## Organic vs. Purchased Data for Travel Time Predictions

- Andres E Chavez
- Andres.Chavez@dot.ca.gov
- Caltrans, District 3

2019 Western States Rural Transportation Technology Implementers Forum

## Glossary

$>$ ATMS - Advanced Traffic Management System.
$>$ BTR - Bluetooth Reader.
$>$ GUI - Graphical User Interface.
$>$ HERE - HERE Technologies. $3^{\text {rd }}$ Party TT Provider.
>IP - Internet Protocol ("IP address").
>LAN - Local Area Network.
$>$ MAC - Media Access Control ("MAC address").
$>$ SLT - South Lake Tahoe.
>TT - Travel Time.
$>$ Waze $-3^{\text {rd }}$ Party TT \& Alerts Provider.
$\rightarrow$ WiFi - Wireless LAN.

## Overview

Organic Data
$>$ Loops - Single vs. double
$>$ Bluetooth
$>$ WiFi
Purchased Data
$>$ Waze
$>$ HERE
Bluetooth/Waze/HERE
>South Lake Tahoe Case Study

## Single Loops

CT Standard Plans 2010 - ES 5B


WINDING DETAIL TYPE A LOOP DETECTOR CONFIGURATION

Cutout on New Pavement


# Single Loop <br> Time Measurement 



$T_{\text {in }}-$ Vehicle's front enters loop.

$T_{\text {out }}$ - Vehicle's rear exits loop.

Therefore $t_{\text {tot }}=t_{\text {out }}-t_{\text {in }}$.

Notice that $\mathrm{t}_{\text {in }}$ and $\mathrm{t}_{\text {out }}$ can be affected by multiple variables.

## Cattrans

# Single Loop 

## Distance Measurement


$L_{L}=$ Length of the loop.

$$
L_{v}=\text { Length of the vehicle. }
$$

$$
D_{\text {tot }}=L_{\text {tot }}=\text { Length of loop }+ \text { vehicle. }
$$

Note that the total distance traveled is NOT just the length of the loop, nor is it just the length of the vehicle. It's a combination of both. In other words, $t_{\text {tot }}$ start counting when the front of the vehicle enters loop, and Stops counting when the front of the vehicle has traveled the length of the loop plus the length of the car.

Cattrave

# Single Loop <br> Speed Calculation 

Speed calculation is simply: $\quad S=\frac{D_{t o t}}{\boldsymbol{t}_{\boldsymbol{t o t}}}$
Unfortunately, there are many factors that introduce variance into $t_{t o t}$
$>$ Vehicle Length.*
$>$ Vehicle Height.*
$>$ Lane Alignment.
> Loop Installation.
$>$ Detector Sensitivity.
$>$ Detector Setting.

Dual Loop calculation eliminates most of these issues.


## Caltrans

## Dual Loops

## Ramp Metering Design Manual



TYPE A MAINLINE LOOP DETECTORS

## Cutout on New Pavement



## Caltrars <br> \section*{Dual Loop <br> <br> Time Measurement} <br> <br> Time Measurement


$T_{T}$ - Vehicle enters Trailing Loop.

Notice that $t_{T}$ and $t_{L}$ can be affected by multiple variables, but they cancel out.

## Dual Loop <br> Speed Calculation

Distance is fixed at $20^{\prime}$. Hence : $\boldsymbol{S}=\frac{\mathbf{2 0} \boldsymbol{f t}}{\boldsymbol{t}_{\boldsymbol{t o t}}}$
Hence these potential issues with Singe Loop Calculations have been eliminated.
$\checkmark$ Vehicle Length.
$\checkmark$ Vehicle Height.
$\checkmark$ Detector Sensitivity.*
$\checkmark$ Sensitivity Setting.*
While these have been mitigated and their impact minimized.
$\checkmark$ Lane Alignment.
$\checkmark$ Loop Installation.*
However, a new variable has been introduced.
> Loop Distance.
But, this can be corrected in the configuration. Hence it is not an issue.

