

BLOCKSTATISTICS

Data Analysis de dades generades pels sistemes BlockSat

Víctor Villalobos Coca

Resum—L'aplicació BlockStatistics utilitza les dades generades pels sistemes BlockSat, un sistema dissenyat per SENER INGENIERIA Y SISTEMAS que va integrat a tres locomotores diferents de mercaderies de Ferrocarrils de la Generalitat de Catalunya. BlockSat genera dades contínuament sobre el comportament de les locomotores on està instal·lat. BlockStatistics agafa els logs dels sistemes BlockSat i els emmagatzema en una base de dades MySQL i mostra a l'usuari valors estadístics sobre el log o logs que ha introduït: inici i final dels trajectes, velocitats mitjanes, etc. Utilitzant l'eina Grafana, es manté un servidor on es poden visualitzar gràfica i interactivament les dades introduïdes a la base de dades. Amb aquesta eina, l'usuari pot veure el comportament de les diferents locomotores en un període de temps determinat. Així doncs, BlockStatistics dona forma i valor a les dades generades i serveix com a una base sòlida per una versió futura amb més dades on s'utilitzin algorismes d'aprenentatge computacional per un anàlisi profund de les dades obtingudes.

Paraules clau—BlockSat, Estadística, Representació gràfica, Anàlisi de dades, Ferrocarril, Python.

Abstract—BlockStatistics is an application which uses the data generated by BlockSat Systems, which have been designed by SENER INGENIERIA Y SISTEMAS and have been integrated at three different railway engines of Ferrocarrils de la Generalitat de Catalunya. BlockSat Systems are generating data continually about the behaviour of the unit where it's been installed. BlockStatistics gets the logs generated by BlockSat and stores them in a MySQL database and shows different statistics values obtained by analysing the log or logs that have been introduced, such as: start and end of the trips, average speed, etc. Using Grafana, a server is maintained where the data can be visualized in a graphic and interactive way. Using this tool, the user can see and analyse the behaviour of the different railway engines in a specific period of time. In summary, BlockStatistics gives form and value to the data generated within the railway engines and can be used as a point of start to develop a future version which uses more data and machine learning algorithms to be able to do an even more deep analysis of the gathered data.

Index Terms—BlockSat, Statistics, Graphical Representation, Data Analysis, Railway, Python.



1 INTRODUCCIÓ

La gestió del tràfic ferroviari es realitza mitjançant sistemes de senyalització específics instal·lats a les infraestructures físiques, combinats amb un sistema central que processa la informació recollida de camp, i des del que es telecomanden les instal·lacions. Amb aquests sistemes es regulen els senyals, les agulles, i es donen les limitacions de velocitat en cada tram de via.

BlockSat és un sistema desenvolupat per SENER Ingeniería y Sistemas, que permet el control de tràfic centralitzat i la regulació sense necessitat d'instal·lar un sistema específic a nivell d'infraestructura de via. Es compona d'un equip embarcat i un sistema central. Aquest sistema es troba instal·lat a les locomotores de mercaderies de Ferrocarrils de la Generalitat de Catalunya (FGC) i al seu centre de comandament. BlockSat genera dades en temps real sobre el comportament del tren en el que està instal·lat i proveeix diferents funcions utilitzant aquestes dades, entre elles:

- És capaç d'aturar el tren si va a una velocitat superior a la permesa en el tram en el que es troba.
- Permet la localització exacte del tren des dels Centres de Comandament de la companyia.
- Proveeix a l'empresa d'uns logs on es guarden totes les dades generades pel sistema.

Però tot i disposar de les dades de registre de Blocksat avui en dia, encara no existeix una eina per a l'anàlisi d'aquestes.

BlockStatistics, el projecte desenvolupat que aquí es presenta, obté i processa les dades generades per BlockSat per tal de permetre un anàlisi de les circulacions ferroviàries. L'anàlisi d'aquesta informació permet aportar un valor afegit a la explotació, i detectar comportaments del material, locomotores, trams, formes de conducció, tot relacionant les dades obtingudes.

BlockStatistics és per tant un sistema d'anàlisi de les dades de circulació de les locomotores de mercaderies de FGC.

E-mail de contacte: victor.villalobos@protonmail.com

Menció realitzada: Enginyeria del Software.

Treball tutoritzat per: Yolanda Benítez (Ciències de la Computació) i Héctor García (SENER Ingeniería y Sistemas)

Curs 2018/19

2 ESTAT DE L'ART

L'Internet de les coses (Internet of Things, IoT), fa referència a una xarxa d'objectes de la vida quotidiana connectats entre ells mitjançant Internet. Quan es porta aquest concepte a l'àmbit industrial, es converteix en el conegut com IloT: Industrial Internet of Things. La idea principal del IloT és la mateixa que la del IoT: connectar els diferents dispositius (en aquest cas, els d'una indústria concreta) per a que s'obtinguin dades i estats de cadascun d'ells. La diferència és que, en una xarxa IloT, no hi ha connexió a Internet, sinó que el conjunt d'objectes connectats està únicament constituït pels sistemes industrials que és connectin.

Per aconseguir una xarxa IloT, els dispositius de la majoria d'indústries haurien de canviar-se, ja que la base hardware no suporta aquesta interconnexió entre subsistemes. L'evolució a una xarxa IloT sol ser molt cara, però en la majoria de casos, acaba sortint rentable per a l'empresa, ja que comporta una sèrie de beneficis (manteniment preventiu, localització d'errors, monitorització de la xarxa, etcètera) que a la llarga amortitzen el cost de la inversió.

El sistema BlockSat és una alternativa a substituir els dispositius dels trens i vies per aconseguir uns resultats semblants als que s'aconseguirien en el cas d'utilitzar sistemes interconnectats: mitjançant sensors de diferents tipus, s'obtenen dades diverses de les locomotores on està instal·lat i es poden relacionar entre elles per trobar problemes desconeguts fins ara, així com les seves causes i veure a quins sistemes afecten.

BlockSat divideix de forma lògica la infraestructura de via en unitats virtuals anomenades tram. Cada tram es correspon amb una secció física del trajecte, amb les seves coordenades de geo-posicionament. Per a cadascun d'aquests trams, l'operador ferroviari pot establir les limitacions de velocitat que corresponguin. La informació de tram s'utilitza així mateix per a fer la representació gràfica a nivell de lloc central al centre de comandament.

Amb el sistema embarcat, BlockSat geo-posiciona la ubicació física en temps real, i permet definir accions sobre la locomotora.

En quant a l'anàlisi d'informació i pel que afecta als sistemes ferroviaris, la tecnologia de sistemes consolidada per a la explotació ferroviària, genera cada vegada més quantitat de dades referents a com aquests sistemes desenvolupen les seves funcions: sistemes de ticketing, sistemes d'informació al client, sistemes de control de tràfic ferroviari, sistemes de telecomandament d'instal·lacions, etc.

