



Máster Universitario en  
Sistemas Ferroviarios

# **Desarrollo de una herramienta de apoyo al diseño y optimización de conducción automática de trenes en la que se integren modelos de machine learning**

TRABAJO FIN DE MÁSTER

Máster Universitario en Sistemas Ferroviarios  
ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA ICAI

CURSO: 2º MII + MSF

Autor: Sofía Rodríguez Pita

Director/es:

Asunción Paloma Cucala García

Antonio Fernández Cardador



**TÍTULO:** Desarrollo de una herramienta de apoyo al diseño y optimización de conducción automática de trenes en la que se integren modelos de machine learning

**AUTOR:** Sofía Rodríguez Pita

Firma:

Handwritten signature of Sofía Rodríguez Pita in cursive script.

**DIRECTOR:** Asunción Paloma Cucala García

Firma:

Handwritten signature of Asunción Paloma Cucala García, featuring stylized initials 'APCG'.

**CODIRECTOR:** Antonio Fernández Cardador

Firma:

Handwritten signature of Antonio Fernández Cardador, featuring stylized initials 'AFC'.

## FICHA TÉCNICA

Apellidos y nombre del autor: Rodríguez Pita, Sofía

Apellidos y nombre del director/es del Trabajo Fin de Máster: Cucala García, Asunción Paloma y Fernández Cardador, Antonio

Programa cursado y curso académico: 2º MII + MSF

Título del trabajo: Desarrollo de una herramienta de apoyo al diseño y optimización de conducción automática de trenes en la que se integren modelos de machine learning

Resumen de las principales ideas desarrolladas (unas 20 líneas)

En este trabajo se desarrolla una herramienta de apoyo al diseño de las marchas que ejecutan los sistemas de conducción automática de trenes ATO. En ella se integran modelos de machine learning, que se utilizan para hacer predicciones. Para la optimización de las marchas, se predice el tiempo de recorrido y el consumo esperado que tienen entre dos estaciones con parada. Es posible realizar la predicción y la optimización consecutivas o por separado. En ambos casos se pueden obtener gráficas o tablas que indiquen el conjunto de soluciones óptimas en consumo y tiempo de recorrido.

Junto a este modelo se integra otro, que modeliza el comportamiento de un equipo ATO. Este modelo se puede utilizar en el simulador de marcha del tren integrado en la herramienta, para evaluar las marchas obtenidas con el primer modelo, o para replicar el comportamiento del ATO en una interestación. Es posible obtener gráficas del recorrido del tren, como la de espacio-velocidad, o espacio-salida del ATO; o tablas con los resultados de la simulación.

Se ha tratado de desarrollar una herramienta adaptable a cualquier situación. Para ello, se pueden cambiar los ficheros utilizados, permitiendo probar otros modelos de machine learning, o simular otras líneas. También se pueden cambiar las características del tren utilizado en la simulación, para acomodar la herramienta, dentro de sus posibilidades, al caso real estudiado.

Gracias a la utilización de modelos de machine learning, las predicciones son más rápidas que las realizadas por modelos y simuladores convencionales. Es posible identificar las marchas más eficientes para someterlas a un análisis posterior. En el modelado del equipo de ATO, al tratarse de una red neuronal ajustada a partir de registros entrada-salida, la solución funciona de forma independiente a los distintos fabricantes, permitiendo replicar su comportamiento en cualquier situación.

Desarrollo de una herramienta de apoyo al diseño y optimización de conducción automática de trenes en la que se integren modelos de machine learning

TRABAJO FIN DE MÁSTER

---

# Índice

|     |  |    |
|-----|--|----|
| 1.  | Introducción .....   | 1  |
| 1.1 | Estado del arte .....  | 2  |
| 2.  | Objetivos y planificación .....                                | 5  |
| 2.1 | Objetivos .....  | 5  |
| 2.2 | Planificación de tareas.....                                   | 5  |
| 3.  | Modelos de machine learning integrados .....                   | 6  |
| 3.1 | Modelo para la predicción de tiempo y consumo.....             | 6  |
| 3.2 | Modelos para la predicción de la salida de ATO .....           | 7  |
| 4.  | Requisitos .....   | 11 |
| 5.  | Plataforma utilizada y diseño .....                            | 14 |
| 5.1 | Apartado de análisis .....                                     | 15 |
| 5.2 | Apartado de predicción de tiempo y consumo .....               | 16 |
| 5.3 | Apartado de predicción de salidas del ATO.....                 | 18 |
| 5.4 | Apartado para el guardado y la carga de archivos.....          | 23 |
| 5.5 | Explorador de archivos.....                                    | 25 |
| 6.  | Validación de los requisitos y pruebas de funcionamiento ..... | 27 |
| 7.  | Caso de ejemplo .....  | 35 |
| 8.  | Conclusiones y aportaciones .....                              | 40 |
| 9.  | Bibliografía .....   | 42 |

# Índice de figuras

|  |    |
|--|----|
| Figura 1: Crecimiento esperado de kilómetros con GoA4 en el mundo [8].....                             | 2  |
| Figura 2: Comparación del comportamiento de la red neuronal y el RandomForest.....                     | 7  |
| Figura 3: Predicción de ATO usando una red de tracción y otra de freno .....                           | 8  |
| Figura 4: Predicción de ATO utilizando una red auxiliar de curva de freno .....                        | 9  |
| Figura 5: Página principal de la interfaz .....  | 14 |
| Figura 6: Pestaña de salidas del ATO .....   | 15 |
| Figura 7: Pestaña de 'Análisis'.....   | 16 |
| Figura 8: Gráfico generado desde la pestaña de 'Análisis' .....  | 16 |
| Figura 9: Apartado de predicción de tiempos y consumos .....   | 17 |
| Figura 10: Gráfico obtenido de puntos que conforman el Pareto front entre los predichos.....           | 18 |
| Figura 11: Pestaña de simulación.....  | 19 |
| Figura 12: Simulación de una interestación sin datos reales.....                                       | 19 |
| Figura 13: Comparación de los valores simulados y los reales medidos en una interestación ...          | 20 |
| Figura 14: Categorías a incluir en el Excel de resultados de la simulación .....                       | 20 |
| Figura 15: Gráfico de una interestación con deriva .....   | 21 |
| Figura 16: Subapartado de cálculo sin simulación en el apartado de predicción de salidas del ATO ..... | 22 |
| Figura 17: Apartado para cambiar las características del tren.....                                     | 23 |
| Figura 18: Apartado para cambiar la curva de tracción máxima.....                                      | 23 |
| Figura 19: Sección para guardar los archivos generados .....   | 24 |
| Figura 20: Apartado de carga de archivos.....  | 25 |
| Figura 21: Explorador de archivos .....  | 26 |
| Figura 22: Predicción individual de tiempo y consumo.....  | 27 |
| Figura 23: Predicción individual de la salida del ATO .....  | 27 |
| Figura 24: Tablas del Excel de entradas para predecir tiempos y consumos .....                         | 28 |
| Figura 25: Predicción consecutiva con dos modelos diferentes .....                                     | 29 |
| Figura 26: Gráfico combinado de una simulación .....   | 29 |
| Figura 27: Tabla obtenida de resultados de simulación .....  | 30 |
| Figura 28: Categorías a incluir en el Excel de resultados de la simulación .....                       | 30 |
| Figura 29: Comparación entre simular teniendo datos reales o solo los datos de la línea .....          | 31 |
| Figura 30: Ejemplo del desplegable para seleccionar los ejes del gráfico .....                         | 32 |
| Figura 31: Excels obtenidos de la optimización.....  | 33 |
| Figura 32: Uso del apartado de 'Análisis' para representar los resultados de la simulación .....       | 33 |

|   |    |
|---|----|
| Figura 33: Gráfica de la predicción de tiempo y consumo desde la pestaña de 'Análisis' .....  | 34 |
| Figura 34: Configuración de la pestaña de 'Tiempo/consumo' para simular el caso de ejemplo    | 35 |
| Figura 35: Marchas óptimas del caso de ejemplo .....  | 36 |
| Figura 36: Predicción individual de la marcha del caso de ejemplo .....                       | 37 |
| Figura 37: Configuración de la pestaña de 'Salidas ATO' para simular el caso de ejemplo ..... | 37 |
| Figura 43: Simulación del caso de ejemplo.....  | 38 |

# Índice de tablas

|   |    |
|---|----|
| Tabla 1: Errores porcentuales de los modelos utilizados para la predicción de las salidas del ATO ..... | 9  |
| Tabla 2: Requisitos principales .....   | 11 |
| Tabla 3: Requisitos adicionales.....  | 12 |
| Tabla 4: Marchas originales en el caso de ejemplo.....  | 35 |
| Tabla 5: Soluciones a la predicción de tiempos y consumos para el caso de ejemplo.....                  | 36 |
| Tabla 6: Resultados de la simulación del caso de ejemplo .....  | 38 |
| Tabla 7: Resumen de resultados del caso de ejemplo .....  | 39 |



# 1. Introducción

La tendencia hacia la inteligencia artificial y el machine learning es natural en sectores como el ferroviario, en el que se recogen una gran cantidad de datos. Estos datos se pueden aprovechar para multitud de usos. El más representativo es el mantenimiento, ya que con históricos se puede predecir el desgaste. Empresas como D-RAIL ofrecen servicios para el mantenimiento de infraestructuras usando inteligencia artificial [1].

Pero las capacidades de la inteligencia artificial no se quedan en el simple análisis de datos. Soluciones como el reconocimiento de voz o imágenes se pueden utilizar en el ferrocarril. Ya existen proyectos con cámaras instaladas en el tren que son capaces de detectar obstáculos, como el tren de Thales en Karlsruhe [2], o el de Rail Vision y SBB en Gossau [3]. En ellos se van identificando los elementos que va encontrando el tren, pudiendo detectar incluso la posición de las agujas de los cambios.

Recientemente, en [4], se ha hecho una revisión de tecnologías existentes, junto con las necesidades y expectativas a las que hace frente el sector en temas de inteligencia artificial.