## Loop Pros

## Existing

 Infrastructurecalculations
counters
census

Tried \& True Technology

Cattrans

## Loop Cons

$>$ Inaccurate Calculations $>$ Single Loops
> Defaces New Concrete
$>$ Weakens Concrete
$>$ Requires Lane Closure
> Incorrectly Wired
> Damaged by Contractors
> Uneven Burial Depth
> Potential for Exposed Loops
$>$ Repair Cost
$>$ Not all Vehicles Detected
$>$ Speed at Single Point


Bluetooth/WiFi Eliminates Most of These....

## Bluetooth/WiFi Pros

$\checkmark$ Easy install - Less than 30 minutes per site.
$\checkmark$ Relatively inexpensive - $\$ 2400 /$ site.
$\checkmark$ BTR - $\$ 2200$
$\checkmark$ Antenna - \$160
$\checkmark$ Bracket - $\$ 40$
$\checkmark$ Off pavement - Any cabinet with power suffices.
$\checkmark$ Quick repair - Less than 30 minutes per site.
$\checkmark$ No Lane Closure - Outside Clear Recovery Zone or guarded by rail.
$\checkmark$ Non Intrusive - Does not interfere with Travelers' Phones/Cars.
$\checkmark$ High Deployment - Anyone over 10 years has smart phone.
$\checkmark$ Anonymous - Can encrypt MAC address.
$\checkmark$ Single Detector may detect both directions.

## Bluetooth/WiFi Basics

## Caltrans

1.Detector Captures MAC Addresses (48 unique bits).
2.Forwards to Server.
3.Downstream Detector Captures MAC Address.
4.Forwards to Server.
5.Server Calculates Travel Time.
6.Server Exports Travel Time.


## Iteris Solution

## caltrars

Choice of Vendors. Only Iteris' Velocity had non-cloud option.
> Caltrans owned and operated in-house VM server.
$>$ Readers inside Caltrans metal cabinets.
$>$ Low Bandwidth Requirement.
$>$ Bluetooth or WiFi Detectors.
$>$ Data is Pushed to Server.
$>$ Linux OS.

$5 \pi$

## Iteris Solution

## Caltrans



Western States Rural Transportation Technology Implementers Forum


## Location Selection

$>$ Bidirectional Detection (Mostly).
> < 150', Bidirectional Ok





## Location Selection

## Directional Detection

$\gg 150^{\prime}$.
$>$ Elevation Obstruction.
$>$ Trees.
> Buildings.

## Kingvale Caltrans Yard



## Location Selection

## Caltrans

BTR Location Considerations:
> On/off points
$>$ Frontage Road
$>$ Traffic Lights
> 2.4Ghz Noise
$>$ Power (Wired vs Solar).
$>$ If Solar, Snow Implications.
$>$ Communication
> Low Bandwidth
> 93 Byte Frame/Hit


## BTR <br> Detector Filters

## Caltrans Cuans

F

## Data Filters <br> - $25 \%$ Buffer -45\% Buffer <br> $\square$ Interquartile Range - None

| Speed | Lower <br> Limit | High <br> Limit |
| :---: | :---: | :---: |
| 48 | $\mathrm{n} / \mathrm{a}$ | $\mathrm{n} / \mathrm{a}$ |
| 55 | 36 | 60 |
| 42 | 41.25 | 68.75 |
| 55 | 31.5 | 52.5 |
| 42 | 41.25 | 68.75 |
| 48 | 31.5 | 52.5 |
| 50 | 36 | 60 |
| 55 | 37.5 | 62.5 |
| 52 | 41.25 | 68.75 |
| 48 | 39 | 65 |
| 43 | 36 | 60 |
| 30 | 32.25 | 53.75 |
| 28 | 22.5 | 37.5 |
| 20 | 21 | 35 |

## Percent Filter Example

 Is data point within $25 \%$ of previous sample? i.e.. Less than $125 \%$ or greater than $75 \%$ ? If outside range, it is an outlier.25\% Buffer


| Sample | Speed |
| :---: | :---: |
| 1 | 48 |
| 2 | 55 |
| 3 | 42 |
| 4 | 55 |
| 5 | 42 |
| 6 | 48 |
| 7 | 50 |
| 8 | 55 |
| 9 | 52 |
| 10 | 48 |
| 11 | 43 |
| 12 | 30 |
| 13 | 28 |
| 14 | 20 |