L'anàlisi de les dades d'explotació i el seu tractament (Big Data) en els àmbits de la operació, el manteniment i la seguretat, és però un aspecte en desenvolupament i que la majoria d'empreses ferroviàries no tenen implementat.

En particular, l'anàlisi del comportament de les circulacions (operació) pot aportar dades significatives per a la optimització de la explotació ferroviària: identificant velocitats reals per trams, possibles ajustos en la planificació del servei o en els torns de personal assignats.

Així mateix, la optimització d'acceleracions i frenades, és un factor fonamental per a l'ajust del consum energètic en els sistemes ferroviaris, i de fet, és una de les principals avantatges quan s'implanten sistemes de conducció automàtica (ATO).

3 OBJECTIUS DEL PROJECTE

3.1 Objectius Principals

Els objectius principals d'aquest projecte es centren en construir el sistema informàtic que permeti la obtenció i anàlisi de les dades reals generades pels sistemes BlockSat de la línia de mercaderies de FGC i la presentació dels resultats en una interfície gràfica d'usuari.

En concret:

- Elaborar una eina o eines que permeti obtenir i analitzar les dades obtingudes pel registre de sistemes BlockSat reals instal·lats a locomotores en servei de Ferrocarrils de la Generalitat de Catalunya.
- Elaborar un portal web on es puguin visualitzar els resultats obtinguts anteriorment, amb interfície tipus quadre de comandament i mitjançant l'ús de diferents taules i gràfiques.

En particular, les dades de trajectes reals disponibles que s'han analitzat són les generades en el ramal de mercaderies entre les estacions de Manresa Alta i Sallent, que és on es troba actiu el sistema BlockSat, i on ara s'estan realitzant les circulacions per fiabilitzar-ne el seu funcionament.

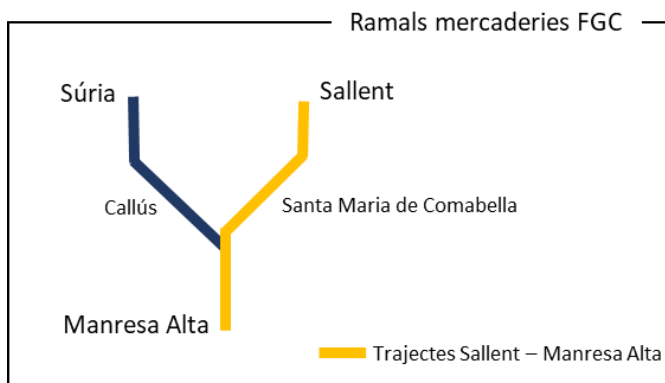


Figura 0. Trajectes monitoritzats per BlockSat

3.1 Objectius derivats

Els objectius derivats contenen aquells aspectes rellevants que es poden inferir com a resultat de l'anàlisi realitzat sobre les dades, i sobre els que el responsable de la operació ferroviària pot incidir i que poden representar una millora de la qualitat del servei i la preservació del medi ambient:

- Trobar trams on la majoria dels trens van a més velocitat de la possible i veure el motiu. (Prioritat alta)

- Trobar patrons de conducció erronis (moltes acceleracions i frenades innecessàries) que es repeteixen a diferents maquinistes i que, d'ésser corregits, podrien suposar un estalvi energètic considerable i, per tant, la reducció d'un percentatge de les emissions de gasos contaminants i ajudant així a preservar el medi ambient. (Prioritat alta)
- Veure si la conducció dels maquinistes varia segons l'horari de servei i com això influeix en la conducció. (Prioritat baixa)

4 METODOLOGIA

4.1 Tecnologies utilitzades al projecte

4.1.1 Eines de desenvolupament utilitzades

Per al desenvolupament del codi del projecte, s'ha utilitzat el llenguatge de programació Python, ja que aquest llenguatge de programació té compatibilitat amb les eines MySQL i Grafana, així com múltiples llibreries útils per a l'anàlisi de les dades.

La base de dades està localitzada al mateix sistema. La base de dades utilitzada és MySQL, administrada mitjançant l'eina PhpMyAdmin. Tant la base de dades com l'eina gràfica (Grafana) es troben en local (127.0.0.1). Els recursos d'Apache i MySQL estan proporcionats per l'eina XAMPP, una eina open-source que habilita l'ús d'aquestes eines en mode local.

4.1.2 Representació de les dades

Per a la realització de les gràfiques de les dades obtingudes, s'ha utilitzat Grafana, una eina per representar les dades de manera gràfica especialitzada en IoT, entre altres. Aquesta eina s'ha fet servir per construir el portal web on es visualitzen les dades obtingudes dels sistemes incorporats als trens. Grafana permet l'ús de fonts de dades per visualitzar aquestes en gràfiques interactives, en l'interval de temps desitjat per l'usuari i mostrant mitjançant colors les diferents fonts de les dades.

Grafana ha permès una representació dinàmica de l'anàlisi dels resultats (actualitzant-se segons la base de dades s'omple amb dades addicionals). Aquest anàlisi s'ha dut a terme mitjançant una taula que relaciona cada trajecte amb diferents valors estadístics que s'han generat a partir de les dades, com ara velocitats màximes, velocitats mitjanes, moments on el tren ha anat a una velocitat superior a la velocitat màxima del tram, etc.

5 PLANIFICACIÓ

Per tal de disposar d'un camí a seguir i obtenir els resultats en el temps desitjat, s'ha realitzat una planificació del projecte on es detallen les diferents tasques, subtasques i fites a assolir.

Aquesta planificació s'ha realitzat utilitzant l'eina Micro-

soft Project, i ha sofert canvis al llarg de l'execució del projecte, ajuntant-se a dates noves que han anat sorgint i reestructurant el camí a seguir.

El projecte BlockStatistics ha tingut una durada aproximada de 4 mesos, on s'han distingit diferents fases del projecte:

1. Fase I: Definició funcional
2. Fase II: Desenvolupament
3. Fase III: Documentació i testeig
4. Fase IV: Tancament TFG

Amb els objectius definits i amb les fases establertes, les tasques principals de la planificació final del projecte han estat les següents:

Tasca	Inici	Final
Fase I: Definició Funcional		
Document definició del projecte	17/09/18	21/09/18
Elicitació de requisits	24/09/18	11/10/18
Anàlisi d'eines software	12/12/18	22/10/18
Instal·lació entorn desenvolupament	23/10/18	29/10/18
Fase II: Desenvolupament		
Importació de dades	30/10/18	15/11/18
Presentació de dades	16/11/18	12/12/18
Simulació i anàlisi de dades	13/12/18	25/01/19
Fase III: Documentació i proves		
Proves generals del sistema	26/01/19	28/01/19
Documentació	28/01/19	31/01/19
Fase IV: Tancament TFG		
Informe Final TFG	23/01/19	27/01/19
Elaboració de la presentació del projecte	04/02/19	10/02/19
Elaboració del dossier	07/02/19	11/02/19

Taula 1. Planificació del projecte

A l'annex A1 es recull la planificació complerta feta amb Microsoft Project.

