage Out’

Distortion of underlying travel patterns:

- Trip Length Distribution
- Individual cell
- Trip Ends level

Caused by:

- Deficiencies in the prior trip matrix
- Inaccurate and/or inconsistent counts
- Poorly ‘validated’ highway network

Before

After

https://en.wikipedia.org/wiki/Garbage#/media/File:Vuilnis_bij_Essent_Milieu.jpg (CC BY 1.0)

Transport Appraisal Guidance (TAG)

For Matrix Development

- See TAG Unit M2.2 (Forthcoming change: Nov'19)

For Matrix Estimation

- See TAG Unit M3.1 (May'14), section 8 et seq
- Key features include:
  - (Only) Use robust Count Data for model development
  - Validation of the Prior Trip Matrix (Tests A-D)
  - Validation of Network and Route Choice
  - Then, and only then:
  - Application of Matrix Estimation to:
    - *Refine the model fit by limited adjustments of prior trip matrix*
    - *and take into account the accuracy of the counts*
    - *grouping as necessary*

TAG Unit M3.1 Section 8 Estimation (ii)

Lots of sensible advice in the TAG Unit

- *"The changes brought about by matrix estimation should not be significant. The criteria by which the significance of the changes brought about by matrix estimation may be judged are given in Table 5 (para 8.3.14)."*
- SATME2 .LPM file provides tabulated reports to assist in understanding the changes
- Remember: 'Garbage in = Garbage Out'!

Extract: Table 5

Significance of Matrix Estimation Changes

Measure	Significance Criteria
Matrix zonal cell values	Slope within 0.98 and 1.02 Intercept near zero R^2 in excess of 0.95
Matrix zonal trip ends	Slope within 0.99 and 1.01 Intercept near zero R^2 in excess of 0.98
Trip length distributions	Means within 5% Standard deviations within 5%
Sector to sector level matrices	Differences within 5%

All exceedances of these criteria should be examined and assessed for their importance for the accuracy of the matrices in the Fully Modelled Area or the area of influence of the scheme to be assessed. Where the exceedances are important, the development of the prior matrix should be reconsidered. Where they are not considered to be important, the reasons should be documented in the Model Calibration and Validation Report.

Outliers should also be examined, even when the criteria are met. Explanations about the relevance of the outliers to the intended uses of the model should be included in the Model Calibration and Validation Report.

Assigned Paths

- A quick recap from SATURN 101 last year

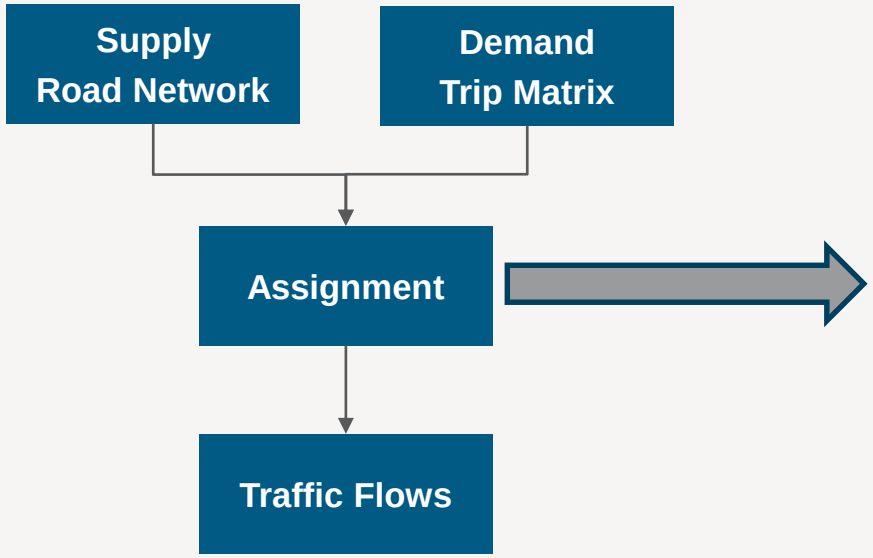
SATURN 101 Series from last year …

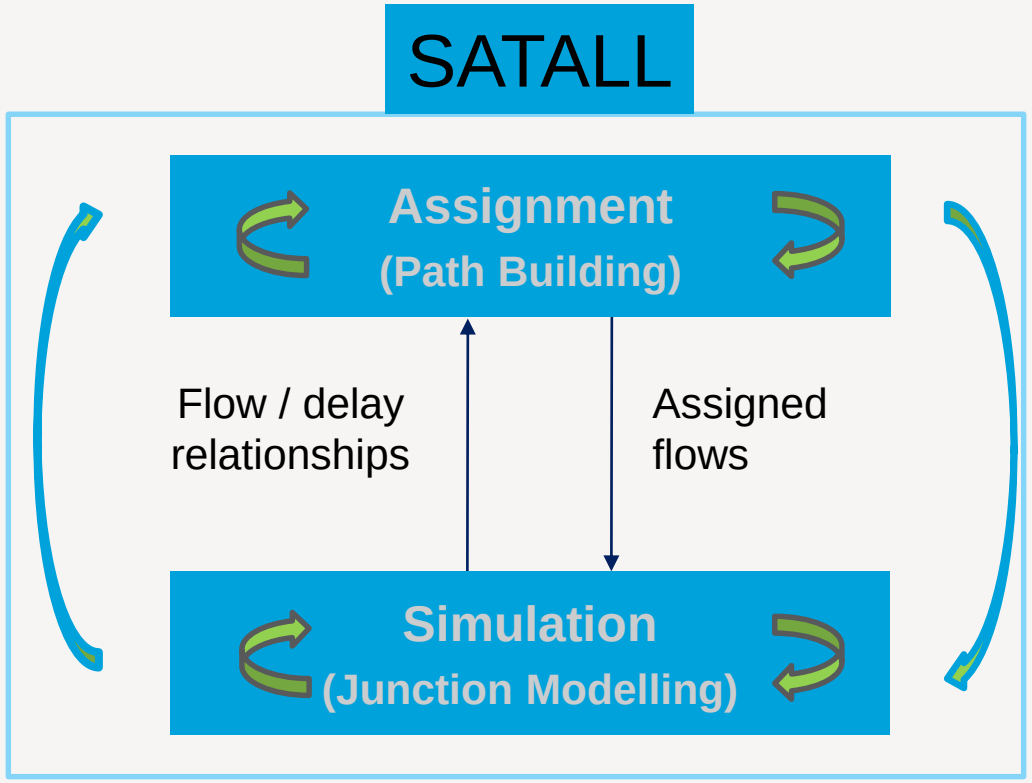
Background Essentials

- Building Blocks
- Path Building

Assignment with Buffer Networks

- Town v Country





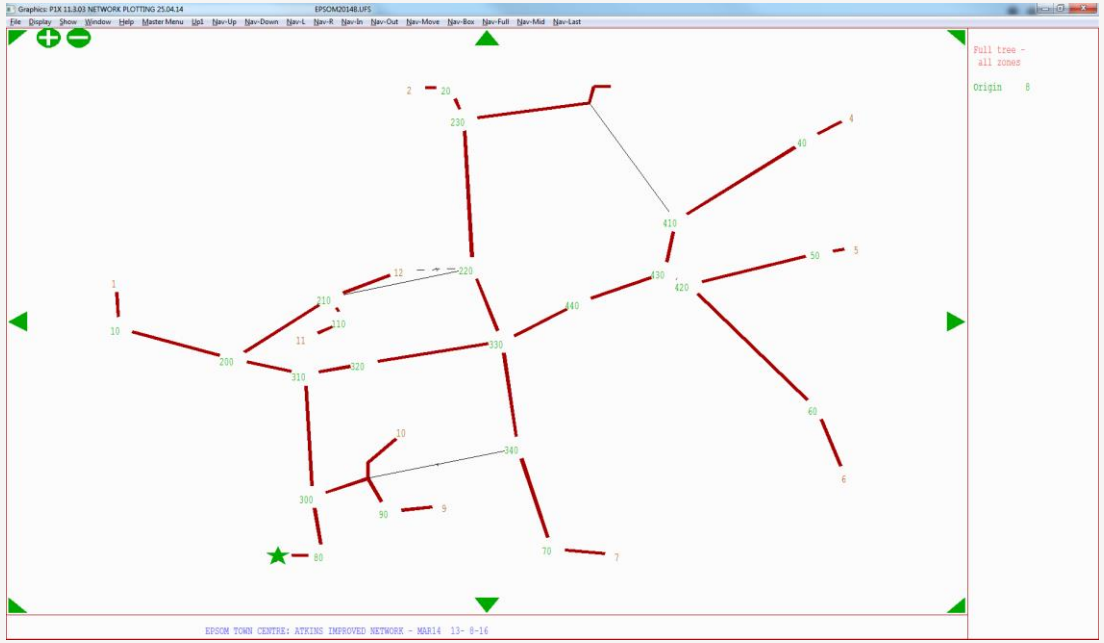
Background Essentials (i)