## Interquartile Filter Sample

Is data point more than 1.5 interquartile ranges (IQR) below the first quartile or above the third quartile? If so, it's an outlier.

| Variable | Value |
| :--- | :--- |
| $1^{\text {st }}$ Quartile (Median of lower half of samples) | 42 |
| $3^{\text {rd }}$ Quartile (Median of higher half of samples) | 51.5 |
| IQR (3 ${ }^{\text {rd }}-1^{\text {st) }}$ ) | 9.5 |
| Low Threshold (Q1 $-1.5^{*}$ IQR) | 27.75 |
| High Threshold (Q3 +1.5*IQR) | 65.75 |

Interquartile


## Caltrars

## BTR Detector Penetration (Same Side and Both Sides)



## Caltrans

## BTR Detector Penetration (Opposite Side and Both Sides)

```
Loop Counted Vehicles = 14,229+4,328=18,557
BTR Matches = 1,438 (+ 324)=1,762
Percent Penetration = 7.75% (+ 1.75%) = 9.50%
```

```
Loop Counted Vehicles = 14,229 +969 = 15,198
BTR Matches = 1482 (+ 341) = 1,823
Percent Penetration = 9.75% (+ 2.24%)=12.0%
```




## Bluetooth Issues

## caltrars

## Filtering Algorithm Issues

> Excessive Speeds Reported (Outliers?).


## Caltrans

## Bluetooth Issues

## Speeds graphs were not smooth

$>$ Average affected by single outlier.
$>$ Average changed significantly minute over minute.
$>$ Frontage Road Interference.
US50 Westbound - Silva Valley WB to CMS2
$01 / 03 / 2019$



# Rural vs. Urban - May 14 ${ }^{\text {th }}, 2019$ 



## Caltrans

## WiFi Detour


$>$ Deep Penetration.
$>$ Clients regularly broadcast WiFi Probe Requests.
$>$ Already Associated to an AP?
> Iteris Detects Some Associated Clients.
$>$ Designed for Greater Range than Bluetooth.
> 10's of feet vs 100's of feet.
> Same Frequencies, but Less Channels Than
Bluetooth.
Initially Deployed on I80. However...
caltrans

## WiFi Detour



## WiFi

## BT




## WiFi Issues

## UC Davis Off Ramp



## BTR HW Issues

>HW Failures.
$>$ Motherboard (Close to 20\%/Year)
> Power Supply (Wall Wart)

- WiFi/USB Dong (Consumer Grade)
$>$ No Reset Button
>Serial Interface Discouraged


## Caltrans

## Environmental Issues

## $\rightarrow$ Snow <br> > Knock Downs > 2.4 GHz Noise?

## Bradshaw Rd:

BTR would fail every 2 to 3 weeks. Swapped out all supporting hardware. Root cause was never isolated. At least 8 units bricked.


## SW Issues

$>$ IP Address Maintained in Two Files.
$>$ Lack of Reset Button.
$>$ Cleartext Password.
$>2 x$ Reboots (by Design).
$>$ Can Bypass Login via Links.
$>$ OS Randomly Corrupted.
> Duplicate MAC Addresses on Road.
$>$ GUI Displays Last Captured MAC (Stale Data).

Caltrans

## Bluetooth Cons Summary

$>$ Erratic Travel Times.
$>$ Low Rural Penetration.
$>2.4 \mathrm{GHz}$ Interference.
$\rightarrow$ HW Failure.
$>$ Snow Pack.
> Rural Power.
$>$ Few Cabinets in Rural Areas.
> Duplicate MAC Addresses.

Let's try something to eliminate these cons, enter FREE crowd source data in the form of Waze.