6 DESENVOLUPAMENT DEL PROJECTE

6.1 Requisits del projecte

L'elicitació de requisits s'ha realitzat consultant els diferents actors del projecte, per definir els requisits generals. Així mateix, s'han identificat quines són les dades necessàries a tractar a partir de la informació disponible i quines són les dades que calia generar. Finalment, es van identificar les tecnologies a utilitzar.

6.1.1 Requisits Funcionals

REQ0010	El sistema haurà d'importar la informació dels logs de registre del sistema BlockSat per tal de poder analitzar i extreure les dades necessàries.
---------	---

REQ0020	El sistema haurà d'identificar els trajectes realitzats per una locomotora
---------	--

REQ0030	El sistema haurà de generar la informació de velocitat màxima per a cada trajecte.
REQ0040	El sistema haurà de detectar l'ultrapassament del límit de velocitat definit per cada tram de Blocksat.
REQ0050	El sistema haurà de generar la informació de moment d'inici i fi per a cada trajecte.
REQ0060	El sistema haurà de generar la informació del nombre d'acceleracions per a cada trajecte.
REQ0070	El sistema haurà de generar la informació del nombre de frenades per a cada trajecte.
REQ0080	El sistema haurà de generar la informació del temps amb velocitat sostinguda.
REQ0090	El sistema haurà de presentar la informació en forma gràfica.

6.1.2 Requisits tècnics i restriccions

REQ0110	El sistema BlockStatistics ha d'ésser executable en un sistema Ubuntu.
REQ0120	Les dades de FGC s'hauran d'extreure directament des dels logs, tractant el format amb el que es guarden.
REQ0130	L'aplicació de visualització de les dades haurà d'ésser configurable en un servidor.
REQ0140	L'aplicació de visualització de les dades haurà de disposar de diferents perfils d'usuari.
REQ0150	Les dades que es visualitzin a l'aplicació de visualització s'hauran d'actualitzar automàticament conforme es vagin afegint nous logs a la base de dades. (Actualització dinàmica).
REQ0160	La base de dades a utilitzar haurà de ser del tipus relacional.

6.2 El sistema BlockStatistics

BlockStatistics està compost per tres mòduls que treballen conjuntament gràcies a la base de dades MySQL. Aquests mòduls són:

1. **Parse&InsertBD:** És el procés encarregat de llegir els logs generats per BlockSat i extreure, utilitzant expressions regulars (Regex), les dades sobre el trajecte del tren que estan presents al

log. Un cop el procés ha filtrat i extret les dades, aquestes són inserides a la base de dades MySQL.

2. **Simulate&InsertBD:** És el procés encarregat de la generació de trajectes on es simula el comportament d'un tren. Aquestes dades tenen el mateix format que les originals de BlockSat. Un cop s'han generat les dades, s'introdueixen a la base de dades MySQL.
3. **GetStatistics:** És el procés que s'encarrega de l'anàlisi de les dades introduïdes a la Base de Dades MySQL i les interpreta per extreure, per tren i trajecte, informació sobre el nombre d'acceleracions, frenades, etc.

6.2.1 Dades simulades i dades de FGC

En el moment en el que el sistema BlockSat embarcat s'activa, aquest va registrant les dades en fitxers logs cada dos segons i aquests s'envien al CTC (Control de Tràfic Centralitzat) de FGC. Analitzant les dades obtingudes dels logs, s'ha detectat que, en determinats moments, el sistema BlockSat perd puntualment el senyal i es deixen de rebre dades al CTC, amb lo qual no hi ha informació contínua de tots els trams. Donat que el sistema no enregistra la data d'inici i final d'un trajecte de manera específica, la traçabilitat dels trajectes es feia molt feixuga i imprecisa a nivell lògic.

Per superar aquest fet, s'ha creat addicionalment un procés de simulació de comportament ferroviari per a poder completar la realització de l'anàlisi de dades. Aquest procés, **Simulate&InsertBD**, genera dades de trajectes dels trams on estan situades les locomotores BlockSat i els insereix a la base de dades MySQL. Aquestes dades són les que s'utilitzen a el procés GetStatistics, així que els resultats complets que s'han obtingut per trajecte i tren corresponen als trajectes simulats mitjançant Simulate&InsertBD. D'aquesta forma el projecte ha pogut presentar les dades i resultats finals.

Tot i aquesta dificultat a l'hora d'analitzar les dades reals obtingudes dels logs de FGC, aquestes han seguit essent essencials i de valor molt important per el projecte, i han estat introduïdes constantment a la base de dades MySQL i, en el moment en que es generin logs amb amb dues variables que indiquin l'inici i el final d'un trajecte, es podrà utilitzar el procés GetStatistics per obtenir estadístiques sobre els trajectes reals, donat que el tractament dels logs està complet.

A Grafana, aquesta situació ha quedat plasmada en la identificació de dos *Dashboards* diferents: "Dades Reals FGC", on es veuen gràficament els valors dels logs dels trajectes de FGC i el dashboard "Dades simulades", on es mostren gràficament els trajectes simulats i una taula on s'observen les diferents qualitats de cada trajecte.



Figura 1. Dashboard Dades Reals FGC



Figura 2. Dashboard Simulades

ID	Tipus	Tipus Trajecte	Data Trajecte	Temps Inicial	Temps Final	Velocitat Màxima	Velocitat Mitjana	Velocitat Mínima	Velocitat Estànd. Desv.	Velocitat Estànd. Desv. Rel.	Velocitat Estànd. Desv. Abs.
1001	1	1	2019-10-25 10:00:00	2019-10-25 10:00:00	2019-10-25 10:00:00	100	50	0	10	20%	20
1002	1	1	2019-10-25 10:00:00	2019-10-25 10:00:00	2019-10-25 10:00:00	100	50	0	10	20%	20
1003	1	1	2019-10-25 10:00:00	2019-10-25 10:00:00	2019-10-25 10:00:00	100	50	0	10	20%	20
1004	1	1	2019-10-25 10:00:00	2019-10-25 10:00:00	2019-10-25 10:00:00	100	50	0	10	20%	20
1005	1	1	2019-10-25 10:00:00	2019-10-25 10:00:00	2019-10-25 10:00:00	100	50	0	10	20%	20
1006	1	1	2019-10-25 10:00:00	2019-10-25 10:00:00	2019-10-25 10:00:00	100	50	0	10	20%	20
1007	1	1	2019-10-25 10:00:00	2019-10-25 10:00:00	2019-10-25 10:00:00	100	50	0	10	20%	20
1008	1	1	2019-10-25 10:00:00	2019-10-25 10:00:00	2019-10-25 10:00:00	100	50	0	10	20%	20
1009	1	1	2019-10-25 10:00:00	2019-10-25 10:00:00	2019-10-25 10:00:00	100	50	0	10	20%	20
1010	1	1	2019-10-25 10:00:00	2019-10-25 10:00:00	2019-10-25 10:00:00	100	50	0	10	20%	20

Figura 3. Taula d'estadístiques per Trajecte

6.3 Arquitectura lògica

El sistema BlockStatistics està separat en dos blocs de programari diferent, enllaçats per la base de dades MySQL. Per una banda existeix el bloc d'inscripció a la base de dades, compost pels algorismes Parse&InsertBD i Simulate&InsertBD. En concret, Parse&InsertBD rep les dades des dels logs generats a les locomotores de FGC que contenen el sistema BlockSat, convertint-se així en la peça que uneix el projecte amb les locomotores.