Desde la Unión Europea se está tratando de regular los usos de la inteligencia artificial en todos los ámbitos. Su uso en ferrocarriles se trata brevemente en *Artificial intelligence in transport* [5], en donde se centra en sus usos para facilitar el mantenimiento.

La Comisión Europea, a través de los programas Horizonte Europa y Europa Digital, tiene el objetivo de invertir 1 billón de euros al año en inteligencia artificial [6]. Este apoyo a la IA se ha visto reflejado en las estrategias nacionales de los respectivos Estados Miembros. En España con la Estrategia Nacional de Inteligencia Artificial (ENIA), presentada a finales de 2020, con la que se prevé una inversión de 600 millones de euros. La importancia de la IA en el transporte se recoge en el eje estratégico 5 [7].

Junto a la inteligencia artificial, se aprecia una gran inclinación hacia la conducción automática en metropolitanos. De todos los kilómetros que hay de líneas de metro, en 2018 había un 10% totalmente automatizado, esperándose un 48% para 2022 [8]. En la Figura 1 se puede ver el crecimiento esperado que va a haber de kilómetros con GoA4. Aunque la mayor parte es debido a la región Asia-Pacífico, si se tiene en cuenta la población de cada zona, Europa es la que tendrá mayor número de kilómetros por habitante.

Esta tendencia centra la atención en el ATO (Automatic Train Operation). Desde la Unión Europea se está tratando de incorporar un ATO al ERTMS. En el programa de la ERA, para el periodo 2021-2023, hay planificado tener a finales de 2021 las especificaciones del ATO GoA2. En este mismo momento se prevé una revisión de las ETIs, por lo que es de esperar que estas especificaciones se incluyan en la versión de 2022 [9].



Figura 1: Crecimiento esperado de kilómetros con GoA4 en el mundo [8]

Dentro de la iniciativa Shift2Rail se encuentra el proyecto X2RAIL-4. En él el objetivo es la obtención de un ATO GoA4 para 2023. Según la planificación y cuentas anuales de Shift2Rail para el año 2021, en 2023 deberían finalizar las pruebas de un ATO GoA3 o 4 en líneas piloto [10].

Tener un ATO, aunque sea GoA2 y esté menos automatizado, ya supone un gran avance frente a no llevarlo. Incorporar un ATO a una línea aporta numerosos beneficios, como un aumento de la capacidad y la puntualidad de los trenes. Aunque lo especialmente relevante hoy en día, bajo el marco del *Green Deal* y el objetivo de la UE de ser *carbon neutral*, es el aumento en la eficiencia energética que se consigue. Por ejemplo, en Metro de Madrid gracias al ATO se consiguió ahorrar de media un 18% de energía [11]. En Metro, en todas las líneas hay ya un GoA2 [12], y se está llevando a cabo una renovación progresiva del ATO [13].

El ahorro que se puede conseguir está ligado al diseño de las marchas del ATO. Para llegar a las mejores marchas es necesario un proceso de optimización, para lo que es preciso disponer de algoritmos que las analicen. Este diseño se realiza offline, seleccionando las marchas que llevan a un menor consumo, cumpliendo además los requisitos de tiempo de recorrido deseado y de confort del viajero. Aquí entra en juego el *eco-driving*, que manteniendo la calidad del servicio intenta utilizar la menor energía posible.

### 1.1 Estado del arte

Dentro del *eco-driving*, la marcha base es en llano, complicándose la resolución del problema cuando hay declividades, o velocidades máximas que el tren tiene que respetar [12]. Una dificultad adicional en los simuladores que se utilizan para evaluar las marchas es el cómo modelar el equipo de ATO. Cada fabricante tiene su ATO propio. Hay casos en los que el fabricante comparte el funcionamiento interno, para realizar estudios con modelos muy detallados, en los que se replica el comportamiento del ATO de forma precisa. ALSTOM tiene en marcha un proyecto con un ATO en trenes regionales. Es un proyecto en el que participan varias instituciones, algunas de las cuales tratarán de incorporar soluciones de IA [13].

Sin embargo, no es habitual disponer de esta información. Thales, por ejemplo, tiene colaboraciones con otras entidades, pero recientemente no ha habido ninguna que incorporara un ATO [14]. Una alternativa para poder simular su comportamiento, cuando no se conoce el funcionamiento interno, es usar un modelo tipo caja negra. Una forma de obtener este modelo es

a través de inteligencia artificial. Hay otros ámbitos, como la biología, en los que en los que la IA se está utilizando para modelar sistemas complejos [15]. En el caso de la conducción automática de trenes, se puede ajustar un modelo a través de medidas registradas en el tren, y utilizarlo para imitar el comportamiento de un ATO.

Desde UNIFE se promueve la utilización de la inteligencia artificial para automatizar la conducción, con vistas a flexibilizar la operación y aumentar la capacidad [15]. El informe de 2019 también hace hincapié en la importancia de automatizar la conducción a través de inteligencia artificial [16].

El uso de IA para temas relacionados con la conducción se puede ver en [12], donde únicamente se menciona la posibilidad de utilizarla, y en [13]. Concretamente en [13], se utiliza una red neuronal para hacer predicciones de las desviaciones de los maquinistas frente a un perfil ideal de conducción. En [14] se utiliza también una red neuronal, en este caso para predecir consumos del tren, obteniéndose errores medios del 1.5%.

Dentro de la inteligencia artificial, en este trabajo se utiliza el machine learning, que entrena algoritmos para hacer predicciones o clasificaciones. A su vez, dentro del machine learning, se utiliza aprendizaje supervisado, y dentro de este, regresiones. Los problemas que se resuelven en este trabajo buscan obtener unas salidas a partir de unas entradas. A la hora de entrenar el algoritmo, éste aprende viendo las entradas y sus salidas asociadas.

Se han utilizado principalmente redes neuronales. En ellas hay neuronas, que están agrupadas en capas. Cuando a la red le llegan las entradas, las va haciendo pasar por las diferentes capas y neuronas. En este camino los valores se van combinando, y multiplicando por diferentes parámetros, hasta que se obtiene la salida, que es la predicción.

El otro algoritmo al que se hace referencia en este trabajo es el *RandomForest* [17], que utiliza varios árboles de decisión. Los árboles de decisión empiezan en un punto común, a partir del cual se van dividiendo las ramas. El avanzar por una rama o por otra depende de si los valores de las entradas están por encima o por debajo de un umbral. Los puntos finales son las hojas, que contienen el valor de la predicción. A partir de las entradas, el algoritmo va avanzando por el árbol hasta que llega a una hoja. El *RandomForest* entrena varios árboles, cogiendo como predicción la media de lo que predigan todos.

Estos modelos de machine learning presentan la ventaja de su rapidez. Las predicciones con ellos son casi instantáneas. Surge en este punto la solución de utilizar estos modelos para replicar el funcionamiento de sistemas más complejos y más lentos. Dado el gran volumen de datos disponibles, se pueden entrenar modelos de machine learning y utilizarlos como aproximación previa a una simulación detallada.

La fase de entrenamiento suele estar separada del uso final de los modelos. El proceso comienza con la obtención del *set* de datos que se va a utilizar para entrenar. Debe contener un número de puntos suficiente como para que el modelo vea, idealmente, todos los contextos a los que se va a enfrentar. Este *set* de datos se divide a su vez en los utilizados para entrenar el modelo, y los utilizados para calcular errores. El modelo sólo ve el segundo *subset* cuando ya está entrenado. Esta división asegura que los errores cometidos al predecir se calculan con puntos que el modelo no ha visto antes.

Durante el entrenamiento, se comparan las predicciones que hace el modelo con las salidas verdaderas. En función de lo lejos que se haya quedado, se corrigen más o menos los ajustes internos, para que la predicción se vaya acercando a los resultados correctos.

Al terminar cada iteración por los datos, se compara el error con el umbral marcado por el programador. El entrenamiento termina cuando el modelo alcanza la precisión de predicción

exigida. Se marca un número máximo de iteraciones, ya que hay casos en los que el error no baja por debajo de este umbral fijado. Pasadas las iteraciones, el entrenamiento para, independientemente de los errores que tenga el modelo.

El siguiente paso es la validación, que se hace con el *subset* de datos específico mencionado anteriormente. Se calculan los errores del modelo al predecir estos puntos nuevos, lo que permite evaluar su comportamiento.

Los parámetros internos se van ajustando en el entrenamiento. Pero hay otras variables, como el número de capas y neuronas de la red neuronal, que también influyen en los errores obtenidos. Los valores que toman estas variables son indicados por el programador. En el proceso se entrenan multitud de modelos, modificando esta configuración, y seleccionando los que tienen errores más bajos en este proceso de validación.

Una vez entrenados, es posible guardar los modelos, para volver a utilizarlos con los mismos parámetros y errores que tenían cuando fueron entrenados. El desarrollo y entrenamiento de los modelos de machine learning utilizados en este trabajo se puede ver en [18]. En el apartado 3 se ha incluido una descripción más detallada de los modelos utilizados, y los errores que tienen.

## 2. Objetivos y planificación

### 2.1 Objetivos

El principal objetivo del trabajo es el desarrollo de una herramienta software que pueda ser utilizada por el diseñador de marchas de conducción automática ATO y que cumpla las siguientes funciones:

- Predicción de tiempos y consumos para el diseño y optimización de marchas de conducción ATO.
- Modelado de un equipo de conducción ATO genérico, y simulación de su comportamiento ejecutando una marcha en casos reales.

En ambos casos, las predicciones se realizan con modelos de machine learning, con el objetivo de que sean lo más rápidas y precisas posible. Los objetivos están muy relacionados con los requisitos que tiene la herramienta, que se pueden ver en el apartado 4.

### 2.2 Planificación de tareas

Se incluyen las tareas identificadas y su planificación temporal. En los apartados de diseño de la arquitectura, y de implementación de los algoritmos, se planificó una extensión, paralela a las pruebas de funcionamiento. Dependiendo de los resultados de las pruebas podía ser necesario reabrir estos apartados.