## What is Waze?

$\checkmark$ Crowd Source
$\checkmark$ Alerts
$\checkmark$ Traffic Conditions


## $0-0$



Pros Over Bluetooth
> Rurally Available
> No HW
> No Comm
> No Power
$\rightarrow$ No interference
$>$ Immune to weather
> Non Fixed Endpoints

Cattrans

## Waze Segment Definition

Segment Definition
$>$ Name
$>$ Start Lat
$>$ Start Long
$>$ End Lat
$>$ End Long
$>$ Start Dir
$>$ End Dir

- Start Cross Street
>End Cross Street

Long Turnaround Time.

cultrars ATMS = Velocity(Waze)

| 65 Lincoln to Roseville | Free flow as usual |  |
| :---: | :---: | :---: |
| 65-Ferrari-Ranch-Rd to 65-Blue-Oak-Blvd 5.29 miles | $5 \mathbf{m i n} \mid 57 \mathrm{mph}$ $5 \mathrm{~min} \mid 56 \mathrm{mph}$ |  |
| 70 Olivehurst to 99 <br> Riego <br> 70-CMS503 to <br> 99-Riego-Rd <br> 23.67 miles | Free flow <br> as usual <br> $19 \mathrm{~min} \mid 75 \mathrm{mph}$ <br> $18 \mathrm{~min} \mid 75 \mathrm{mph}$ | $V$ |
| 99 Riego to 99 Elkhorn 99-Riego-Rd to 99-Elkhorn-Blvd 4.45 miles | Free flow as usual <br> $3 \mathrm{~min} \mid 73 \mathrm{mph}$ $3 \mathrm{~min} \mid 73 \mathrm{mph}$ | While 1 <br> > Scrape Data <br> $>$ XML.Convert <br> $>$ Wait (for Velocity) <br> > Append |
| 80 Richards to Dixon 80-Richards-Blvd to 80-Currey-Rd 6.59 miles | Free flow as usual <br> $5 \mathrm{~min} \mid 72 \mathrm{mph}$ <br> $5 \mathrm{~min} \mid 73 \mathrm{mph}$ | > Continue |

- <match_summary_data distance_measurement_unit="Miles"> -<match_summary>
<system_id>Iteris</system_id>
<origin_id>80_Dixon</origin_id>
<dest_id>80_Richards</dest_id>
<origin_roadway>I80</origin_roadway>
<origin_cross_street>Dixon</origin_cross_street>
<origin_direction>Eastbound</origin_direction>
<dest_roadway>I80</dest_roadway>
<dest_cross_street>Richards Blvd</dest_cross_street> <dest_direction>Eastbound</dest_direction> <segment_length_miles>7.6</segment_length_miles> <timestamp>5/17/2019 4:32:33 PM</timestamp> <travel_time> 1025 </travel_time>
<speed_mph std_dev=" 2.87 " $>27<$ speed_mph> <summary_mins>15</summary_mins>
<summary_samples>26</summary_samples>
<map_display>True</map_display>
<substitute_speed>-1</substitute_speed>
</match_summary>


## Waze Hurdles

## RESTRICTED ACCESS



50 CMS10 to Pollock

## Pine:

50-CMS10 to
50-Pollock-Pines
NaN miles

50 JCT 89 Weat to $F$
50-JCT89 to 50 -F-5t
NaN miles

50 Ski Run to Al
Tahoe
faster than usua
| NaN mph
| NaN mph
faster than usual
| NaN mph
| NaN mph
$>$ No JSON or XML Feed
>Python and Selenium Incompatibility
HTTP Scraping
> Unauthorized Users
> 404 Error
$>$ TT not present
$>\mathrm{NaN}$

## Cattane Waze Hurdles

> Turnaround Time
> Segment Inaccuracy (Up to 30\%)
$>$ Rounding Error
> Sub Optimal Route



Bayourd



Slide42

## Waze Cons <br> You Get What You Pay For

$>$ Selenium Incompatibility Issues
$>$ HTTP Scraping
> Unreliable HTTP Feed
$>$ Slow Turnaround Time
$>$ Inaccurate Segments
$>$ ATMS Hack

Let's eliminate these cons. Presenting HERE.

## You are HERE

## What is HERE?

HERE captures location content such as road networks, buildings, parks and traffic patterns. It then sells or licenses that mapping content, along with navigation services and location solutions.