Per altra banda, existeix el bloc d'anàlisi de dades, "Data Analysis", compost per la representació gràfica de les dades reals i simulades (Data Graph) i les estadístiques extretes de les dades simulades mitjançant el procés GetStatistics (Data Statistics). Aquest bloc es representa gràficament en format de gràfic i taula a Grafana.

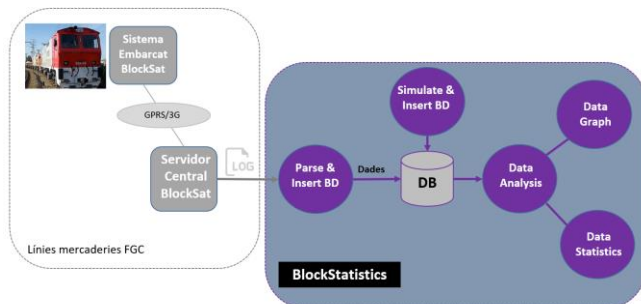


Figura 4. Arquitectura lògica de BlockStatistics

```

2018/10/25-23:43:30.396 [D] (0): Received HMI data 'updates'...
2018/10/25-23:43:30.397 [D] (0): Received 'train update'
2018/10/25-23:43:30.526 [D] (0): Processing message: (msg_type == HMI_DATA)
2018/10/25-23:43:31.071 [D] (0): CtoReceiver.readPendingDatagrams() for CTC# 1 runs in thread: 0x7fed9dc44700
2018/10/25-23:43:31.072 [D] (0): CtoReceiver.readPendingDatagrams() packet received from CTC# 1 : {"message_type":...}
2018/10/25-23:43:31.072 [D] (0): Received HMI data 'updates'...
2018/10/25-23:43:31.072 [D] (0): Received 'ctc_status update'
2018/10/25-23:43:31.073 [D] (0): Processing message: (msg_type == HMI_DATA)
2018/10/25-23:43:33.771 [D] (0): CtoReceiver.readPendingDatagrams() for CTC# 1 runs in thread: 0x7fed9dc44700
2018/10/25-23:43:33.772 [D] (0): CtoReceiver.readPendingDatagrams() packet received from CTC# 1 : {"message_type":...}
2018/10/25-23:43:33.772 [D] (0): Received HMI data 'updates'...
2018/10/25-23:43:33.772 [D] (0): Received 'train update'
2018/10/25-23:43:33.910 [D] (0): Processing message: (msg_type == HMI_DATA)
2018/10/25-23:43:33.996 [D] (0): CtoReceiver.readPendingDatagrams() for CTC# 3 runs in thread: 0x7fed9cc42700
2018/10/25-23:43:33.996 [D] (0): CtoReceiver.readPendingDatagrams() packet received from CTC# 3 : {"message_type":...}
2018/10/25-23:43:33.996 [D] (0): CtoReceiver.readPendingDatagrams() packet received from CTC# 3 : {"message_type":...}
2018/10/25-23:43:33.996 [D] (0): Received HMI data 'updates'...
    
```

Figura 5. Exemple de Log de BlockSat

La base de dades disposa de diferents taules on es guarden els valors de BlockStatistics. Aquestes taules són:

1. **Logbon2:** Aquesta taula conté les dades extretes dels logs de les locomotores de FGC. És utilitzada per la representació gràfica dels trajectes realitzats per les locomotores on estan integrats els sistemes BlockSat.
2. **Simdata:** Aquesta taula conté les dades simulades dels trajectes fets, amb un format equivalent al de les dades reals. Aquesta taula és utilitzada per a la representació gràfica d'aquestes dades així com per a l'obtenció de les dades per l'anàlisi estadístic.
3. **Valoresest:** Aquesta taula conté els valors obtinguts a partir de l'anàlisi estadístic de la taula Simdata. Conté les dades que es mostraran com a valors estadístics dels trajectes simulats.
4. **Velswmax:** Aquesta taula conté la velocitat màxima per cada tram de via. S'utilitza com a complement per a obtenir resultats estadístics i per a la generació de trajectes simulats.

El procés **Simulate&InsertBD** genera, de manera semi-aleatòria, dades d'un trajecte entre una data d'inici i una data final. Depenent de la situació on es trobi, el tren accelerarà, frenarà o mantindrà la velocitat. Un cop s'han generat totes les dades del trajecte, aquestes són pujades a la taula Simdata.

Aquestes dades s'han generat seguint un mètode de generació aleatori ponderat, donant preferència a les acceleracions abans que a les frenades i el mantenir la velocitat, per tal de simular un comportament de portar el tren a prop de la velocitat màxima permesa a cada tram.

En el cas de que el tren estigués anant a més velocitat de la permesa al tram, el procés dona preferència absoluta a la realització d'una frenada, i es manté aquest estat d'emergència fins que la velocitat del tren és menor a la velocitat màxima permesa al tram.

El procés rep com a paràmetres un any, un mes, i un interval de dies, i generarà un trajecte a la mateixa hora per cada dia que estigui dins l'interval, per exemple, si el procés rep per paràmetres: 2019, 1, 10 i 30, el procés generaria 20 trajectes, un per cada dia des del dia 10/01/2019 fins el dia 30/01/2019.