Al haber ido realizando pruebas preliminares según se iban completando los apartados, estas pruebas finales de funcionamiento no revelaron fallos graves. Aunque apareció la necesidad de pequeñas modificaciones, éstas no supusieron un retraso apreciable en la realización del proyecto.

|  | Semana |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |
|--|--------|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|----|
|  | 1      | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
| Especificación de requisitos                         |        |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |
| Elección de la plataforma de desarrollo              |        |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |
| Análisis de alternativas de diseño de la herramienta |        |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |
| Diseño de arquitectura software                      |        |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |
| Implementación de los algoritmos                     |        |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |
| Pruebas de funcionamiento                            |        |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |
| Redacción de la memoria                              |        |   |   |   |   |   |   |   |   |    |    |    |    |    |    |    |    |

### 3. Modelos de machine learning integrados

Una parte fundamental de la herramienta son los modelos de machine learning que integra, principalmente redes neuronales. Estos modelos, como se comentó en el apartado 1.1, están almacenados en archivos independientes, teniendo los errores de predicción que presentaban cuando fueron entrenados. Para el guardado se ha utilizado el módulo *Pickle*, que permite englobar todo lo que es el modelo en un único archivo. Este archivo se puede abrir con la herramienta y utilizar el modelo que contiene.

A continuación se describen los dos tipos de modelos utilizados. Estos tipos corresponden a las dos partes que tiene el objetivo principal del proyecto:

- Predicción de tiempos y consumos
- Predicción de las salidas del equipo ATO

En los dos casos se comentarán las entradas y salidas de los modelos, y los errores esperables al hacer predicciones con ellos.

#### 3.1 Modelo para la predicción de tiempo y consumo

Se trata de una red neuronal, en la que las entradas son las consignas de conducción del ATO: distancia de deriva y reducción de velocidad. La distancia de deriva marca la distancia a la estación, en metros, a la que empieza la deriva. La reducción de velocidad toma valores discretos, representando los km/h por encima o por debajo de la velocidad máxima que se pueden alcanzar. El tren puede conducir a la velocidad máxima, o a la máxima menos esta reducción. En algunos casos se admite circular hasta 3 km/h por encima de la velocidad máxima, pero siempre garantizando la seguridad.

Las salidas que predice la red neuronal son el tiempo de recorrido, en segundos, y el consumo, en kWh, esperables. La red actualmente integrada en la herramienta tiene unos errores medios de un 0.04% en el tiempo y un 0.34% en el consumo [18].

En los errores se trata del MAPE, *mean absolute percentage error*, calculado durante el entrenamiento de los modelos con el *set* de datos apartado específicamente para este fin. Se incluye a continuación la ecuación de este error.

$$MAPE = \frac{\sum | \frac{\text{valores reales} - \text{predicciones}}{\text{valores reales}} |}{n^{\circ} \text{ datos}} \cdot 100$$

También es posible integrar modelos con algoritmos diferentes. En la herramienta, además de la red neuronal, se ha incluido un *RandomForest*. Este modelo tiene errores más bajos, 0.01% en tiempo y 0.05% en consumo. Sin embargo, el preferido por defecto es la red neuronal, ya que tiene mayor capacidad de generalización.

Según un análisis realizado, a la hora predecir valores intermedios de las consignas, que los algoritmos no hayan visto antes, la red neuronal es capaz de aplicar lo que ha aprendido y acercarse a la solución correcta. El *RandomForest* en este caso se limita a predecir las salidas de la consigna más próxima que ha aprendido en el entrenamiento. En la Figura 2 se puede ver esta comparación entre los algoritmos. Tenían que predecir la línea negra, que ninguno de los dos había visto en el entrenamiento. La red neuronal (NN) es capaz de predecir dónde tienen que estar las salidas, mientras que el *RandomForest* (RF) predice la línea más próxima que conoce.

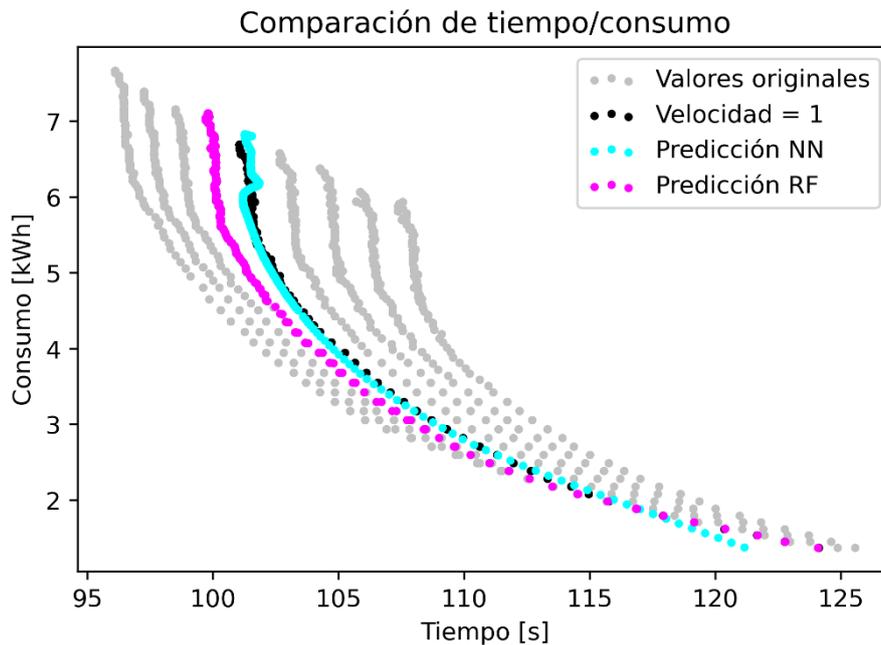


Figura 2: Comparación del comportamiento de la red neuronal y el RandomForest

Viendo estos resultados se decidió utilizar la red neuronal como primera opción. El *RandomForest* se añadió para comprobar que se podían cargar otro tipo de algoritmos, y para tener un segundo modelo con el que comprobar el requisito 4 (cambiar el modelo para predicción utilizado).

### 3.2 Modelos para la predicción de la salida de ATO

Se han utilizado únicamente redes neuronales, aunque en un futuro sería posible entrenar modelos con algoritmos diferentes. En todos los casos, las entradas a la red neuronal son las mismas:

- Velocidad actual del tren
- Aceleración actual del tren
- Velocidad máxima del ATP y del ATO
- Velocidad y distancia objetivo del ATP y del ATO
- Gradiente

En el proceso de entrenamiento, las redes neuronales se evaluaron utilizando el código del simulador. Los errores cometidos a la hora de parar el tren en la estación hicieron que se desarrollaran varias alternativas. La primera, y más simple, con la que se empezó el proceso, fue utilizar una única red neuronal, que predice las salidas del ATO durante todo el recorrido.

A continuación, pasó a utilizarse más de una red neuronal, obteniendo a partir de las predicciones una única salida. Las dos variantes que integra la herramienta son las siguientes:

- Utilizar dos redes neuronales, una que prediga la parte de tracción y otra la parte de freno. Las entradas se pasan a ambas redes a la vez. Como resultado final se coge el mínimo entre lo que predigan las dos. La red de freno se ha entrenado para predecir un valor muy alto cuando la salida es de tracción, y la red de tracción para predecir 0 cuando la salida es de freno. De esta forma, con una salida que debería ser de tracción, el valor más bajo es el de la red de tracción, y para las salidas de freno ocurre de forma análoga con la red de freno. En la Figura 3 se puede ver un esquema del funcionamiento de esta alternativa.

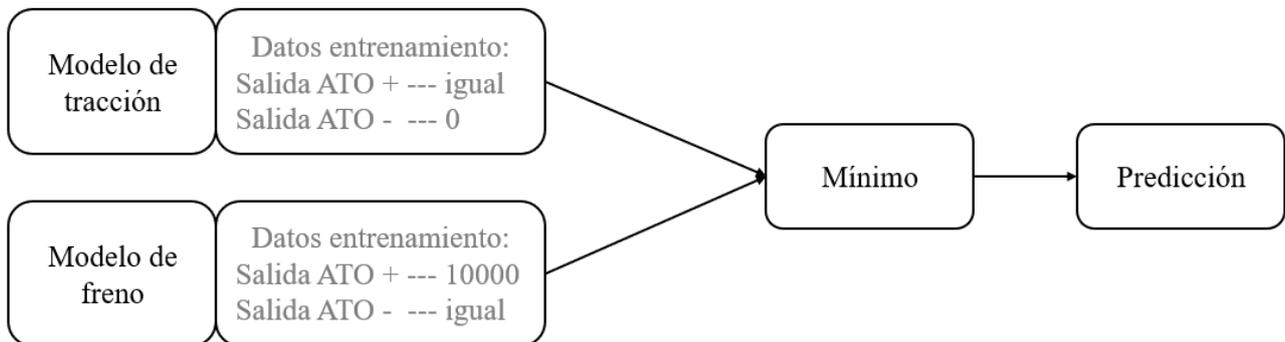


Figura 3: Predicción de ATO usando una red de tracción y otra de freno

- Utilizar dos redes neuronales, una de ellas es del primer tipo (red única), una red que predice durante todo el recorrido; y la otra es una red auxiliar, que predice únicamente la curva de freno cuando el tren llega a menos de 150m de la estación de destino. Desde que se llega a 150m de la estación, se predice de forma simultánea con ambas, cogiéndose la salida más baja. En la Figura 4 se muestra un esquema de cómo esta alternativa obtiene las predicciones. Si el modelo principal no es capaz de frenar lo suficiente, la red de curva de freno da salidas más bajas y para el tren. Las entradas se pasan sólo a la red principal o a ambas en función de la distancia a la estación de destino.