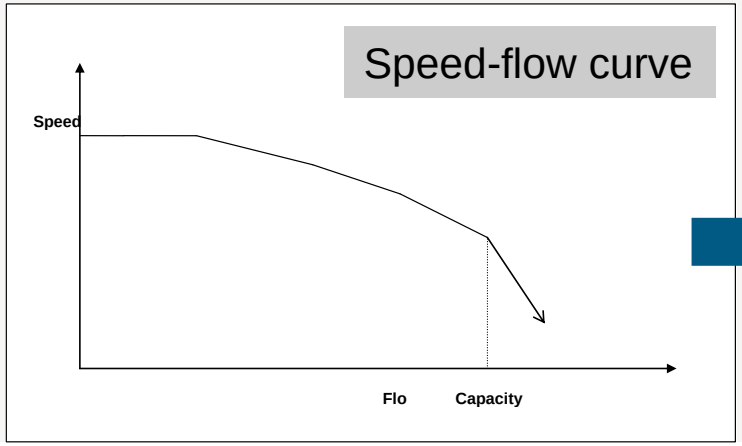
Assignment Trees & Forests

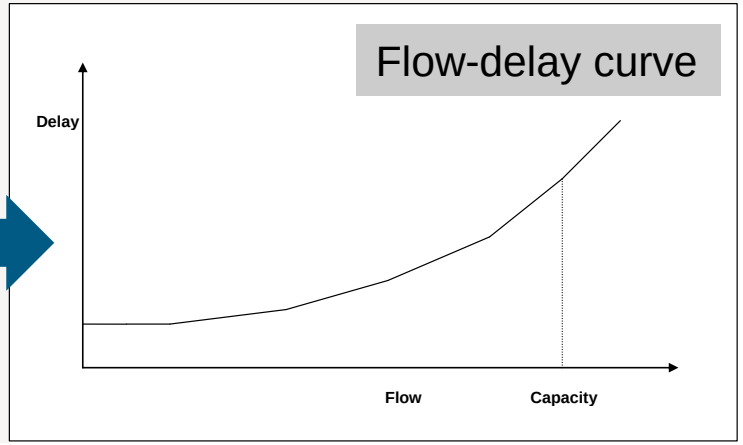
- “Tree” = set of shortest routes from one origin to one (or all) nodes/zones in a network
- “Forest” = collection of trees from a single origin over all assignment iterations

Capacity constraint

- Relationship between vehicle flow and travel time
  - Usually non-linear
  - For example:
    - *COBA-based ‘speed-flow’ curves in Buffer network*
    - *Or more usefully a ‘flow-delay’ curve*







Background Essentials (ii)

Assignment:

› Single All-or-Nothing (AoN) - allocates all the OD-demand to a single route (or ‘path’)

Equilibrium Assignment

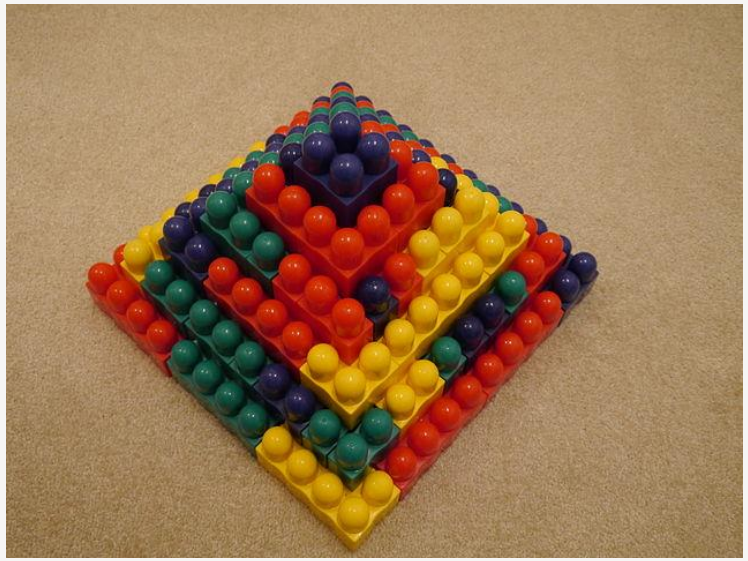
› Series of AoN assignments with paths costs varying through capacity constraint, leading to:

Wardrop Equilibrium

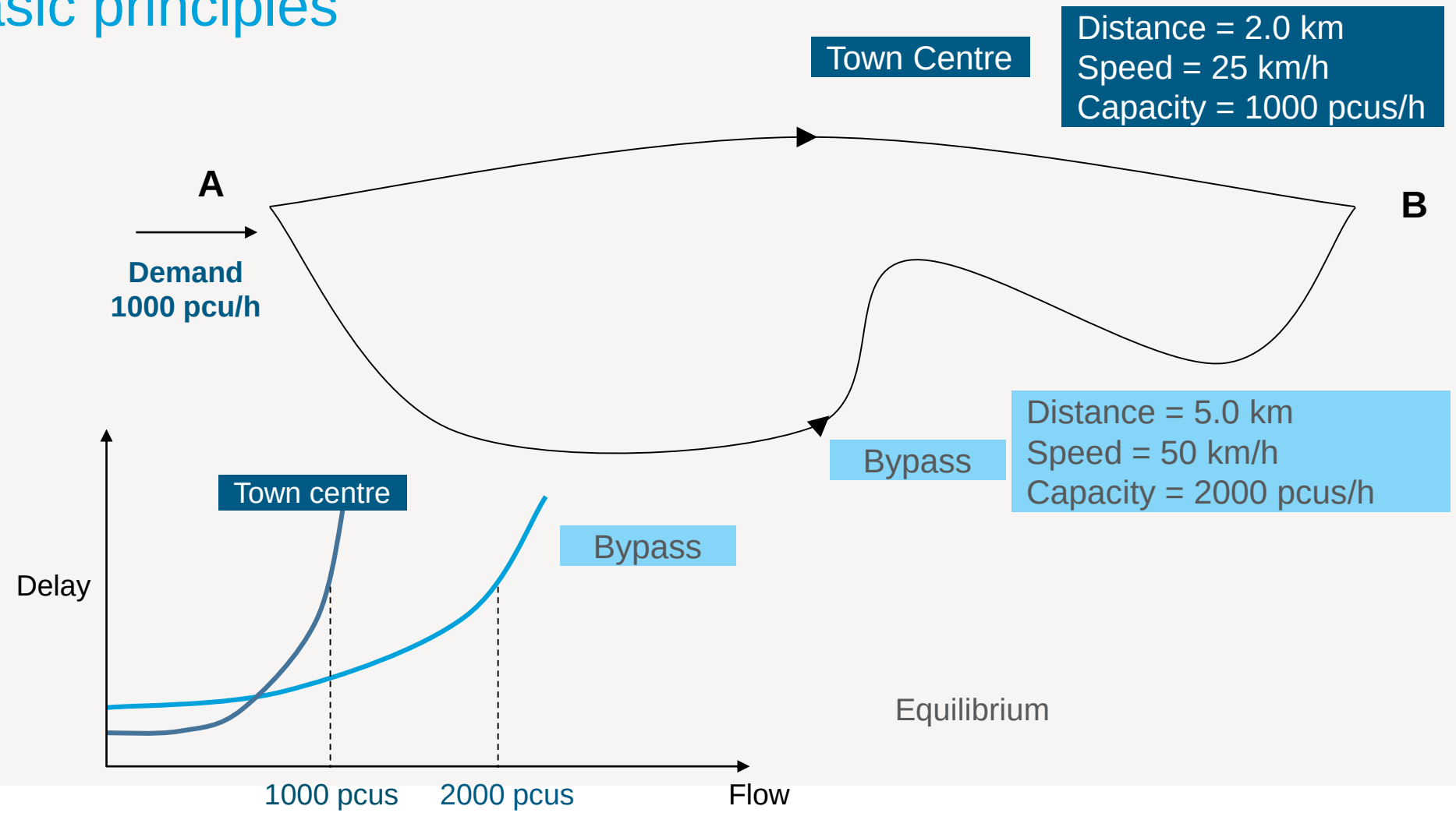
› “Traffic arranges itself on networks such that the cost of travel on all routes used between OD pair is equal to the minimum cost of travel and all unused routes have equal or greater cost” (TAG Unit M3.1)

In SATURN, this mathematical process is undertaken by:

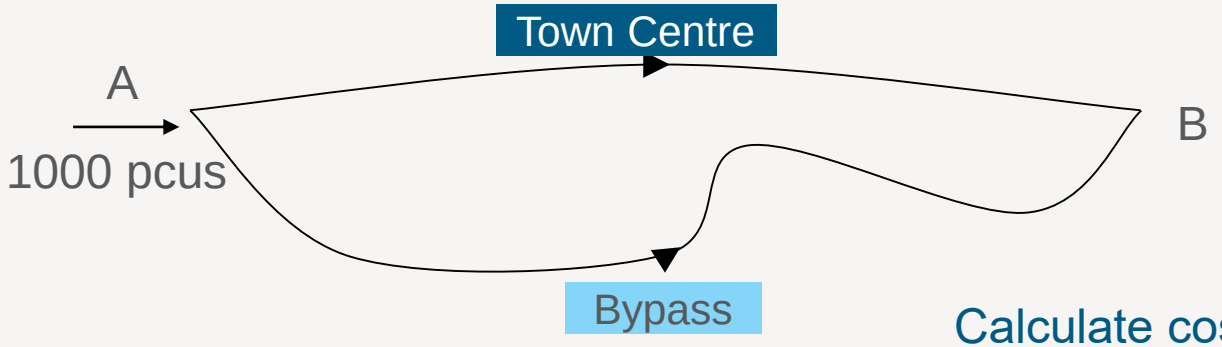
› ‘minimising’ an objective function
› using the Frank-Wolfe algorithm
› to determine the optimum combination (lambda) of the available AoN assignments.