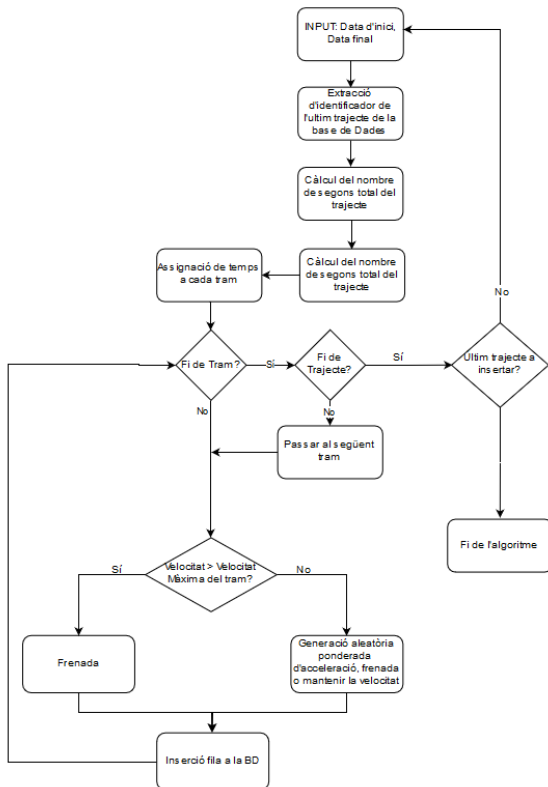


Figura 6. Lògica de l'algorisme Simulate&InsertBD

El procés **Parse&InsertBD** extreu, mitjançant l'ús d'expressions regulars, les dades a partir d'un conjunt de logs reals generats pels sistemes BlockSat embarcats. Un cop troba dades en un log, les va introduint a la base de dades a la taula Logbon2.

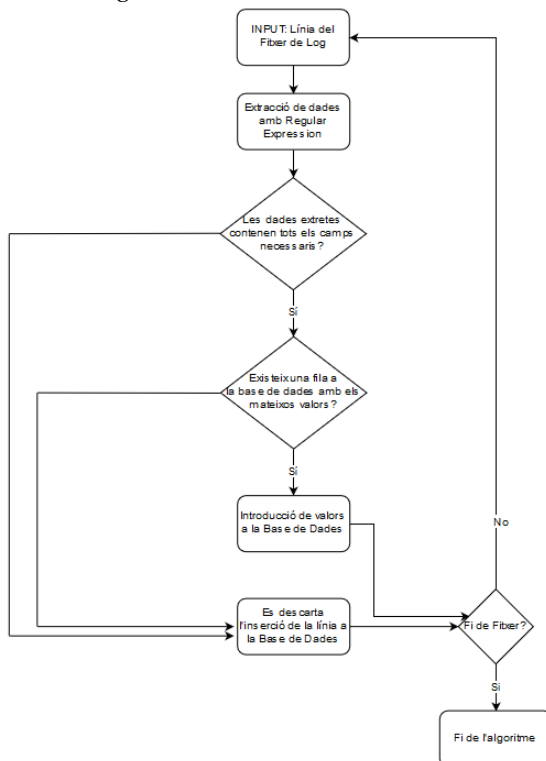


Figura 7. Lògica de l'algorisme Parse&InsertBD

7 RESULTATS

Un cop inserits els logs de FGC i els simulats a la base de dades, s'ha procedit a l'elaboració dels dos dashboards que es presenten al usuari. El primer dashboard, anomenat "Dades reals FGC", conté totes les gràfiques necessàries per visualitzar les dades generades en real a les locomotores.

Les gràfiques sempre utilitzen el mateix codi de colors:

1. La línia taronja representa la velocitat màxima del tram en el que es troba el tren.
2. La línia verda representa la velocitat a la que es desplaça el tren en un determinat moment
3. La línia blava representa el tram on es troba el tren.

Observant els resultats obtinguts, s'observen pèrdues de dades, així com dades generades per error. A la figura 8 podem observar com existeix una pèrdua de dades a partir de les 9:55, i la següent dada de la que es disposa és de dos dies més tard, un altre cop al tram 0. Tot i que la línia es continua, entre la figura 8 i la figura 9 no hi ha dades registrades (a la figura 8 es pot observar la posició del ratolí [la línia vermella] i l'últim punt on es tenen dades, marcat amb un cercle sobre cada línia de resultats).

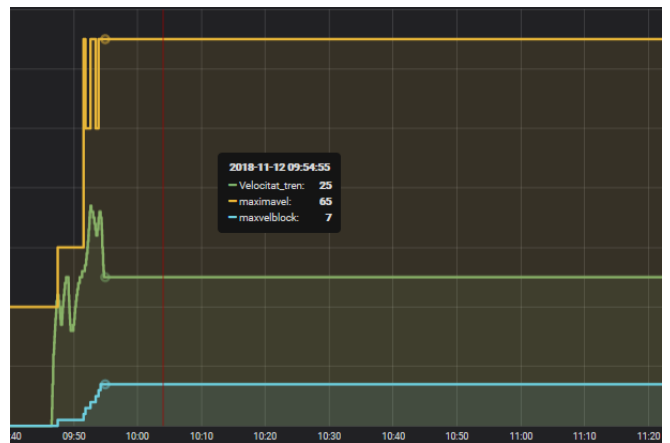


Figura 8: Inici de trajecte incomplet

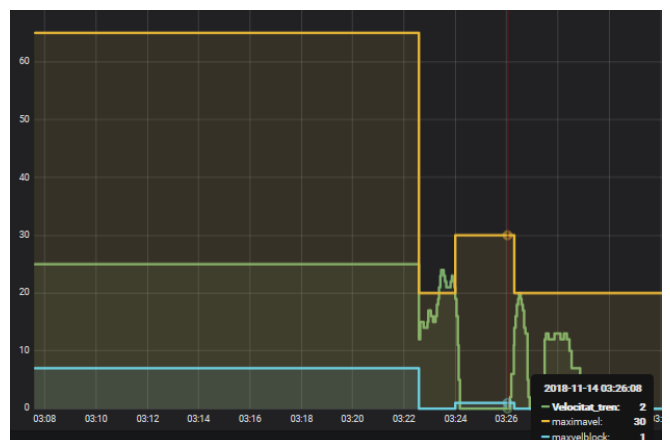


Figura 9: Continuació de les dades de la figura 8

Aquest fet, que es repeteix en més d'una ocasió, és degut

probablement a errors de comunicació entre el CTC (Centre de Tràfic Centralitzat de FGC, a on s'envien els logs generats per BlockSat) i els sistemes embarcats.

Aquesta situació és la que va originar la idea de la simulació de les dades per a poder completar l'anàlisi de dades de manera que, un cop solucionat aquest error, es pugui realitzar immediatament el mateix anàlisi per a les dades reals de FGC.

Un cop introduïdes les dades simulades, seguint el procediment descrit a l'apartat anterior, s'ha procedit a la visualització de les dades simulades juntament amb les estadístiques.