Pros Shared With Waze
> Rurally Available
$>$ No HW
$>$ No Comm
$>$ No Power
$>$ No interference
> Unaffected by weather conditions

Pros Over Waze
> Paid Support - (\$30k/yr. for District 3)
> XML Feed!
> Supported in ATMS 5.3
$>$ Supported in ActiveITS
$>$ Confidence Factor included
> Historical Data available

## HERE TMC's

TMC: Traffic Message Channel.
$\checkmark$ TMC codes are a reference system designed to give a unique alpha-numeric code to road segment for the purposes of assigning traffic information to that segment.
$\checkmark$ Assigned and certified by TISA (Traveler Information Services Association).
$\left.\begin{array}{l|l|l|l|}\hline \text { Country Code } & \text { Table ID } & \text { Direction } & \text { Location } \\ \hline 1 \text { (Numeric or Alpha) } & 05 \text { (Numeric) } & \mathrm{N}(-) \text { or } \mathrm{P}(+) & 012345 \\ \hline \begin{array}{l}\text { Country Code. }\end{array} & \begin{array}{l}\text { Table ID within } \\ \text { The United States uses } \\ \text { Country Code 1. }\end{array} & \begin{array}{l}\text { Direction of travel. }\end{array} & \text { Specific location. } \\ \hline \mathrm{P}(+)=\text { North or East } \\ \mathrm{N}(-)=\text { South or West }\end{array}\right]$

## Cattrans

## TMC Example

105 P 05430


EOberlin Rd

1 = US
$05=$ Northern California
P = Northbound
05430 = Unique identifier within US, CA, NB.

## HERE TMC's Defined

| A | B | C | D | F | G |  |  |
| :---: | :--- | :--- | :--- | :--- | :--- | ---: | ---: |
| 1 | ADMIN1 | ADMIN2 | ADMIN3 | ADMIN4 | ADMIN5 | TMC | TMC_LENGTH |
| 489475 | United States | California | Siskiyou | Uninc Siskiyou County |  | 105P05435 | 2.621803 |
| 489476 | United States California | Siskiyou | Yreka |  | 105P05430 | 1.852229 |  |
| 489477 | United States | California | Siskiyou | Yreka |  | 105P05431 | 0.743655 |


| H | I | J | K | L | M | N |
| :--- | :--- | ---: | ---: | ---: | :--- | :--- |
| LINEAR | PARENT_LIN | TMC_ORDER | ROAD_NAME | ROAD_NUM | ROAD_DIR | POINT_DESC |
| 105P00139 | 105P03009 | 182 |  | I-5 | Northbound | Bailey Hill Rd/Exit 793 |
| 105P00139 | $105 P 03009$ | 177 |  | I-5 | Northbound | Foothill Dr/Exit 775 |
| 105P00139 | 105P03009 | 178 |  | I-5 | Northbound | CA-3/Montague Rd/Exit 776 |


| O | P | Q | R | S | T | U |
| ---: | :---: | :---: | :---: | :---: | :---: | :---: |
| TMC_TYPE | POS_OFF | NEG_OFF | START_LAT | START_LON | END_LAT | END_LON |
| 1 105P05436 | 105P05434 | 41.92239 | -122.57598 | 41.95772 | -122.59409 |  |
| 1 | $105 P 05431$ | $105 P 05429$ | 41.70787 | -122.64236 | 41.73348 | -122.63174 |
| 1 | $105 P 05432$ | $105 P 05430$ | 41.73348 | -122.63174 | 41.74239 | -122.62407 |

TMC Definitions
Released Twice Per
Year

Cattrans

# TMC's Travel Time 

Travel Time File: RealtimeFlowA0105.xml


## RW: Roadway

- LI: Unique String Identifier. Note Embedded +/- Sign.
- DE: Text Description of the Road.
- PBT: Base Timestamp.
- mid: NAVTEQ identifier. DO NOT USE.


## Cattrans

# TMC's Travel Time 

Travel Time File: RealtimeFlowA0105.xml


FIS: List of Flow Items.
> FI: Flow Item.