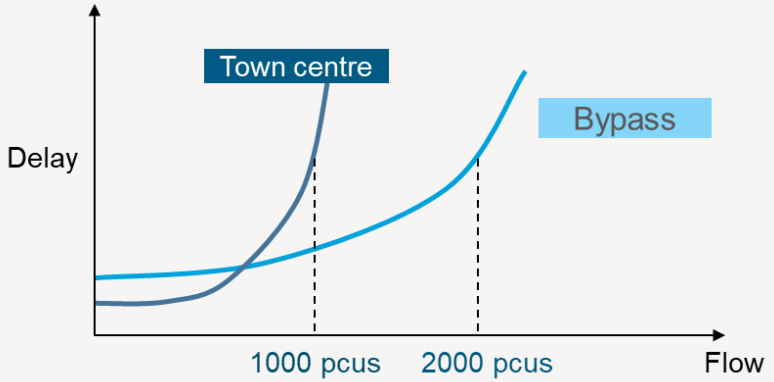


Assignment for a Buffer Network (i) - Basic principles



Assignment for a Buffer Network (ii) - Combining AoN solutions for Equilibrium





Path-building: Successive All-or-Nothings

… allocate 1000 pcus to either Town Centre or Bypass

Path Build	Town Centre	Bypass
1	1000	0
2	0	1000
3	0	1000
4	1000	0

Calculate costs based on flows of …

Iteration	Town Centre	Bypass
1	1000	0
Combine (e.g. 0:100)	(1000)	(0)
2	0	1000
Combine (e.g. 50:50)	(500)	(500)
3	0	1000
Combine (e.g. 66:33)	(333)	(667)
4	1000	0

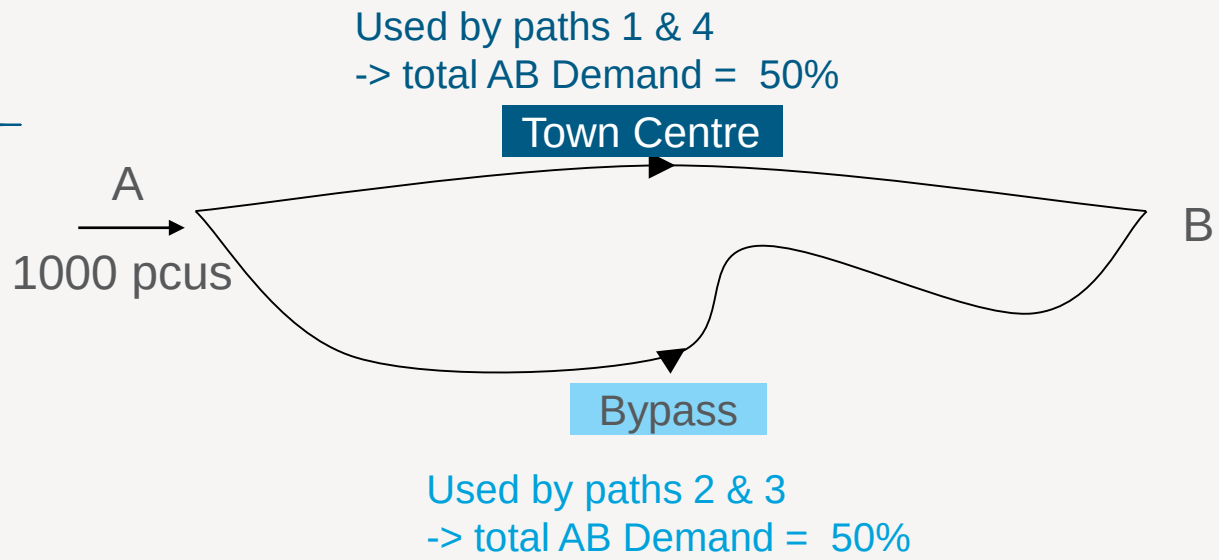
Accumulating the final set of paths

Having calculated the costs based on flows of …

Iteration	Town Centre	Bypass
1	1000	0
Combine (e.g. 0:100)	(1000)	(0)
2	0	1000
Combine (e.g. 50:50)	(500)	(500)
3	0	1000
Combine (e.g. 66:33)	(333)	(666)
4	1000	0
Combine (e.g. 75:25)	(500)	(500)

Now visualise the forest between A, B

- As Method of Successive Averages used, equal weight attached to each iteration
- Link costs based on final combined flows



Matrix Estimation Process

- from section 13 of the User Manual!

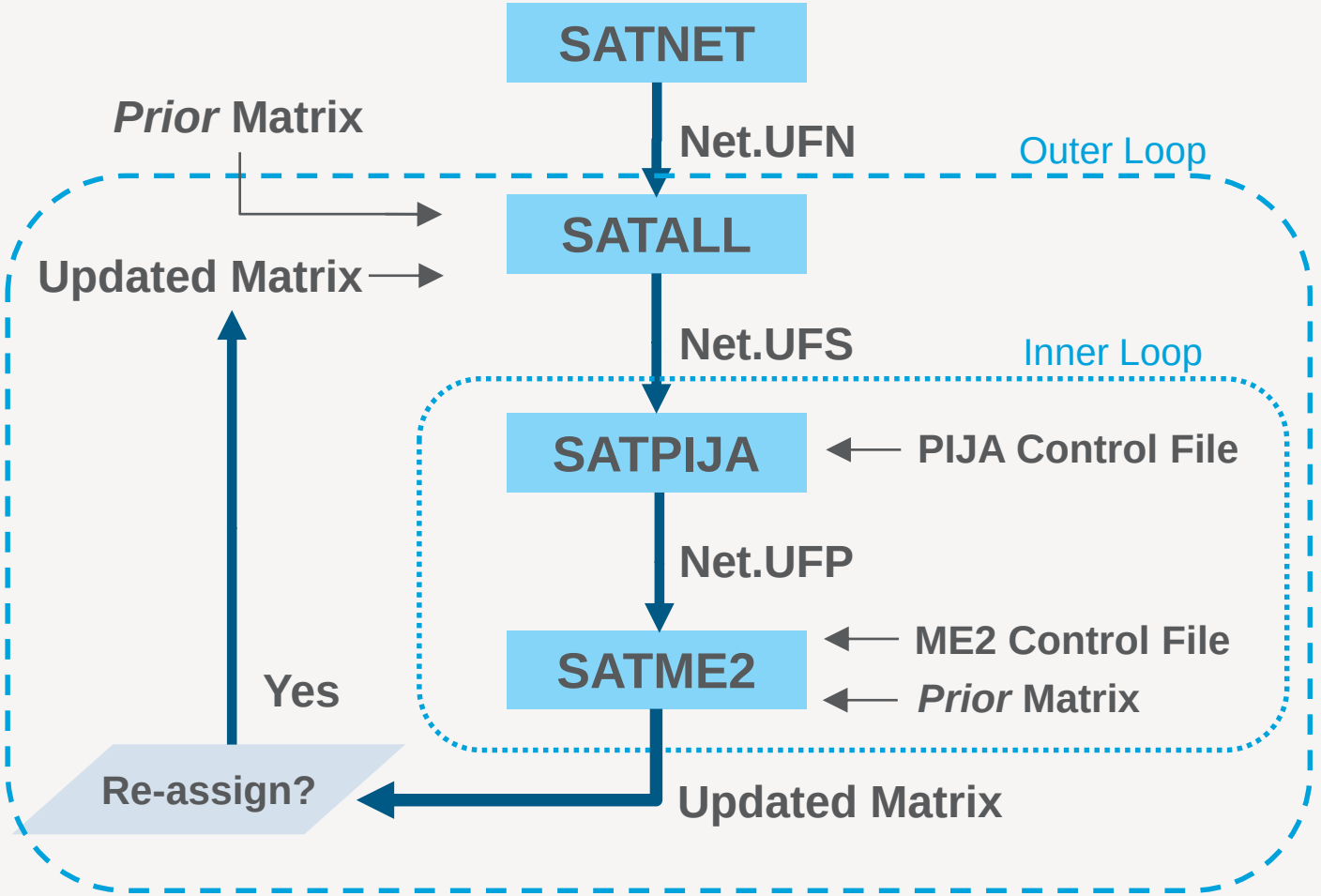
Typical Matrix Estimation Workflow

SATPIJA process

- Simply extracts path information
  - **P**roportion of OD trips (**I**-**J**) using count link **A**