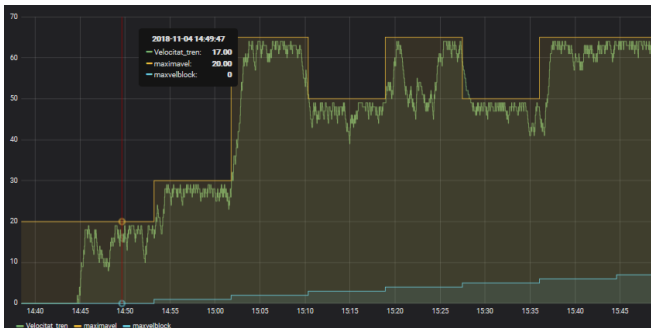


Figura 10: Anada d'un trajecte simulat

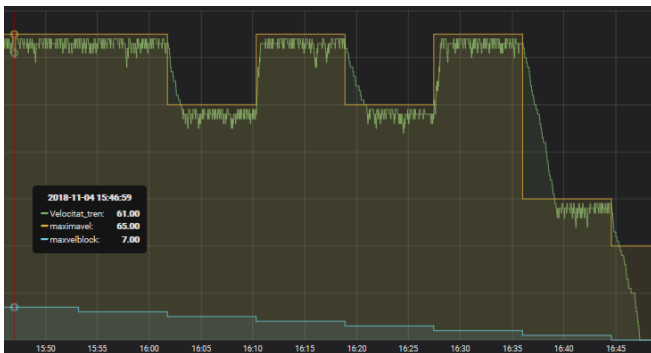


Figura 11: Tornada d'un trajecte simulat

ID_TRAJECTE	ID_TREN	Inici_Trajete	Final_Trajete	Temps_Accelerant	Temps_Frenant	Temps_Manant_Velocitat	Velocitat_Mitja	Temps_Per_Sob
2001	1	2018-11-01 15:44:39	2018-11-01 17:33:07	305	610	1869	46	544
2002	1	2018-11-02 15:44:39	2018-11-02 17:33:07	307	615	1830	46	504
2003	1	2018-11-03 15:44:39	2018-11-03 17:33:07	305	610	1846	46	452
2004	1	2018-11-04 15:44:39	2018-11-04 17:33:07	307	615	1834	46	460

Figura 12: Estadístiques del trajecte del 4/11/18 (marcat)

A la figura 12, podem observar una part de la taula que conté les dades estadístiques extretes a partir de les dades simulades. Pel trajecte mostrat a les figures 10 i 11, podem observar que, en un trajecte de 2h el tren ha:

- Estat accelerant durant 307 segons
- Estat frenant durant 615 segons
- Mantingut la velocitat durant 1834 segons
- Estat per sobre de la velocitat màxima permesa durant 460 segons
- Tingut una velocitat mitja de 46 km/h.

Observant els resultats obtinguts, podem observar com el conductor del tren no accelera en cap moment per superar

la velocitat màxima del tram en el que es troba, però per contra, la frenada es produeix tard i per tant ha estat 460 segons per sobre de la velocitat màxima permesa: El conductor no frena abans d'arribar al canvi de tram, fet que provoca que, en el cas de que la velocitat del tram al que entra és inferior a la velocitat màxima, hagi de frenar urgentment el tren per tal de mantenir-se a una velocitat permesa.

Aquest comportament, tot i ésser simulat, és un exemple d'un possible patró de comportament que es podria trobar un cop es puguin analitzar correctament les dades obtingudes per BlockSat. En concret la optimització de les acceleracions i frenades és un fet bàsic per a la optimització del consum energètic.

A l'annex A3 s'inclouen captures de pantalla d'ambdós Dashboards, el de les dades de FGC i el de les simulades.

8 EVOLUCIÓ FUTURA

BlockStatistics és una base sòlida per a una futura aplicació de reconeixement intel·ligent de patrons de conducció.

Com s'ha indicat, en el moment que s'enregistren en el sistema embarcat les marques d'inici/fi per acotar el trajecte, el sistema BlockStatistics està preparat per mostrar els resultats estadístics d'aquests trajectes reals.

Utilitzant algorismes de *Machine learning*, uns logs més precisos i que indiquin l'inici i final d'un trajecte, juntament amb un professional d'operació de FGC que entreni el procés per a que aquest identifiqui què significa cada part d'un log, es pot desenvolupar una aplicació que, en temps real, mostri per pantalla trajectes, estadístiques dels mateixos i que avisi als operaris en cas de que trobi patrons de conducció erronis.

Tot i que aquesta versió de BlockStatistics estigui executada en un entorn local, l'aplicació està preparada per ésser utilitzada a nivell més ampli, allotjant la base de dades i Grafana en un servidor. S'han definit dos tipus d'usuari a Grafana per tal de facilitar l'ús empresarial de l'eina: administrador i usuari. L'administrador podrà editar panells, afegir-ne de nous o crear nous dashboard, mentre que l'usuari els visualitzarà.

9 CONCLUSIONS

Aquest projecte ha suposat un aprenentatge profund i intens per a mi. Des del primer moment, m'he hagut d'adaptar a diferents situacions que fins el moment no se m'havien plantejat mai: treball amb logs industrials, entendre com organitzar la informació, i, entre altres, adaptar-me a treballar en entorns de treball poc ortodoxos com ara Grafana.

És cert que el fet de que BlockSat sigui un sistema innovador i que estigui en fase de fiabilització i ajustos ha suposat que en els logs reals analitzats, en alguns casos, mancava informació d'alguns trams, però tot i així han proporcionat una informació molt valuosa que es podrà aprofitar en un futur per explorar altres funcionalitats de l'aplicació.

Per superar la dificultat trobada en la definició de trajectes reals, he hagut de generar un mòdul addicional, no previst inicialment, per a la simulació de dades que m'ha permès concloure el projecte i que, tot i deixar de tenir utilitat en el moment en que els logs de BlockSat continguin dades que delimitin els trajectes de les locomotores i continguin informació continuada de tots els trajectes, m'ha servit, a nivell tècnic, per a entendre mecanismes alternatius de generació de comportaments amb variables que canvien en funció del temps. A nivell personal, aquesta dificultat m'ha fet avaluar la situació i trobar la millor sortida possible, que en aquest cas ha estat fer un mòdul d'estadística fàcilment modificable per mostrar les estadístiques sobre les dades reals.

L'eina desenvolupada és fàcilment adaptable a diferents entorns, doncs essent Grafana una eina dinàmica per a la visualització de dades, es podran visualitzar totes aquelles estadístiques generades en un futur de manera fàcil i senzilla, així com fer un anàlisi a llarg termini de tots els trajectes realitzats durant el temps que s'utilitzi l'eina.

10 AGRAÏMENTS

Voldria agrair a la meva tutora de la UAB, Yolanda Benítez, pels consells i la tutorització que m'ha prestat al llarg del desenvolupament del projecte.

Vull agrair també al meu tutor de SENER, l'Héctor Garcia, per la oportunitat de treballar en aquest projecte així com l'ajuda i els coneixements que m'ha aportat durant totes les etapes del mateix.