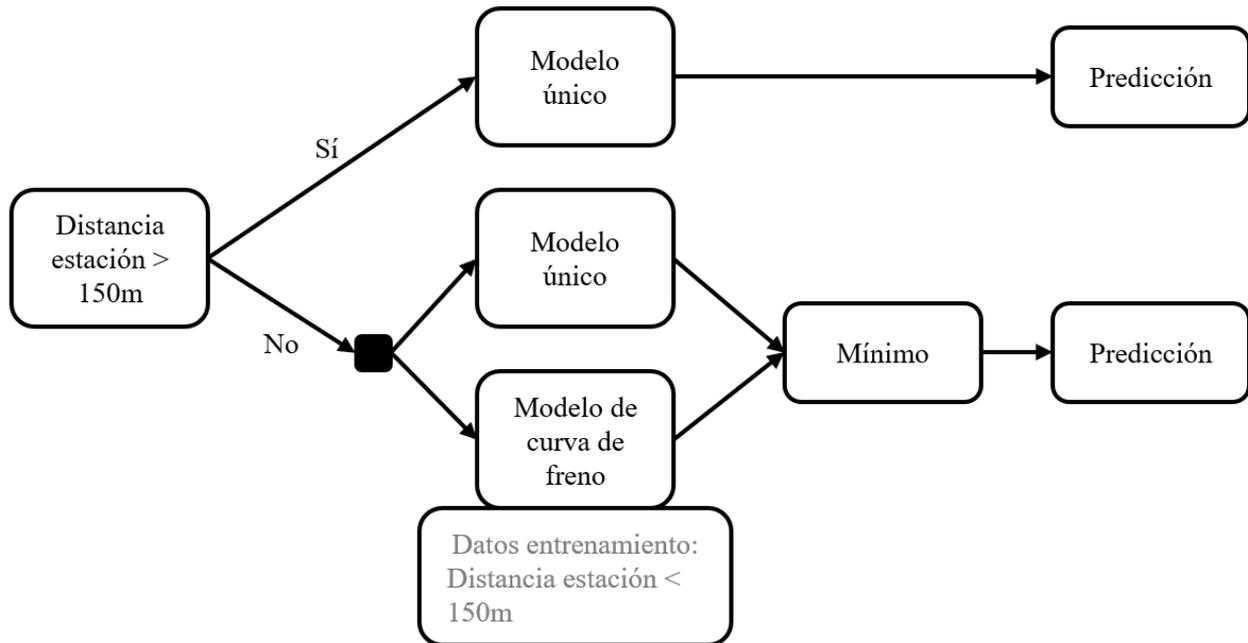


Figura 4: Predicción de ATO utilizando una red auxiliar de curva de freno

Todos estos modelos se evaluaron, al igual que los que predicen tiempos y consumos, con un *set* de datos diferente, al terminar el proceso de entrenamiento. Sin embargo, y por ser los errores más relevantes para esta aplicación, en la Tabla 1 se incluyen los valores de los errores obtenidos en la simulación. Estos errores son de tiempo y distancia al punto de parada. Para calcular el error de tiempo, se han comparado los resultados de las simulaciones con datos reales disponibles en el mismo recorrido. El error en distancia mide lo que el tren en la simulación se desvía del punto de parada.

El error en ambos casos es el MAPE, cuya ecuación aplicada a esta situación sería

$$MAPE = \frac{\sum \left| \frac{\text{valores reales} - \text{valores simulados}}{\text{valores reales}} \right|}{n^{\circ} \text{ estaciones}} \cdot 100$$

Aunque se entrenaron multitud de modelos, en la Tabla 1 sólo se muestran los errores de los modelos incorporados en la herramienta.

Tabla 1: Errores porcentuales de los modelos utilizados para la predicción de las salidas del ATO

|                        | Tiempo [%] | Distancia [%] |
|------------------------|------------|---------------|
| Modelo único           | 4.7        | 8.2           |
| Tracción + freno       | 0.3        | 5.8           |
| Único + curva de freno | 0.1        | 1             |

Como se puede observar, la forma más exacta de predecir es utilizando la alternativa que incorpora una curva de freno adicional para asegurar la parada. Las otras alternativas se han incluido por los posibles desarrollos que puedan tener. La herramienta está abierta a futuras

Desarrollo de una herramienta de apoyo al diseño y optimización de conducción automática de trenes en la que se integren modelos de machine learning

TRABAJO FIN DE MÁSTER

---

mejoras de los modelos, por lo que se han mantenido las opciones con más error. La de un único modelo presenta la ventaja de la sencillez a la hora de entrenarla, ya que no hay que dividir los datos y tratarlos varias veces.

## 4. Requisitos

Dentro de la herramienta, como ya se ha comentado, se diferencian dos grandes secciones de predicción: por un lado la predicción de tiempos y consumos (*TC*) y por otro la predicción de salidas del ATO (*ATO*), y simulación de la marcha del tren (*Sim*). Los requisitos se han definido teniendo en cuenta esta división. Adicionalmente, se ha añadido un apartado de representación gráfica (*An*) para visualizar los datos de forma independiente.

Los requisitos principales se enumeran a continuación en la Tabla 2.

Tabla 2: Requisitos principales

|    | Categoría | Requisito   |
|----|-----------|---|
| 1  | TC        | Mostrar de forma individual la predicción de tiempo y consumo           |
| 2  | TC        | Introducir datos a través de fichero Excel                              |
| 3  | TC        | Incorporar las predicciones realizadas y entradas a Excel de resultados |
| 4  | TC        | Cambiar el modelo para predicción utilizado                             |
| 5  | ATO       | Mostrar de forma individual la predicción de salida ATO                 |
| 6  | ATO       | Introducir datos a través de fichero Excel                              |
| 7  | ATO       | Incorporar las predicciones realizadas y entradas a Excel de resultados |
| 8  | ATO       | Cambiar el modelo para predicción utilizado                             |
| 9  | Sim       | Cambiar los datos del tren (masa, coefs A B C)                          |
| 10 | Sim       | Gráficas básicas de resultados  |
| 11 | Sim       | Sacar Excel con resultado de la simulación                              |
| 12 | Sim       | Simulación de interestación, con o sin datos reales de circulación      |
| 13 | An        | Elegir las categorías a incluir en el gráfico                           |

Junto con la realización de las predicciones, se ha priorizado que la herramienta permita cambiar los modelos de machine learning utilizados. Esta característica obedece a la posibilidad de mejorar los modelos, y obtener otros en un futuro con errores más bajos. Tiene que ser posible realizar predicciones o simulaciones consecutivas con modelos diferentes, seleccionándolos desde la herramienta.

Aunque se ha escindido la parte de simulación (*Sim*) de la de salidas del ATO (*ATO*), el requisito de poder cambiar los modelos de predicción es común a ambas. El apartado de simulación se construye sobre la base del que predice las salidas del ATO.

Para poder predecir a la vez gran cantidad de puntos, se ha añadido el requisito de que se puedan introducir los inputs a los modelos a través de un fichero Excel. En este archivo se encontrarán las entradas a la red neuronal, que predecirá las salidas correspondientes.

En el subapartado de simulación se podrán obtener los resultados en un fichero Excel o en forma de gráficas. Las gráficas básicas que se deben poder obtener son las siguientes:

- Espacio-velocidad
- Espacio-tiempo
- Espacio-salidas de ATO
- Espacio-gradiente

Siendo el espacio el correspondiente a la longitud de la interestación simulada.

El requisito nº9 permite cambiar las características del tren. Este requisito está asociado al de cambiar de línea (14). Al poder cambiar de línea, se hacía necesario también poder cambiar las variables del tipo de tren que se utiliza. El 9, sin embargo, tiene más prioridad que el 14. Si las características del tren se definen como variables internas de la herramienta, no es posible cambiarlas, ya que una vez generado el ejecutable .exe no se tiene acceso al código. El por qué el 14 es menos prioritario se explica más adelante tras los requisitos adicionales.

Dentro del requisito nº11 se pueden diferenciar a su vez dos partes, obtener un Excel completo con todas las salidas, o un Excel reducido, en el que se puedan elegir las categorías que se quiere que aparezcan.

Otros requisitos de menos prioridad se presentan en la Tabla 3. Las prioridades están establecidas en una escala de 1 a 3, siendo el 1 correspondiente a los más prioritarios, que se han separado en la Tabla 2.

Tabla 3: Requisitos adicionales

|    | Categoría | Requisito   | Prioridad |
|----|-----------|---|-----------|
| 14 | Sim       | Cambiar el Excel de datos de la línea                             | 2         |
| 15 | An        | Cambiar el Excel de datos   | 2         |
| 16 | Misc      | Seleccionar donde guardar los Excels/gráficos                     | 2         |
| 17 | TC        | Mostrar el cluster de óptimas                                     | 2         |
| 18 | Sim       | Elegir las salidas a incluir en el gráfico                        | 3         |
| 19 | Sim       | Simulación de varias interestaciones seguidas y gráficas de todas | 3         |
| 20 | Misc      | Explorador de archivos  | 3         |
| 21 | TC        | Gráfica de resultados   | 3         |

La categoría *Misc* corresponde a requisitos que son generales de la herramienta, pudiendo usarse en todos los apartados.

Los requisitos 14 y 15 buscan que este cambio sea lo más cómodo posible. Se puede cambiar el Excel aún sin estos requisitos, por lo que se les ha dado menos prioridad. Pero habría que entrar en la carpeta de la herramienta, buscar el archivo y sustituirlo por otro con el mismo nombre y estructura. Con este requisito, se pueden incorporar Excels con nombres diferentes, pudiendo seleccionar el que se quiere utilizar.

Para el requisito 14 sí que está definida la estructura del Excel, que debe tener siempre las mismas columnas, que contienen los datos necesarios para hacer la simulación. Con el requisito 15, sin embargo, se busca que se puedan representar gráficas de cualquier Excel, con cualquier número de columnas.

El requisito nº17 se refiere a la posibilidad de diferenciar las entradas que resultan en menores tiempos y consumos. Aunque no se llegan a seleccionar las marchas, la herramienta debe de ser capaz de presentar el espacio de soluciones óptimas.