## $>$ TMC: Traffic Message Center. <br> $>$ CF: Current Flow.

caltans

## TMC's Travel Time

TMC: Traffic Message Channel.
$>$ PC: Point Location Code = TMC ID (stripped).
$>$ DE: Description.
$>$ QD: Queuing direction. (Opposite of traffic flow).
$>$ LE: Length. Units defined above.
catrans

## TMC's Travel Time

## CF: Current Flow

> TY: Always "TR" for normal lanes. (RM, EX, Etc.)
> SP: Capped Average Speed.
$>$ SU: Uncapped Average Speed.
$>$ FF: Free Flow Speed.
$>$ JF: Jam Factor. -1 to 10.
$>$ CN: Confidence Factor. 0.1 to 1.0 (DO NOT IGNORE)
$>$ TS: Travers ability Status. "O"pen or "C"losed.

## WARNING!

## Caltrans <br> Not All TMC's Are Created Equal

NPMRDS INRIX
105+05430 and 105P05430

HERE
105P05430


## HERE in the Office - Actual CMS TT

## Caltrars

TT 50/89 JCT to Meyers
Jan 21st, 2019-12:00 to 23:59


From: Paula Peterson <tahoepaula@
Sent: Monday, January 21, 2019 8:15 PM
To: Nelson, Steve@DOT <steve.nelson@
Subject: Message boards in South Lake Tahoe

Hello...hope you had a nice holiday!
There is something off with the message board times posted in SLT. It took people between 2.5 hours and 3 hours to get from the $Y$ to Meyers for most of the day but the sign said 11 minutes, or sometimes 14 minutes. its great if the signs are accurate so people know...many are turning back tonight as they've been on US50 for hours and not getting far. Of course that is a bigger issue, I'm just curious about the timing.

SouthTahoeNOW.com
Your One Stop for Lake Tahoe News \& Information

## Standstill on South Lake Tahoe area highway and streets; Groups looking into solution

Submitted by paula on Tue, 01/22/2019-8:44pm

## MINUTES TO:



A sign alerting motorists of the drive time was not accurate and well shor of the actual.


## Paula Peterson

SOUTH LAKE TAHOE, Calif. - It's almost a perfect storm for traffic in Lake Tahoe: extra visitors in town for the holiday weekend and epic ski conditions with snow and chain requirements over US50 and Echo Summit.
On Monday, locals and visitors alike were part of that storm, leaving motorists stranded along US 50, Lake Tahoe Blvd., and all surface streets in Meyers that have a link to the highway over Echo Summit.

This isn't a new problem, but one that rears its ugly head on many Sundays and holidays throughout the year. And it's not just a South Lake Tahoe problem but one seen in Truckee and other towns across the west as populations grow.


## Cattrans

## TMC Gotchas

$>$ Non Sequentially Numbered.
$>$ Parallel Paths. $-\rightarrow$
$>$ Hardcoded Endpoints.
> 6 Month TMC Updates.

Create problems for ATMS. ActiveITS is more resilient.

Do NOT ignore Confidence Factor


## Waze $2^{\text {nd }}$ Pass Pros

## Caltrars:



## Waze $2^{\text {nd }}$ Pass Cons

Not Resolved in $2^{\text {nd }}$ Pass
$>$ Lack of Confidence Factor
$>$ ATMS Integration (Velocity Spoofing)
$>$ No Support for ActiveITS


## SLT Revisited

## Coltrans

Waze<br>1 Easy Segment 4.83 Miles


HERE
7 Fixed Segments
4.81

> Bluetooth - Velocity 5 Readers, 4 Segments 5.01 Miles


## Cltans

## SLT Revisited - Free Flow

TT Comparison
Friday April 5th, 2019-8:00 to 16:15


Caltrans

## SLT Revisited - Heavy Flow

TT Comparison
Sunday March 1st, 2019-00:00 to 23:59



## Conclusion \& Next Steps

$>$ BTR's are out.
$>$ Phasing out Loops. (ActiveITS)
$>$ Jury is out on Waze vs. HERE.
$>$ Tach Runs
$>$ Free flow with Traffic.
$>$ Free flow without Traffic.
$>$ Bad Weather.
$>$ Holiday Weekend (July 4 ${ }^{\text {th }}$ ).

$$
Q \& A
$$