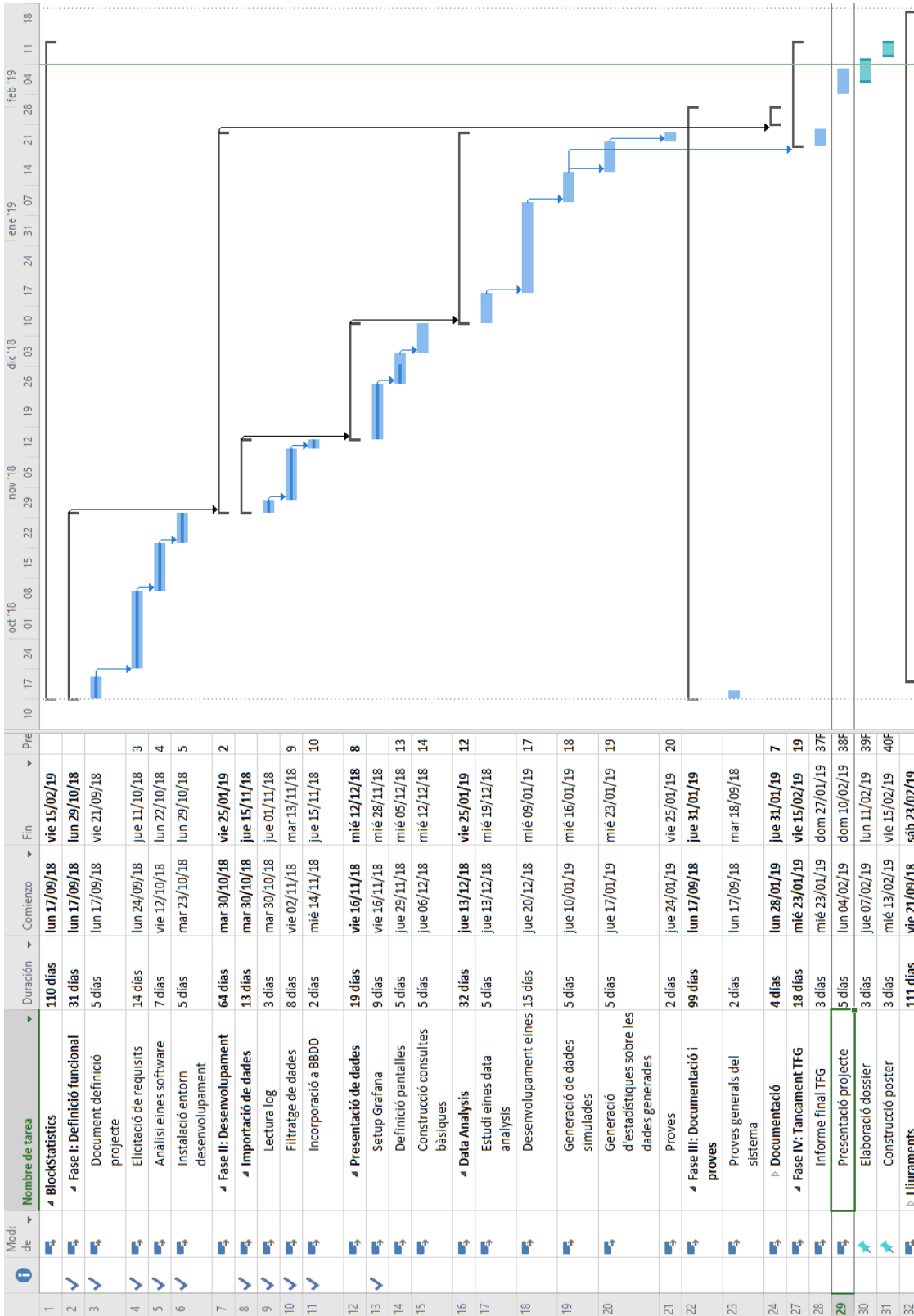
11 BIBLIOGRAFIA

- [1] Big Data junction. Converting raw data into meaningful insights remains to be a challenge for railway industry. [last visited 12/2018] <http://www.railnews.in/big-data-junction-converting-raw-data-into-meaningful-insights-remains-to-be-a-challenge-for-railway-industry/>
- [2] Big Data in Railways. European Union Agency for Railways. Antonio D'Agostino. [last visited 12/2018] https://www.era.europa.eu/sites/default/files/activities/docs/cor_big_data_en.pdf
- [3] BlockSAT - Manual de operación. SENER. D0210047-223IT-MN-2301
- [4] BlockSAT - Requisitos d'usuari i de sistema. SENER. D0210047-223CO-RQ-2202
- [5] Creació de dashboards a Grafana - Albert Coronado [last visited 10/12/2018], <https://www.albertcoronado.com/>
- [6] Digital Ocean, How to configure Grafana, [last visited 27/10/2018], <https://www.digitalocean.com/community/tutorials/how-to-install-and-configure-grafana-to-plot-beautiful-graphs-from-zabbix-on-centos-7>
- [7] Grafana, Features of Grafana, [last visited 01/10/18], <https://grafana.com/>