La parte del explorador de archivos se añadió para facilitar el uso de la herramienta. Es posible acceder a los archivos entrando en la carpeta de la aplicación, pero resulta incómodo buscar entre los ficheros que contiene. En un principio, este requisito (20) se formuló de otra manera. El enunciado original sería:

*Poder cargar Excels y modelos desde la propia herramienta, y limitar al máximo el tener que buscar y abrir la carpeta en la que está la herramienta*

En la fase posterior de organizar la herramienta, decidiendo qué apartados incluir y de qué forma, este requisito se materializó en varias opciones. Una de ellas era un explorador de archivos, barajándose otras como un apartado de *drag and drop*, o una caja para escribir la ruta del archivo. La elegida finalmente fue el explorador, reescribiendo el requisito para darle claridad.

## 5. Plataforma utilizada y diseño

Se ha utilizado el lenguaje Python y el entorno Spyder. Los modelos de machine learning embebidos en la herramienta han sido obtenidos con estos recursos, por lo que la integración más sencilla es en este mismo entorno. Cabe destacar, además, la posibilidad de utilizar el módulo PyQt de Qt [19], pensado para crear aplicaciones y GUIs (Graphical User Interface).

Para diseñar la parte gráfica de la herramienta se ha utilizado Qt Designer, que permite crear ventanas arrastrando los diversos *widgets* disponibles. Para el diseño de ventanas con multitud de elementos es una herramienta útil, ya que permite organizar los *widgets* de forma visual. Utilizando el propio PyQt se puede traducir la interfaz creada con Qt Designer a un fichero *.py*, que a su vez se puede editar con Spyder.

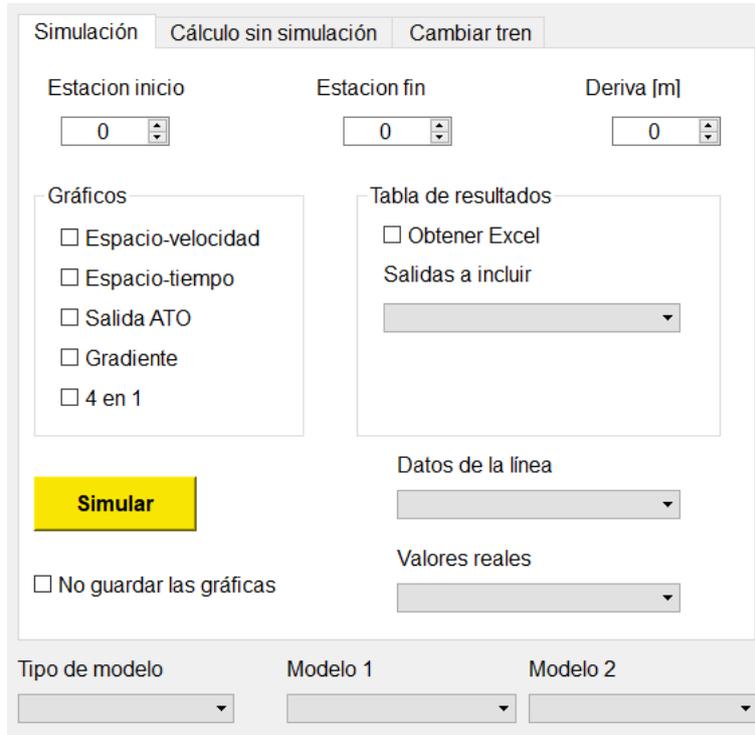
Dados los requisitos se ha decidido estructurar la interfaz como se puede ver en la Figura 5.



Figura 5: Página principal de la interfaz

A la izquierda se encuentran los botones para acceder a cada uno de los apartados. Junto a los ya comentados de Análisis y las predicciones, hay otro especial para cargar o eliminar archivos de la herramienta (apartado 5.4). La zona verde para guardar los archivos se explica con más detalle también en el apartado 5.4. En el espacio de la derecha aparece la pestaña del botón seleccionado. En la Figura 5, al no haber hecho click aún sobre ningún botón, se ve la página por defecto, que es con la que se abre la herramienta.

En el apartado de salidas del ATO se han incluido las predicciones individuales, las simulaciones, y los cambios de datos del tren. En la Figura 6 se muestra el subapartado de *Salidas ATO*. Usando las pestañas superiores es posible cambiar de subapartado. El que se muestra por defecto al pulsar el botón de *Salidas ATO* es el de *Simulación*. El resto de pestañas se pueden ver en el apartado 5.3.



The screenshot shows a software interface with three tabs: 'Simulación', 'Cálculo sin simulación', and 'Cambiar tren'. The 'Simulación' tab is active. It contains several input fields and checkboxes:

- Three spinners for 'Estacion inicio', 'Estacion fin', and 'Deriva [m]', all set to 0.
- A 'Gráficos' section with checkboxes for 'Espacio-velocidad', 'Espacio-tiempo', 'Salida ATO', 'Gradiente', and '4 en 1'.
- A 'Tabla de resultados' section with a checkbox for 'Obtener Excel' and a dropdown menu for 'Salidas a incluir'.
- A yellow 'Simular' button.
- A checkbox for 'No guardar las gráficas'.
- Three dropdown menus for 'Datos de la línea' and 'Valores reales'.
- A 'Tipo de modelo' section with three dropdown menus labeled 'Modelo 1' and 'Modelo 2'.

Figura 6: Pestaña de salidas del ATO

### 5.1 Apartado de análisis

Desde el apartado de *Análisis* se pueden representar gráficas de Excels cargados en la aplicación. En la Figura 7 se puede ver la interfaz utilizada. Se puede modificar el Excel elegido, así como la hoja dentro de éste. Las categorías para los ejes se actualizan automáticamente al cambiar el Excel utilizado. Estas categorías aparecen en un menú desplegable que se muestra al hacer click sobre las cajas que hay bajo el nombre de cada eje.

La opción de elegir la hoja dentro del Excel surgió durante el desarrollo de la herramienta. El requisito original (15) sólo contemplaba poder cambiar el archivo. En las pruebas parciales tras completar esta parte, surgió el problema que llevó a esta modificación.

Es posible leer un Excel con varias hojas, y si todas tienen las mismas columnas, se genera una única tabla de datos, y se representan todos juntos. Pero con Excels que tienen información diferente en cada hoja, se pierde la capacidad de representar los puntos por separado, llevando en algunos casos al fallo de la herramienta.

La solución supone tener que elegir entre representar varias hojas iguales juntas, o permitir leer Excels con hojas diferentes. Finalmente se decidió utilizar esta última, ya que daba más flexibilidad a la herramienta, y aseguraba el funcionamiento adecuado al hacer las gráficas.

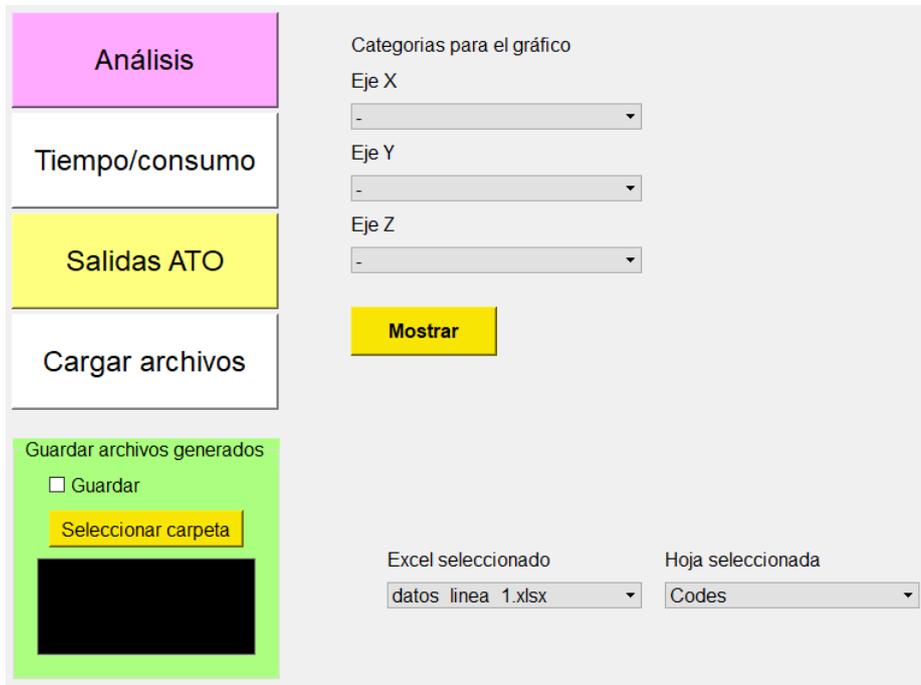


Figura 7: Pestaña de 'Análisis'

Se pueden representar gráficos en 2D o 3D en función de las casillas a las que se les dé valor en los ejes. Todos los gráficos de la herramienta se generan a través de *Plotly* [20], modulo enfocado a la representación gráfica, que incorpora características como la posibilidad de rotar los gráficos en 3D. En la Figura 8 se incluye un ejemplo de gráfico generado con esta pestaña.

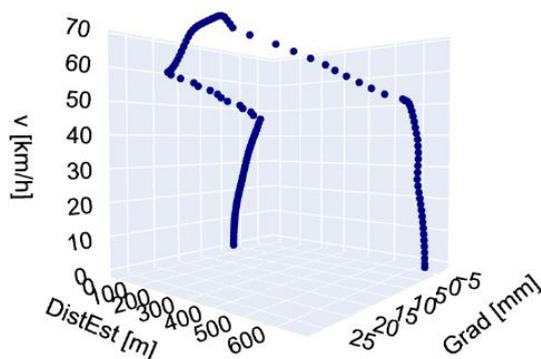


Figura 8: Gráfico generado desde la pestaña de 'Análisis'

## 5.2 Apartado de predicción de tiempo y consumo

Esta sección se ha dividido a su vez en dos partes fundamentales, como se puede ver en la Figura 9. La parte superior permite hacer predicciones y a partir de ellas también optimizaciones. Estas predicciones pueden ser individuales, o introducirse los datos de entrada desde un Excel. En el caso de predicciones individuales, existe la posibilidad de añadir las predicciones realizadas a un Excel de resultados. Si se utiliza la opción de introducir datos a través de un Excel, se puede hacer a continuación una optimización, seleccionando opciones en la caja azul de la derecha como se comenta más adelante.