SATME2 process

- Updates prior matrix to match target counts (assuming paths unchanged)
  - **M**atrix **E**stimation by **M**aximum **E**ntropy
  - Undertaken by either by:
    - *Each User Class (ie Level) or*
    - *Multiple User Classes (IVC)*
- Examples shows a single pass through of PIJA-ME2
  - Inner loops for each Level / IVC
  - Typically, six outer loops undertaken



Matrix Estimation Process (i) – The Calculation

Conditions

We have a prior matrix t_{ij} such that

link flow $V_a = \sum t_{ij}\, p_{ija} \neq$ target count W_a

Problem: find the matrix T_{ij} such that

$$\sum T_{ij}\, P_{ija} = W_a$$

And

$\|T_{ij} - t_{ij}\|$ is minimised

i.e. 'New' T_{ij} is as "near" as possible to 'Old' t_{ij}

Matrix Estimation by Maximum Entropy

$$T_{ij} = t_{ij} \prod_a X_a^{Pija}$$

Such that, new link flow $\hat{V}_a$

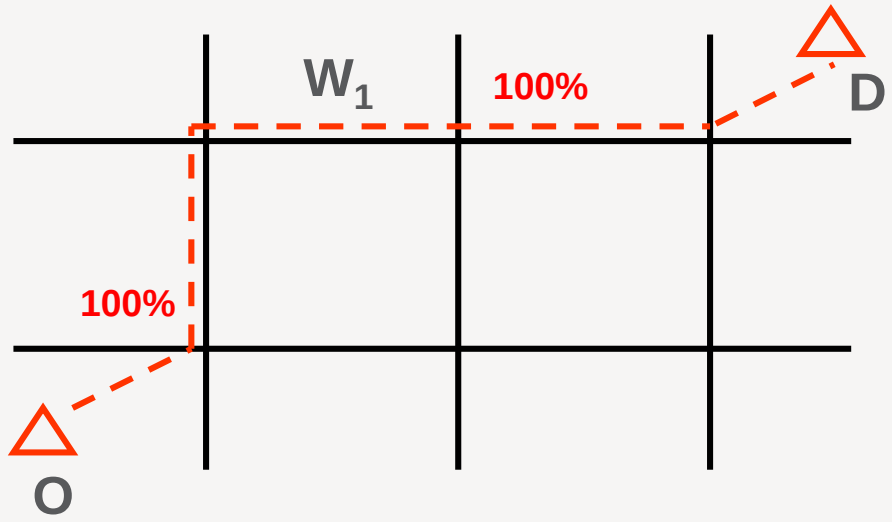
$$\hat{V}_a = \sum_{ij} T_{ij} P_{ija}$$

What's missing?

- The balancing factor X_a for each count
- Calculated by SATME2 …

Matrix Estimation Process (iii) – A Simple Example

Single OD path with one count site



Assuming:

- Target count W_1 = 800 pcu/h
- Prior demand t_{OD} = 1000 pcu/h

Then

- Link Flow V_1 = 1000 pcu/h
- Pija on $link_1$ = 1.00

Hence:

$T_{ij} = t_{ij} \prod_a X_a^{Pija} \rightarrow T_{ij} = 1000\ X_1^{1.0}$

And

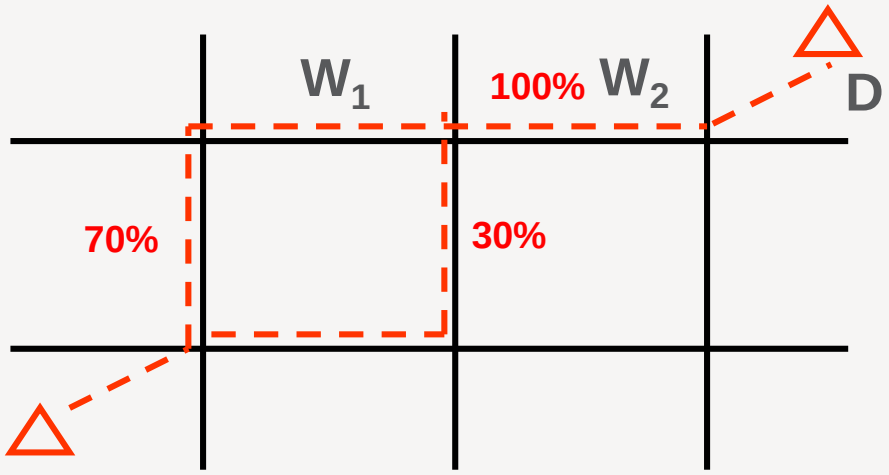
If X_1 = 0.8 then T_{ij} = 800 pcu/h

So

$\hat{V}_a = \sum_{ij} T_{ij} P_{ija} \rightarrow \hat{V}_1 = 800 * 1.00 = W_1$

Matrix Estimation Process (iii) – A More Complicated Example

Single OD pair with two OD paths & two count sites



Assuming:

- Target count W_1 = 800 pcu/h
- Target count W_2 = 1100 pcu/h
- Prior demand t_{OD} = 1000 pcu/h

Then

- Link Flow V_1 = 700 pcu/h
- Link Flow V_2 = 1000 pcu/h
- Pija on $link_1$ = 0.70
- Pija on $link_2$ = 1.00

Hence:

$$T_{ij} = t_{ij}\prod_a X_a^{Pija} \rightarrow T_{ij} = 1000\, X_1^{0.7} X_2^{1.0}$$

… if (say) …

$X_1 = 1.14$ & $X_2 = 1.02$

Therefore:

T_{ij} = 1118 pcu/h

and

$$\hat{V}_a = \sum_{ij} T_{ij} P_{ija} \rightarrow \hat{V}_1 = 1118 * 0.7 = 783 \text{ pcu/h}$$