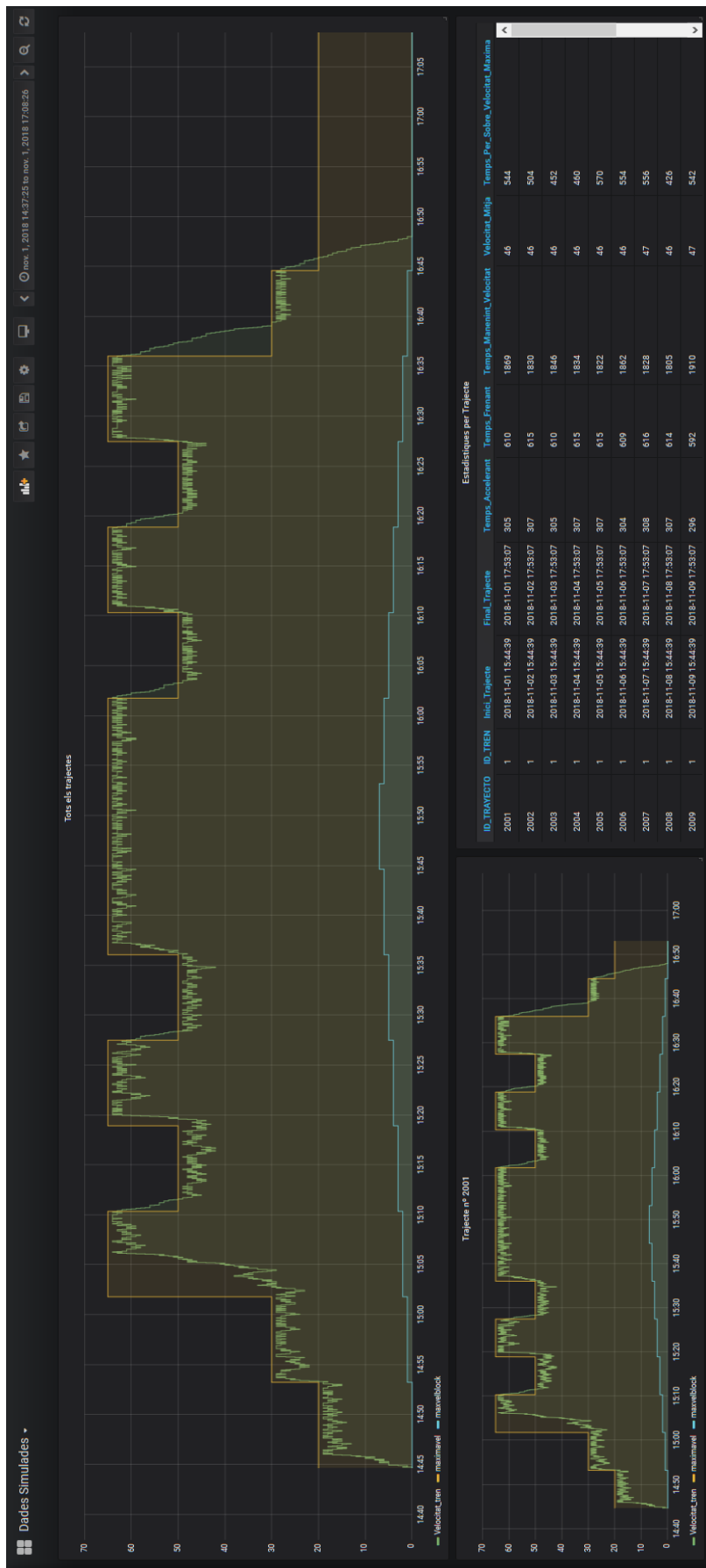
- [8] Grafana, Getting Started with Grafana [last visited 25/10/2018], http://docs.grafana.org/guides/getting_started/
- [9] NodeRed, [last visited 01/10/18], <https://nodered.org/>
- [10] NodeRed, NodeRed go-IoT features [last visited 1/10/18], <https://nodered.org/users/go-iot/>
- [11] Plugins per Grafana, [last visited 02/12/2018] <https://grafana.com/plugins>
- [12] Python, Developer's guide, [last visited 15/10/18] <https://devguide.python.org/>
- [13] Python MySQLConnector [last visited 20/10/2018], <https://dev.mysql.com/doc/connector-python/en/>
- [14] Python MySQL, [last visited 20/10/2018], https://www.w3schools.com/python/python_mysql_getstarted.asp
- [15] Using big data for railway operations and maintenance. Lulea University of Technology. Diego Galar, Uday Kumar, Ramin Karim. [last visited 12/2018] <https://www.globalrailwayreview.com/article/34319/big-data-railway-operations-maintenance/>

ANNEX

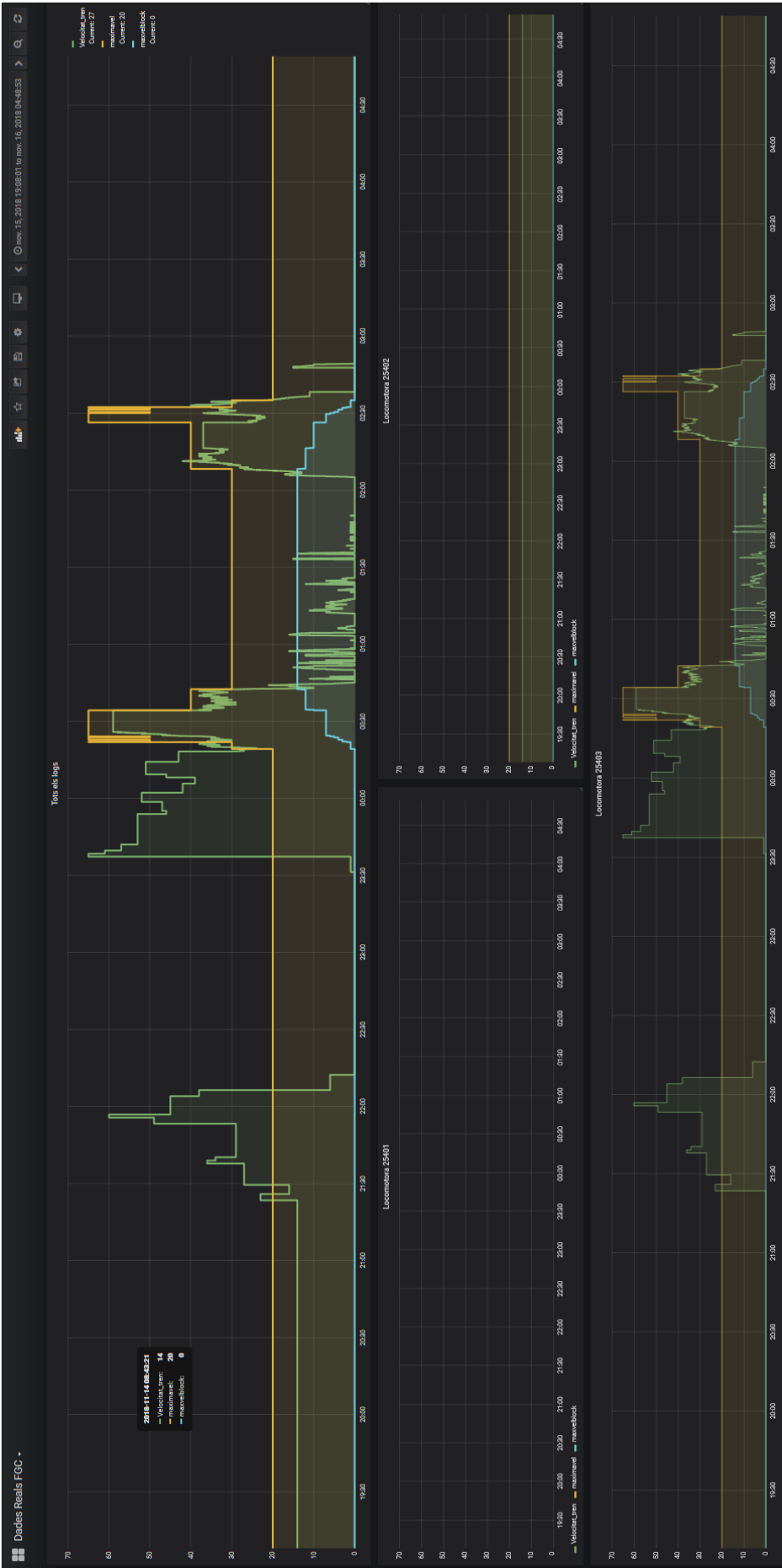
A1. PLANIFICACIÓ COMPLERTA DEL PROJECTE



A3. CAPTURES DE PANTALLA DE GRAFANA



Dashboard – Dades Simulades.



DashBoard- Dades Reals FGC.