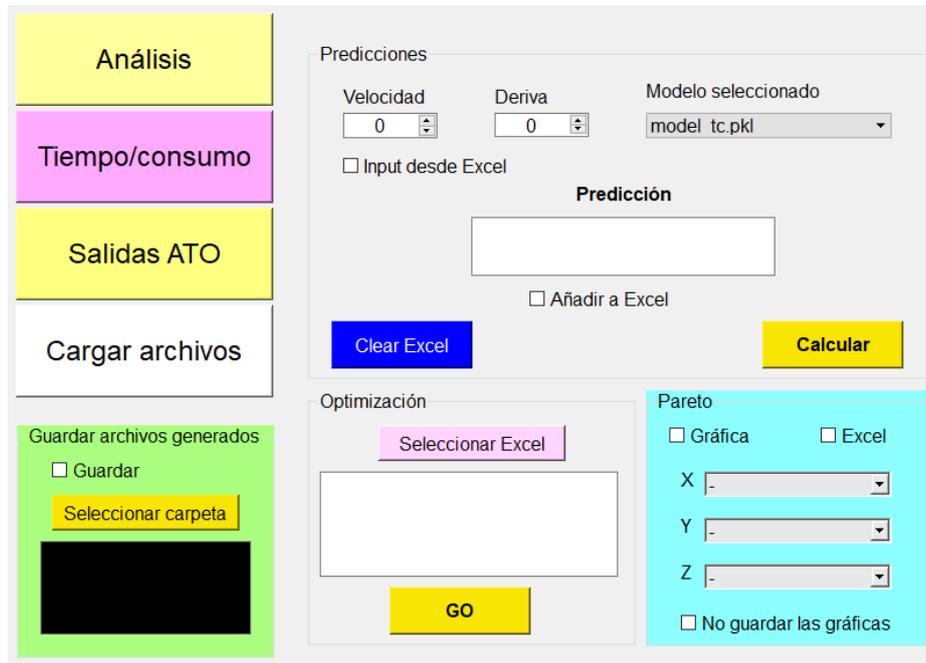


Figura 9: Apartado de predicción de tiempos y consumos

La parte inferior se puede utilizar para hacer optimizaciones de Excels ya existentes. En este caso el fichero debe contener tanto las entradas como las salidas. Estas categorías se explican en el apartado 3. En este mismo apartado también se comentan los tipos de modelos que se pueden utilizar y el empleado por defecto.

El apartado de optimización, resaltado en azul, es común para ambas partes. A través de él se puede obtener el *Pareto front*, conjunto de soluciones óptimas, de una serie de datos, como se puede ver en la Figura 10 en 2D. También es posible obtener el gráfico en 3D si se especifica un eje Z. Adicionalmente, se puede añadir al Excel sometido a optimización una columna que indique si los puntos son óptimos o no. En este último caso, también se genera en el Excel una hoja adicional, que contiene sólo los puntos óptimos. Para elegir una u otra opción, o ambas, se ha de hacer tick en la opción que se prefiera.

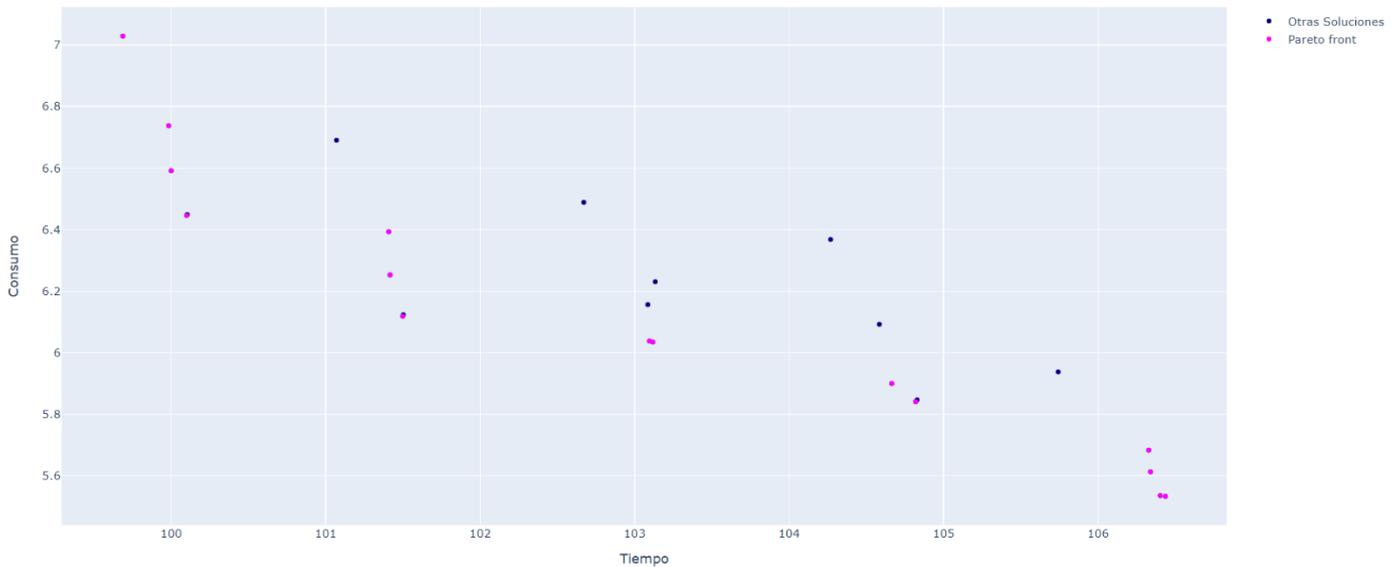


Figura 10: Gráfico obtenido de puntos que conforman el Pareto front entre los predichos

### 5.3 Apartado de predicción de salidas del ATO

Como se ha mencionado anteriormente, este apartado es el más extenso, ya que agrupa varias funcionalidades. En la Figura 11 se puede ver el subapartado de simulación de la marcha de tren. La parte inferior para seleccionar el tipo de modelo es común para todos los apartados. En el punto 3 se relata con más detalle de qué tipo de modelos se trata.

#### 5.3.1 Subapartado de simulación de la marcha del tren

Detrás de la interfaz gráfica se encuentra el código del simulador dinámico de la marcha del tren, en el que se reúne todo lo que puede hacer este apartado. Funciona con incrementos de tiempo de 0.5s. En cada uno de ellos se utilizan las condiciones del tren y características de la vía, obteniéndose el esfuerzo de tracción o freno que intenta aplicar el ATO. Se considera el esfuerzo máximo de tracción que puede dar el tren, modificable en la pestaña mostrada en el apartado 5.3.3.

Las otras fuerzas que entran en el balance son la resistencia al avance ( $A + B \cdot v + C \cdot v^2$ ) y la debida a las declividades. El tren se ha considerado una masa puntual, pudiendo cambiarse su masa y sus coeficientes A, B y C en la pestaña del apartado 5.3.3.

La pestaña de este subapartado en concreto se muestra en la Figura 11. Se ha de especificar el código de la estación de inicio y fin, así como la distancia de deriva. Una distancia de 0 m equivale a no hacer deriva.

A continuación se ha de especificar en qué formato se quieren obtener los resultados, siendo posible obtener gráficas de la marcha del tren, o una tabla con los pasos de simulación. Para elegir qué gráficos o tabla obtener hay que marcar con un tick las opciones.



Figura 11: Pestaña de simulación

Dentro de esta pestaña de simulación se pueden hacer dos tipos de simulaciones, en función de si se tienen datos reales de recorrido o no, pudiendo darse dos supuestos:

- Sólo se rellena el apartado de ‘Datos de línea’, quedando vacío el de ‘Valores reales’: la herramienta entiende que se trata de una simulación aislada, que podría ser incluso de una interestación ficticia, y se limita a correr la simulación y representar los resultados, como se puede ver en la Figura 12.

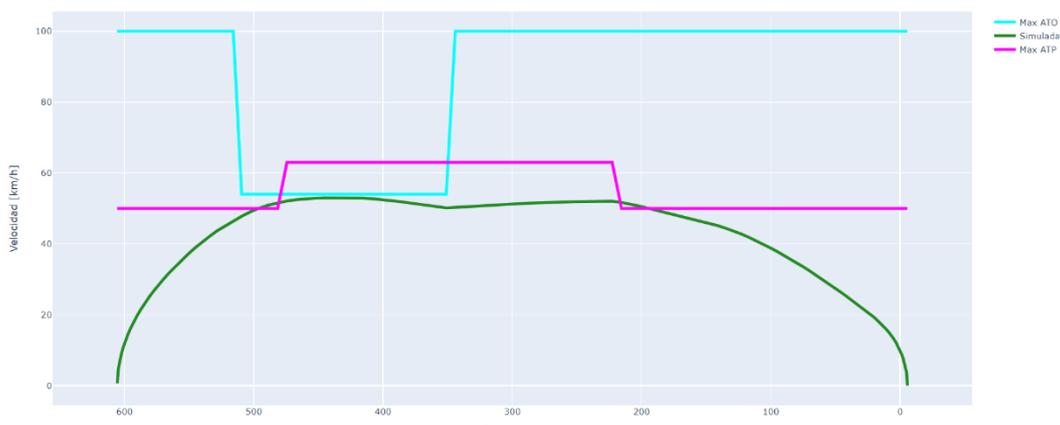


Figura 12: Simulación de una interestación sin datos reales

- Hay un fichero también en ‘Valores reales’: la herramienta simula igual que en el caso anterior, utilizando los datos de línea, pero a la hora de hacer gráficas de los resultados, representa una comparación de los valores reales y los simulados, como se puede ver en la Figura 13.