$$\hat{V}_2 = 1118 * 1.0 = 1118 \text{ pcu/h}$$

SATPIJA 102

- Parameters, Options & Performance

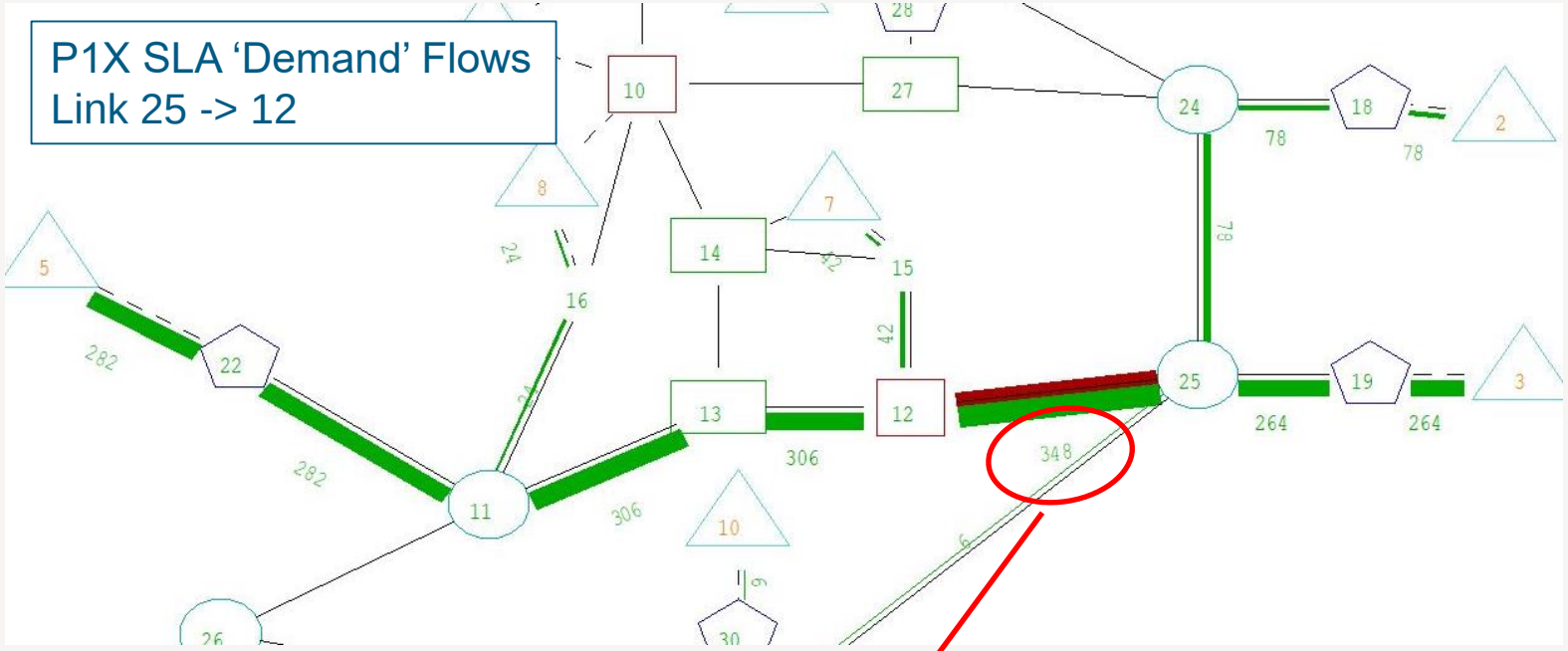
ATKINS
Member of the SNC-Lavalin Group

Extracting Path Information – A Simple Grotley Example

SATPIJA process

- Simply extracts the PIJA factors for each count …
  - Very similar to undertaking separate Select Link Analysis (SLA) for each counted link

- For example, link 25-12
  - Demand Flows
    - *P1X SLA = 348 pcu/h*
    - *SATPIJA = 347.8 pcu/h*



Extract from SATPIJA LPJ File

NO.	A	B	C	COUNT PCU / HR.	TARGET	ASSIGNMENT LINK	SELECTIVE FLOW	SELECTIVE CAPACITY	PER-CENT QUEUED-UP	SAVEIT GEH	FIXED PCU/HR	PASSQ PCU/HR
1	25	12	0	400.0	400.0	141	347.8	952.6	0.0	0.00	0.0	0.0

Understanding SATPIJA – Flow Types

Assigned and Fixed Flows (section 13.1.4.1)

By default (SUBFIX=T)

- Observed counts = Assigned Trips + Fixed Flows (eg bus, PASSQ, PLOD)
- Target count = Observed count – all Fixed Flows
- Option to remove PASSQ flows only – set SUBFIX=F & SUBPQ=T

Demand versus Actual Flows (section 13.1.4.2)

SATME2 adjusts the Prior 'Demand' matrix

- Calculations undertaken using Demand Flows not Actual Flows
- SATPIJA automatically adjusts target count to compensate for upstream losses due to queues
  - Be careful: upstream flow metering may make it difficult / impossible to achieve a demand matrix to match the counts!

Multiple Crossings (section 13.1.8.1)

- If combining counts as screenlines, grouped counts must be on parallel routes not sequential
  - PIJA factors cannot be greater than 1

Extracting Path Information – A More Complicated Example

With Fixed Flows added

- P1X now reports

```
    Downstream Stop Line Flows

Demand          393.6
Arrive          353.3
Queued-up        40.3
Fixed (D)        49.7
Buses (D)        30.0

Capacity        953.1
V/C              37.1 %
```

- LPJ reports (all pcu/h):

1. Bus Flow = 30
2. PASSQ Flow = 19.7
3. Fixed Flow = 49.7
4. (Actual) Count = 400
5. ME2 Target Count = 445.6
   - + 45.6 increase

```
      CALCULATE PIJA FACTORS FOR THE FOLLOWING LINKS OR TURNS:
         AS READ FROM THE CONTROL FILE (SINCE UFFLOW = F)
         WITH FIXED FORMAT WITH DUTCH =  T
         I.E., WITH FORMAT 3I10

   (1) COUNT FLOWS ARE AS INPUT AND ASSUMED TO BE ACTUAL FLOWS

   (2) TARGET FLOWS ARE (ACTUAL) COUNTS FACTORED UP BY QUEUE
       REDUCTION FACTORS TO CONVERT THEM TO DEMAND FLOWS

   (3) SELECTIVE FLOWS AND CAPACITIES ARE TAKEN FROM THE (DEMAND) ASSIGNMENT
       AND FACTORED DOWN TO ACTUAL FLOWS; See 13.3.7.1 in the SATURN Manual.

   (4) PER CENT QUEUED UP MEASURES HOW MUCH TRAFFIC ON THE LINK
       COMES FROM QUEUED TRAFFIC ON THE SIMULATION LINK(S) IMMEDIATELY UPSTREAM

   (5) SAVEIT GEH MEASURES THE DIFFERENCES BETWEEN THE (TOTAL) ASSIGNED DEMAND FLOWS
       AND THE FLOWS AS CALCULATED BY SAVEIT

   (6) TOTAL FIXED FLOWS

   (7) TOTAL PASSQ FLOWS
```

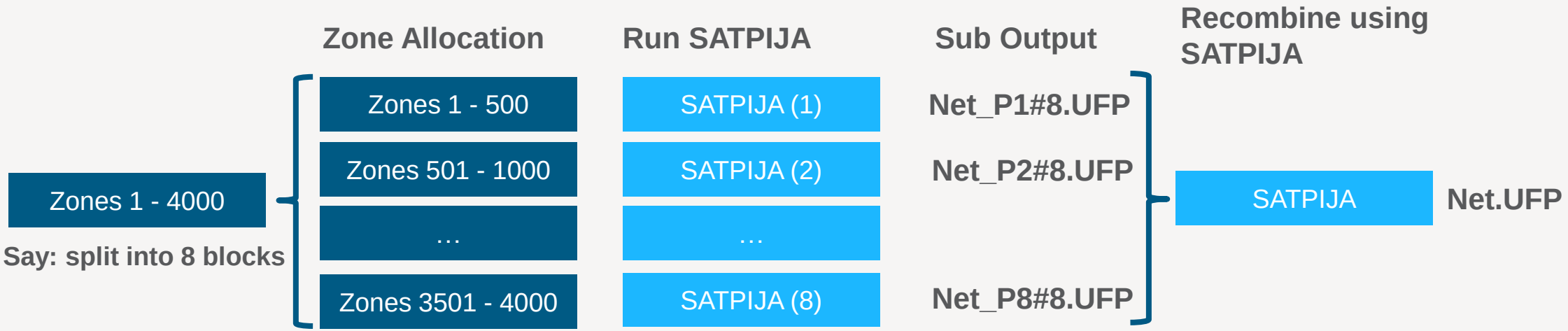
NO.	A	B	C	COUNT PCU / HR. (4)	TARGET (5)	ASSIGNMENT LINK	SELECTIVE FLOW	SELECTIVE CAPACITY (6)	PER-CENT QUEUED-UP	SAVEIT GEH	FIXED PCU/HR (3)	PASSQ PCU/HR (2)
1	25	12	0	400.0	445.6	141	308.7	943.6	76.8	0.00	49.7	19.7

Also check: Count (4) versus Capacity (6)

Speeding-up with distributed SATPIJA_MC

Computationally expensive – need to recreate paths

- For each SATPIJA_MC run, process splits the origins into blocks of zones
  - Run separate SATPIJA for each block to extract path information for all counts & then recombine



Comments:

- Number of blocks set by PC environmental variable *NUMBER_OF_PROCESSORS*
  - By default, maximum # of blocks ('threads') is greater of either 8 or user-defined *SATURNmaxprocessors* value
  - Possible performance improvements with higher values (eg 12+)

SATME2 102

- Internal Processes

ATKINS
Member of the SNC-Lavalin Group

SATPIJA/ME2 Warnings – Kirchoff Violations

Kirchoff's (1st) Law:

› *the sum of currents flowing into that node is equal to the sum of currents flowing out of that node*

Useful for checking traffic counts

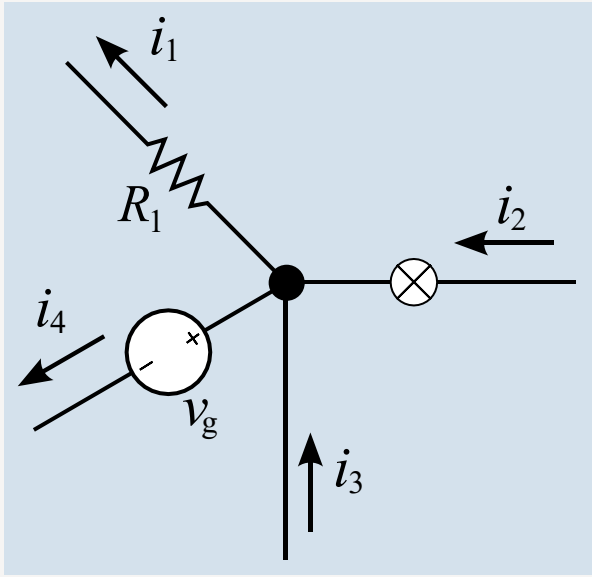
› Inconsistent counts tend to prevent convergence and/or force X_a factors to their minimum / maximum values