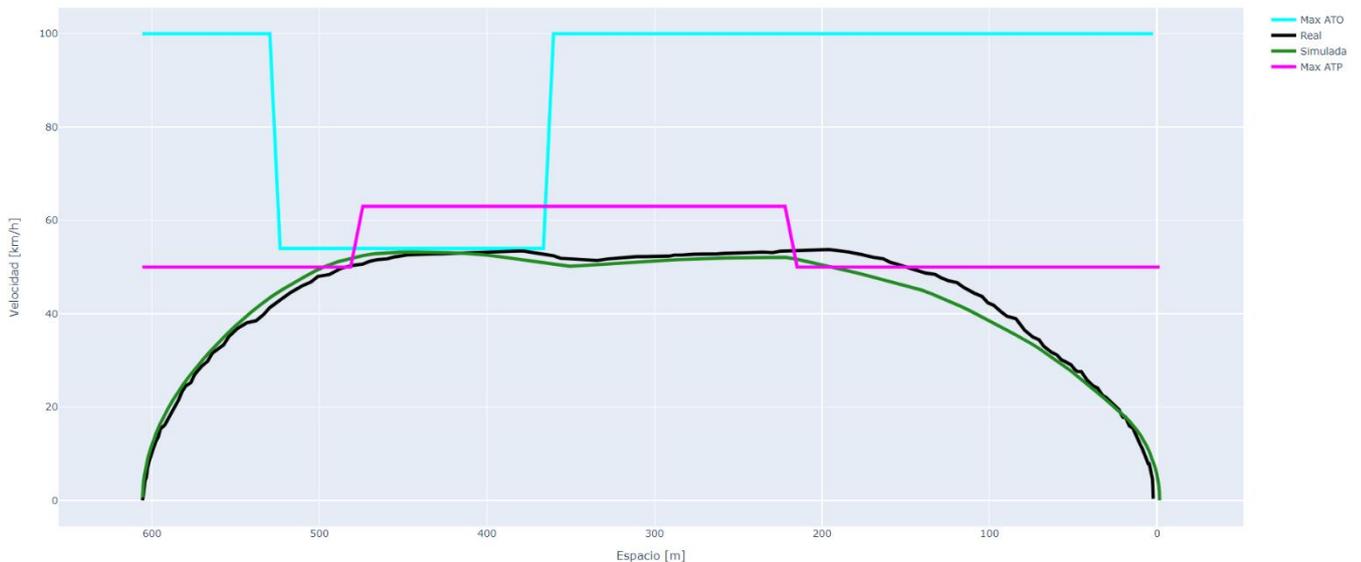


Figura 13: Comparación de los valores simulados y los reales medidos en una interestación

Además de los gráficos hay la posibilidad de obtener el resultado de la simulación en un Excel, en el que se pueden elegir las categorías a incluir. En el menú desplegable se puede seleccionar la opción de ‘Todas’, que es la elegida por defecto si no se selecciona nada. Con ella, en el Excel se incluyen todas las salidas de la simulación, incluyendo parámetros como la resistencia al avance. Al elegir categorías concretas, sólo se incluyen las seleccionadas.

Este desplegable con opción de seleccionar varios ítems se puede ver en la Figura 14. Este tipo de *widget* no es proporcionado por PyQt de forma directa, teniendo que ser desarrollado a partir del desplegable básico. Se ha añadido una función adicional para que el desplegable no se cierre después de seleccionar un ítem. Es posible que se quieran seleccionar varias categorías, y tener que abrir el desplegable tras cada una se vio que era muy incómodo en las pruebas preliminares. Cuando se ha terminado de seleccionar las opciones, basta con hacer click fuera del desplegable para que se cierre.



Figura 14: Categorías a incluir en el Excel de resultados de la simulación

En todos los casos se incluyen a la izquierda dos columnas con los códigos de las estaciones entre las que se está simulando. En la Figura 27 se puede ver un ejemplo de esta tabla, los códigos son las columnas de PSC y NSC, *present* y *next station code*. De querer simular interestaciones ficticias, como es el caso de la Figura 12, se pueden poner códigos de las estaciones aleatorios, siempre que sean consecutivos.

Como se ha comentado, se ha añadido una casilla adicional para poder hacer una deriva. En la Figura 15 se puede ver la gráfica espacio-velocidad de una interestación en la que se ha hecho una deriva a los 500m de la estación de destino.

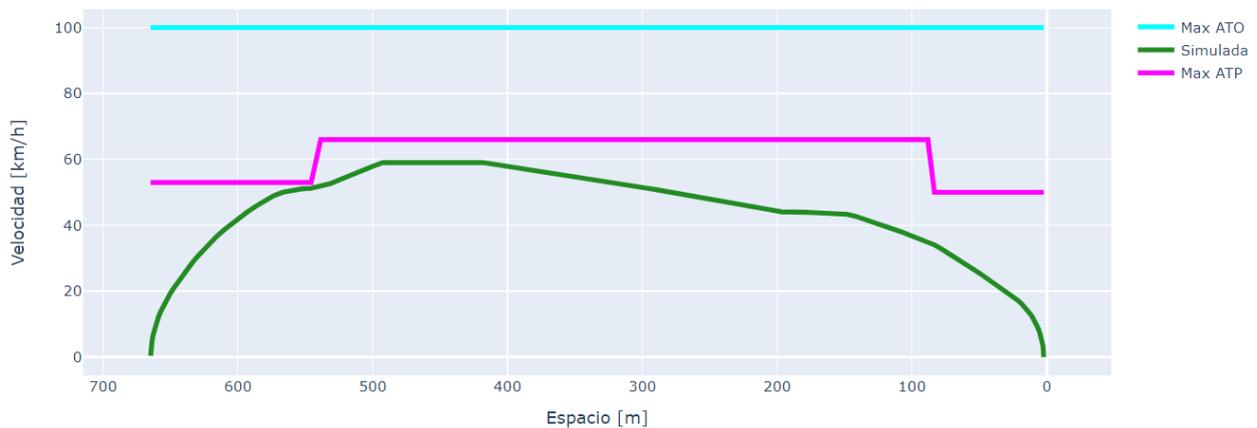


Figura 15: Gráfico de una interestación con deriva

Desde este apartado existe la posibilidad de hacer simulaciones de tramos de interestaciones, de una en una, introduciendo códigos no consecutivos.

### 5.3.2 Subapartado de predicción aislada

En el apartado de *Cálculo sin simulación* se pueden obtener las salidas del ATO de puntos aislados, habiendo, al igual que en la parte de tiempo y consumo, la posibilidad de añadir estos puntos calculados a un Excel, o introducir los puntos a calcular a través de un Excel. En este caso, y como se puede ver en la Figura 16, los datos se introducen a través de una tabla, por la que se puede avanzar con las flechas del teclado.

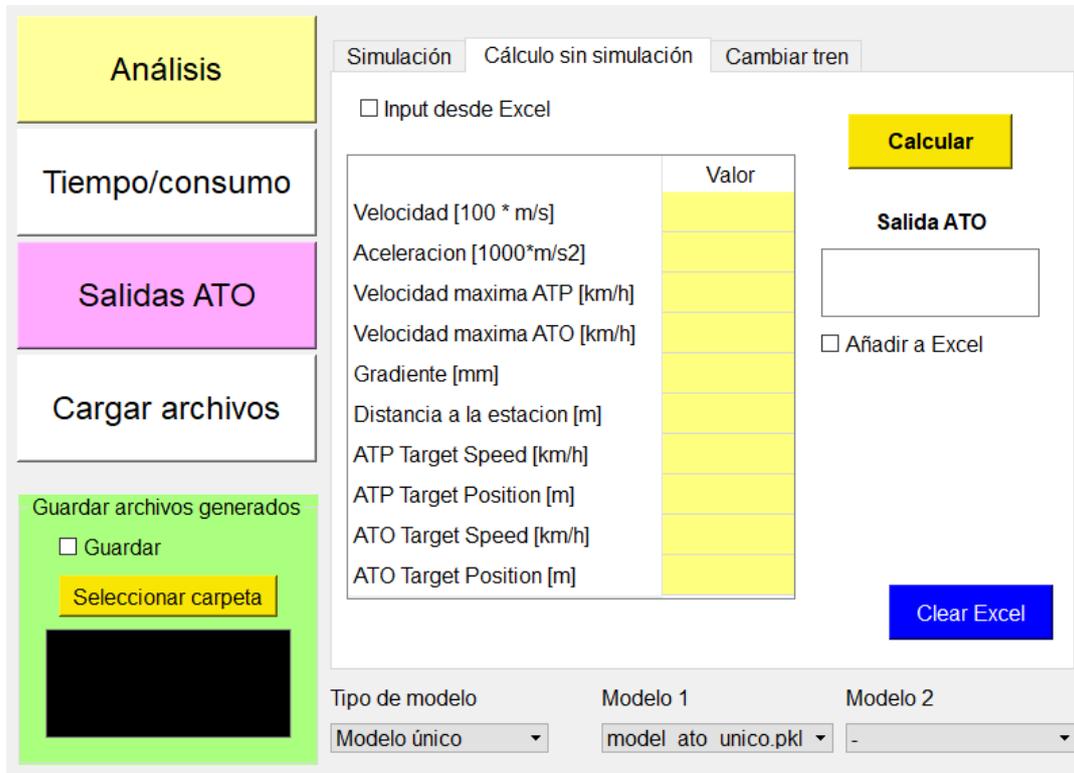


Figura 16: Subapartado de cálculo sin simulación en el apartado de predicción de salidas del ATO

### 5.3.3 Subapartado de modificación del tren

El último apartado, de *Cambiar tren* (Figura 17), se puede utilizar para modificar los parámetros del tren que se utilizan en la simulación:

- Masa, tanto en vacío como inercial
- Coeficientes A, B y C de resistencia al avance
- Tracción máxima que es capaz de dar el tren