› Undertaken on the expanded assignment network

  › Purely mathematical process - does not consider 'real-life' conditions

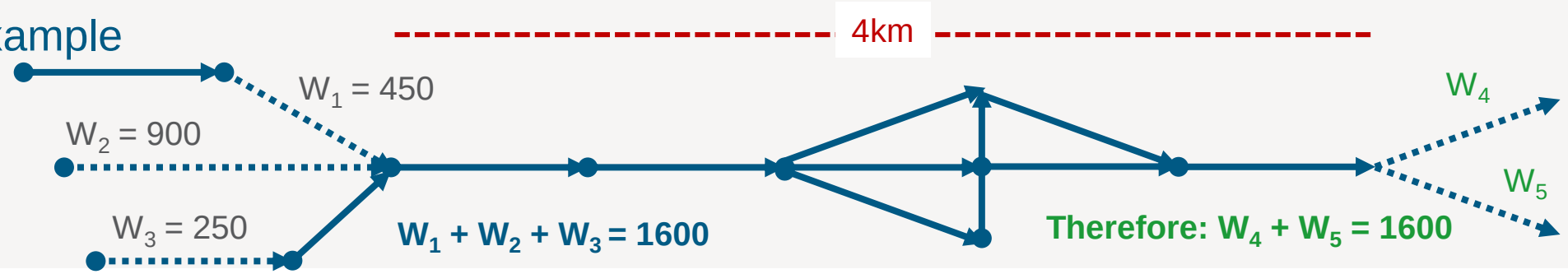
  › Tenacious search process - so may not immediately be obvious

› Don't ignore them as they're flagging a problem: set PRINT = T for more information



By KCL.png: Pflododerivative work: M0tty (talk) - KCL.png
CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=16594174

Skeletal Example



SATPIJA/ME2 Problems – Simulation Centroid Connectors

Bridging counts on Simulation Centroid Connectors (Sim CC)

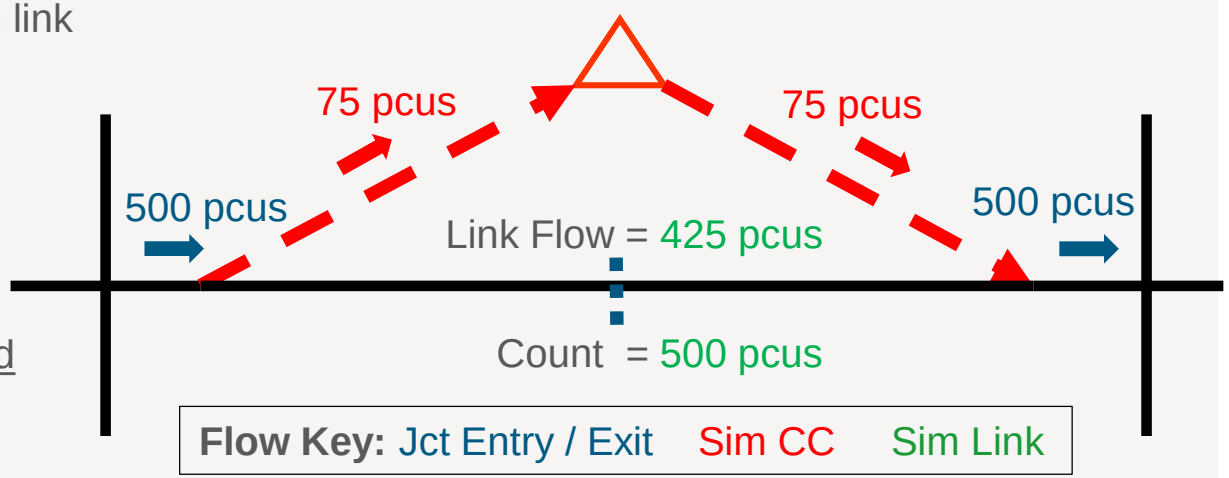
By definition:

- Flow to Sim CC: leaves at the beginning of the simulation link
- Flow from Sim CC: joins at the end of the simulation link

SATPIJA assumes:

- Link Count: taken at the middle of the link
- Actual Flow: mid-link flow ie no Sim CC flows are included

- See Serious Warning



Avoid 'bridged' counts on links with simulation CCs

- If not possible, need to adjust observed count

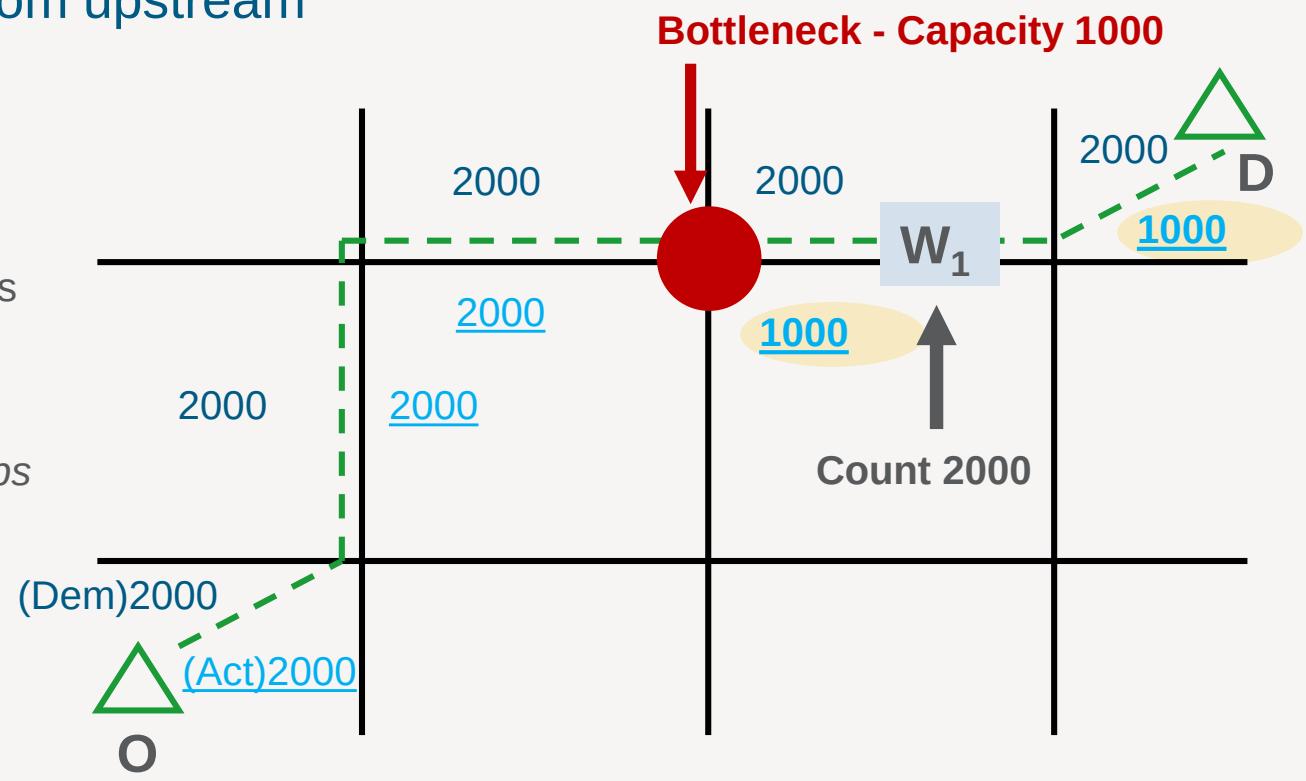
SATPIJA/ME2 Problems – Upstream Flow Metering

Observed Count > permitted discharge from upstream bottleneck

- Generates nasty feedback mechanism

Why?

- ME2 process works on demand flows not actual flows
- Within SATALL – SATME2 looping process
  - For Loop 1
    - *target count adjusted for queues & shortfall in trips*
    - *ME2 increases demand matrix*
  - For Loop 2
    - *Actual Link flow unchanged at 1000 but …*
    - *Increased queues upstream*
    - *So target count increased again to compensate*
  - Exponential growth in demand matrix as further ME2 loops undertaken
  - **Check LPJ for serious warnings !!**



Key: Demand Flows Actual Flows

Matrix Estimation – Parameters, Levels & IVC etc

Key Parameters

Parameter	Description	Default	Comment
ITERMX	Maximum number of ME2 iterations	30	Go high unless runtime a concern – suggest 100 and monitor
EPSILN	Convergence criteria for ME2 process	0.01	Keep low unless runtime a concern
XAMAX	Upper limit on Balancing Factor Xa (and lower 1/XAMAX)	5.0	**Default too high! Set target 2.0 or lower as:** - Greater the value, the greater the maximum / minimum adjustments allowed eg XAMAX = 2 allows changes in range 0.5 to 2.0 - **NB: XAMAX limits are multiplicative for each count passed through**

Using PASSQ (SUBPQ=T) with MUC Assignments

› Total PASSQ flows assumed to be split in proportion to the component UC flows

Using In-Vehicle Class (IVC) rather than Levels (UC)

› Same mathematical approach used – balancing factors calculated based on total IVC flows / pijas

› Prior UC T_{ij} proportions retained

Order of Counts

› Yes, it does make a difference! Counts balanced in order defined in the list so final record undertaken last

SATME2 Reporting: LPM & ME2 Files

SATPIJA .LPJ as previously reported

- Kirchoff Violations, Upstream Capacity Constraints, Bridged Centroid Connectors etc

NO.	A	B	C	COUNT	TARGET PCU / HR.	ASSIGNMENT LINK	S E L E C T I V E FLOW	CAPACITY	PER-CENT QUEUED-UP	SAVEIT GEH	FIXED PCU/HR	PASSQ PCU/HR
1	25	12	0	400.0	445.6	141	308.7	943.6	76.8	0.00	49.7	19.7

SATME2 .LPM

- Summary of Inputs (incl. Frozen Cells, Trip End Constraints etc)
- ME2 calculations
- Pre & Post-ME2 comparisons
  - TAG Unit M3.1 Table 5 reports
  - Trip Length Distributions (if FILTID provided)

SATME2 .ME2 Output File

- Information for every processed count
  - *See section 13.9 for more details*

13.9 "ME2" Output Files

Following **SATURN** 10.5, **SATME2** now automatically outputs an ascii/text file which contains summary data from the ME2 calculations. The intention is that this file may be input to other procedures, for example a spreadsheet or **P1X** (see 11.8.5), for further analyses of the results. The file takes its pathname from the counts control file but with extension .ME2; hence it is referred to as a "ME2" file.

Currently ME2 files contain one record per input count with the following data per count:

- The link A, B and C (if a turn) numbers;
- A single character L, = or G to indicate less than / equal / greater than constraints (13.1.7);
- The input link count (as converted to demand; see 13.1.4);
- The "fixed" flow on that link;
- The (demand) target flow (i.e., with fixed flows etc. removed);
- The (demand) link flow prior to running ME2;
- The (demand) link flow after running ME2;
- The "balancing factor X_a for that link (13.1.1);
- Various link sensitivity statistics as listed in Table 10 (13.3.12);
- The number of the 66666 data set if the link is part of a combined set
- Target flows etc. for combined sets (as above for single constraints).

The end of the data file is indicated by a final record containing (as is standard in **SATURN** data files) 99999.

```
11111
*  A    B    C <=>    COUNT     FIXED    TARGET  PRIOR FL  POST FL       XA      XA-1    DIFF-F    TIJ_IN   TIJ_OUT    DIFF_T    DT/DF     DTSUM  DTSUM/DF     DTIJ2   RATIO3 ACOUNT-IN    ACOUNT KOMBI  TARG666
   25   12    0 =    445.60     55.38    445.60    343.89    444.73      1.30      0.30   100.839   348.000   449.602   101.602     1.008   101.602     1.008   101.602     1.008     400.0     400.0     0      0.00
99999
```

Matrix Estimation
- options for Aggregations

Aggregated Matrix Estimation (i) - Existing

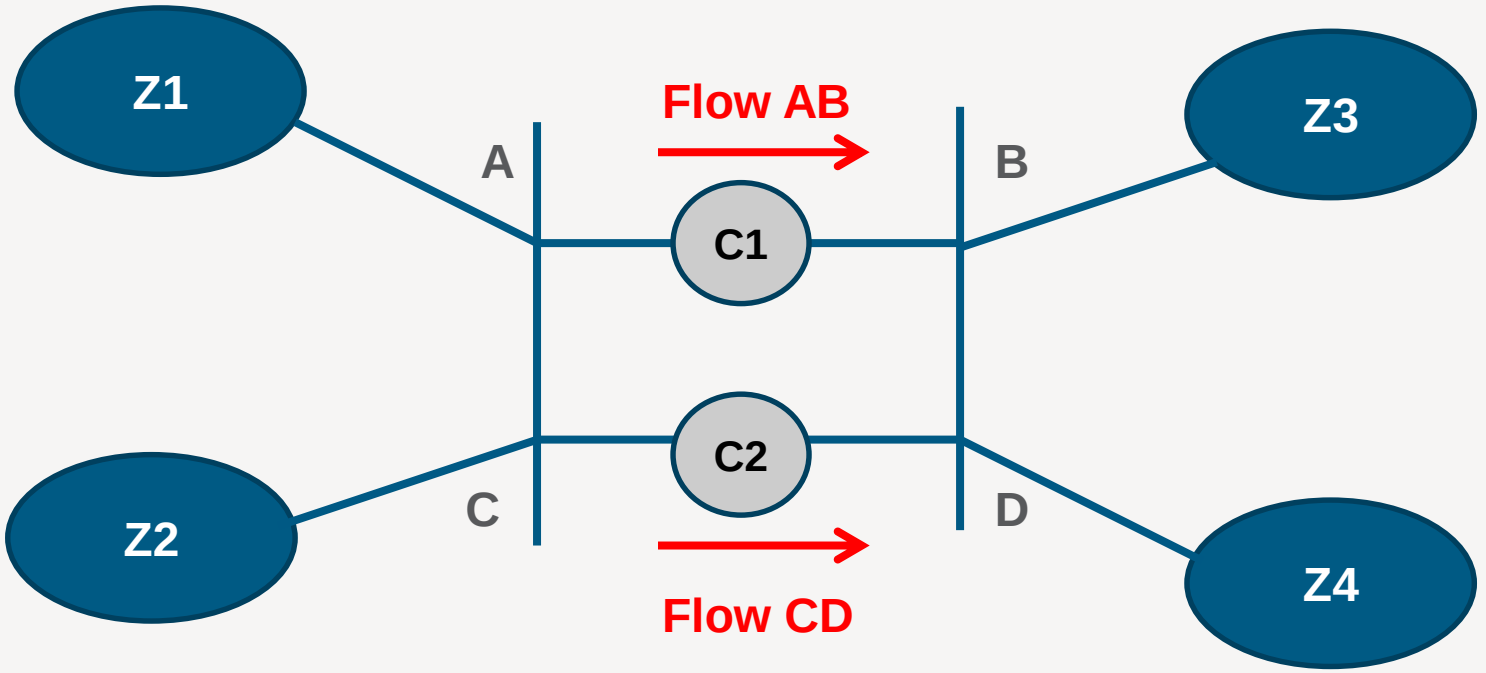
#1 Conventional ME2

› with individual link counts

Assuming robust counts C1, C2

› Differences between
  › Flow AB and count C1
  › Flow CB and count C2

› Assumes route choices are accurate
  › Emphasis on weaknesses in the prior matrix to be addressed by ME2
  › Bold assumption!!



Target Count	Demand Flow	Zone Pairs Adjusted
C1	Flow AB	Z1 or 3 -> Z2 or 4
C2	Flow CD	Z1 or 3 -> Z2 or 4

Aggregated Matrix Estimation (ii) – Screenlines

Conventional ME2

› with screenlines

Combined counts C1 & C2

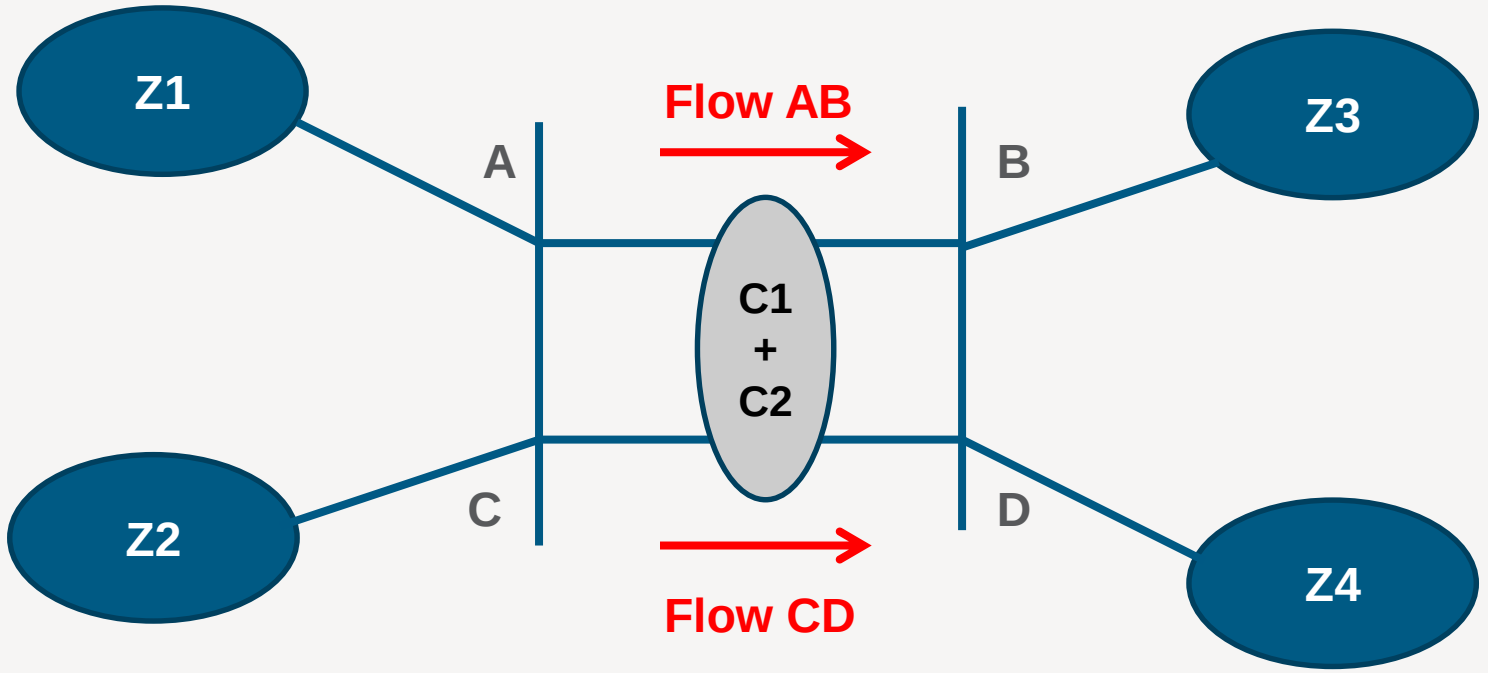
› Acknowledgement that uncertainty in routeing (and flow volumes) and target counts

› As recommended in TAG Unit M3.1

Applied by:

› Screenlines / mini-screenlines
  › Grouped counts in ME2 6666 card

› Judgement required on groupings
  › Specific to location & data
  › See TAG Unit M3.1 section 8



Target Count	Demand Flow	Zone Pairs Adjusted
Total C1 + C2	Total Flow (AB + CD)	Z1 or 3 -> Z2 or 4

Aggregated Matrix Estimation (iii) – Sector-based

Sector-based ME2

› with screenlines

Combined Zones

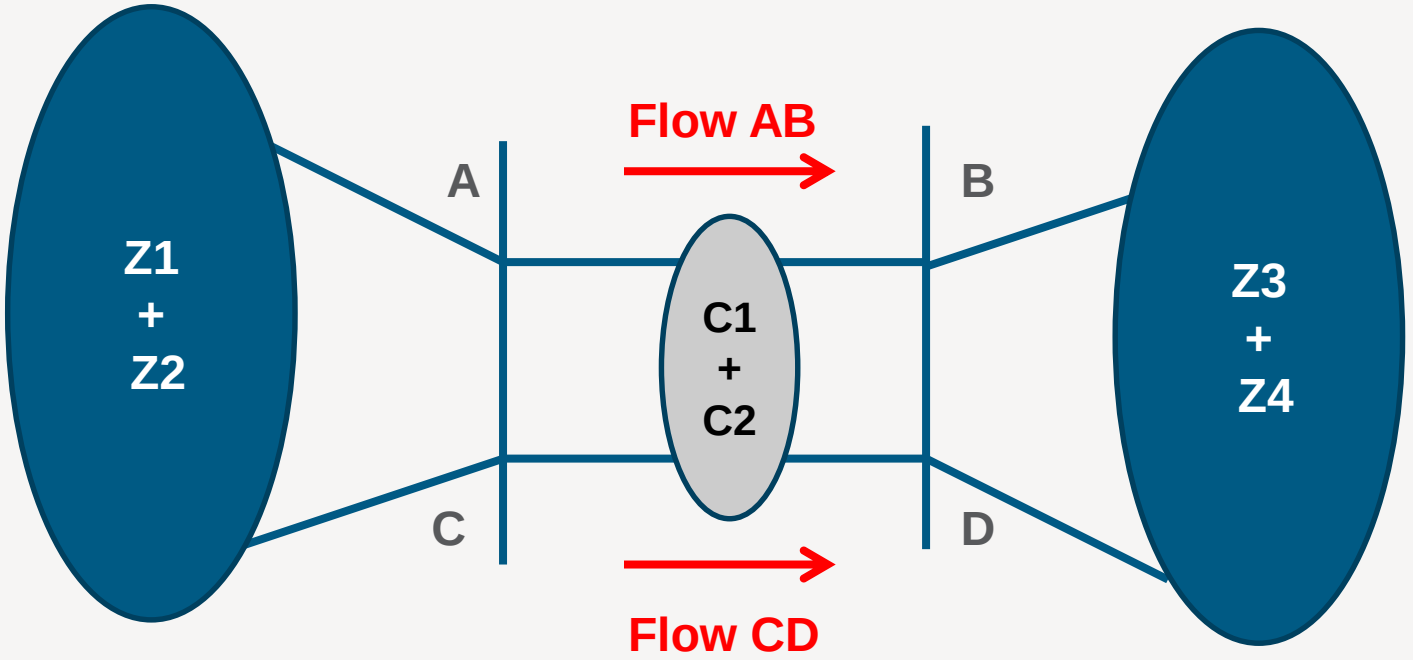
› Acknowledgement that uncertainty in routeing (and flow volumes), target counts and Tij estimates

› New Tij calculated at sector level and retain same zonal proportions (GROUP=T)

Applied by:

› Screenlines / mini-screenlines
  › Grouped counts in ME2 6666 card
  › Grouped zones in .Z2G file
  › Judgement required on groupings

› Introduced in SATURN 11.3
  › Not widely used?



Target Count	Demand Flow	Zone Pairs Adjusted
Total (C1 + C2)	Total Flow (AB + CD)	Total (Z1+Z3) -> Total (Z2+Z4)

Summary

Final Advice with Matrix Estimation

Before thinking of using ME2:

- Read through TAG Unit M3.1, section 8
- Read SATURN User Manual, section 13

Before running ME2, address:

- Deficiencies in the prior trip matrix
- Inaccurate and/or inconsistent counts
- Poorly ‘validated’ highway network

Start small & build-up

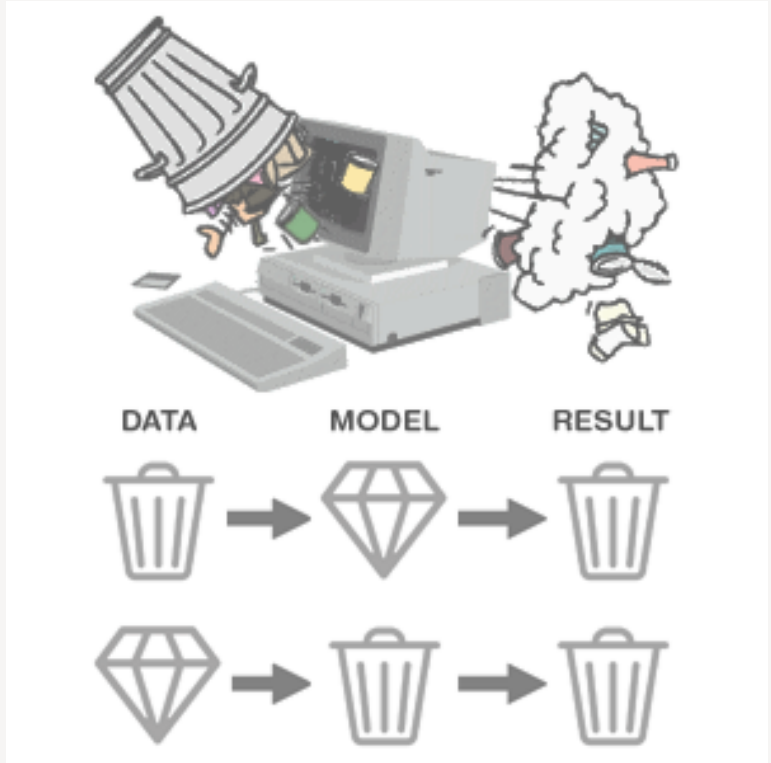
- Key screenlines only & expand in a controlled fashion

Use count / zonal groupings

- Driven by quality of data, prior matrix & network validation

Checks:

- **Remember: Look at the .LPJ & .LPMs!!!**



© https://thedailyomnivore.net/2015/12/02/garbage-in-garbage-out/

Possible:

Questions