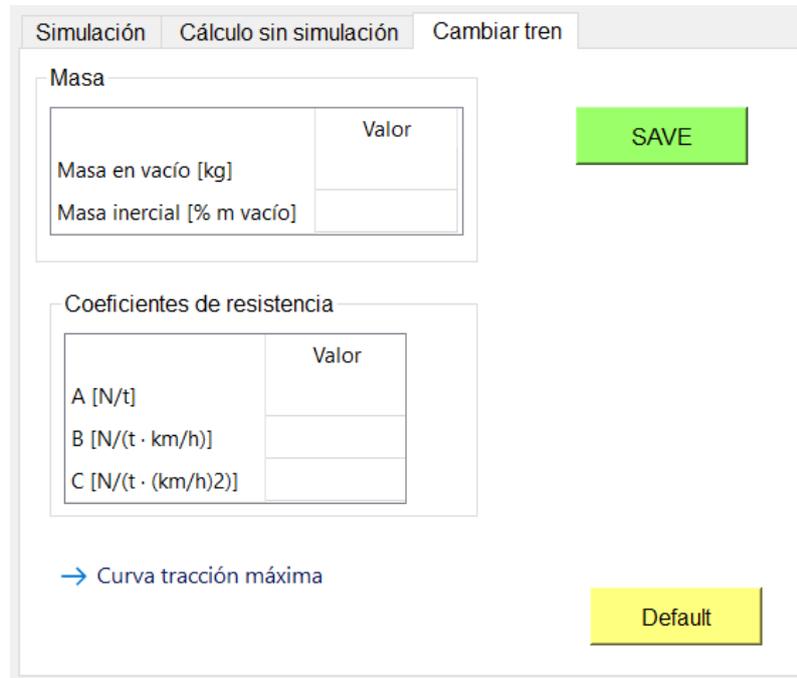


Figura 17: Apartado para cambiar las características del tren

Para registrar los nuevos valores hay que guardar (SAVE) aunque siempre es posible volver a los cargados por defecto (Default). El apartado de tracción máxima se muestra en la Figura 18. En la tabla de la izquierda se pueden introducir los valores que definen el esfuerzo máximo de tracción, apareciendo a la derecha la ecuación asociada que se usará en el simulador, y la representación gráfica.

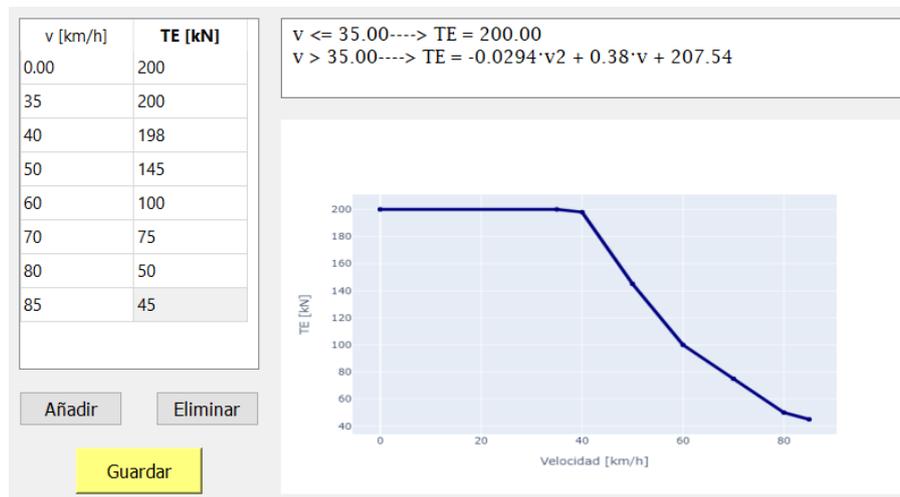


Figura 18: Apartado para cambiar la curva de tracción máxima

#### 5.4 Apartado para el guardado y la carga de archivos

Los ficheros Excel generados se guardan por defecto en la propia carpeta de la interfaz, siendo necesario entrar en ella y buscarlos para copiarlos a otra ubicación. La opción de preguntar tras cada cálculo se rechazó, ya que resultaba molesta al hacer multitud de predicciones y no querer guardar ninguna. La solución finalmente implantada se puede ver en la Figura 19, al marcar el tick de guardar, los Excels generados se guardan:

- En el escritorio si no se indica ninguna carpeta.
- En la carpeta que se seleccione desde el botón de ‘Seleccionar carpeta’. En este caso, en el recuadro negro se muestra la ruta en la que se guardarán los ficheros.

Si no se indica lo contrario, las gráficas se guardan en la misma ubicación que los Excels, aunque esta opción se puede deshabilitar para que únicamente se abran en el explorador.



Figura 19: Sección para guardar los archivos generados

El apartado para la carga de archivos se puede ver en la Figura 20. Hay una lista para los Excels disponibles en la herramienta y otra para los modelos de machine learning, pudiendo añadirse ficheros desde un explorador de archivos y eliminarlos desde las listas. Este apartado facilita el ver qué archivos tiene cargados la herramienta sin tener que abrir la carpeta en la que se encuentra el ejecutable .exe.

Al pulsar *Cargar archivo* se abre un explorador (apartado 5.5), desde el que se puede elegir el Excel o el modelo en *pickle*. Seleccionando las filas de la lista, y con *Eliminar archivo*, se borran los ficheros de la carpeta en la que está la herramienta. Cuando se carga o elimina un archivo, se actualizan las listas de la herramienta, tanto las de este apartado como las de los diversos desplegables.

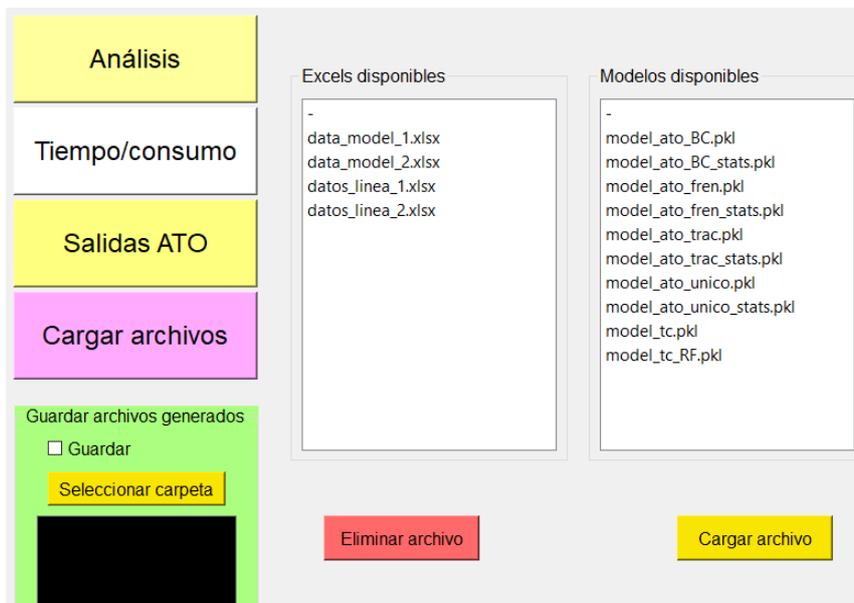


Figura 20: Apartado de carga de archivos

### 5.5 Explorador de archivos

El explorador se utiliza para poder buscar los archivos o carpetas más cómodamente. Se utiliza en el apartado de tiempo y consumo, para elegir el Excel sobre el que hacer la optimización, y en los apartados de guardado y carga de modelos o Excels.

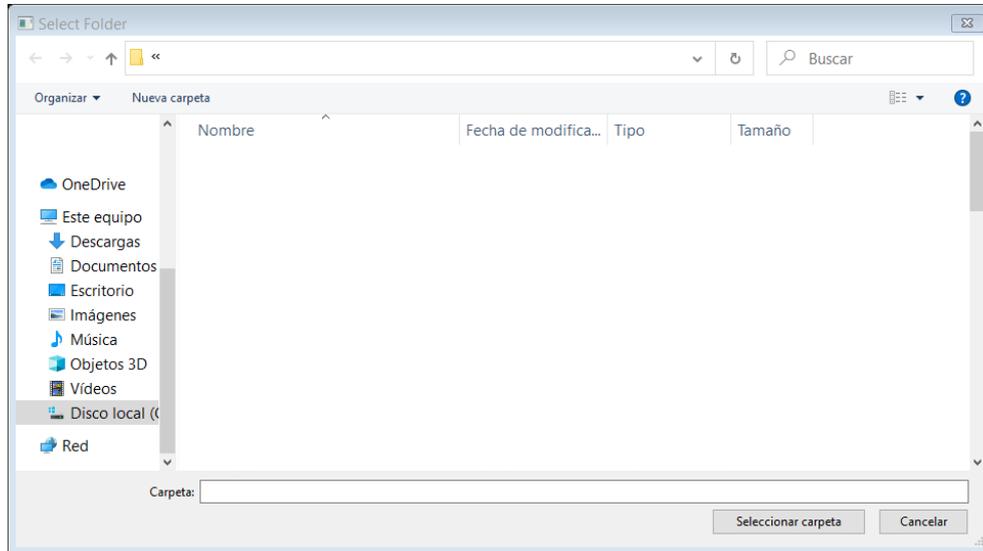
El explorador integrado es el proporcionado por PyQt. El proporcionado por defecto es un explorador de archivos como los que utiliza Windows. En la Figura 21 se incluye el explorador utilizado para seleccionar la carpeta en la que guardar los ficheros.

El explorador es ligeramente diferente cuando se quieren seleccionar archivos.

Desarrollo de una herramienta de apoyo al diseño y optimización de conducción automática de trenes en la que se integren modelos de machine learning

TRABAJO FIN DE MÁSTER

---



*Figura 21: Explorador de archivos*

## 6. Validación de los requisitos y pruebas de funcionamiento

A continuación se incluyen los requisitos y las formas en las que se evaluó si se habían logrado. En los casos en los que la validación se hace de la misma forma se han agrupado los requisitos.

|   |     |   |
|---|-----|---|
| 1 | TC  | Mostrar de forma individual la predicción de tiempo y consumo |
| 5 | ATO | Mostrar de forma individual la predicción de salida ATO       |

Para estos apartados, la prueba consistió en ver si con el modelo por defecto se mostraba el resultado por pantalla. En la Figura 22 se incluye la predicción para el caso de tiempo y consumo. Los valores reales que tenía que haber predicho son 103.3s y 5.64 kWh.

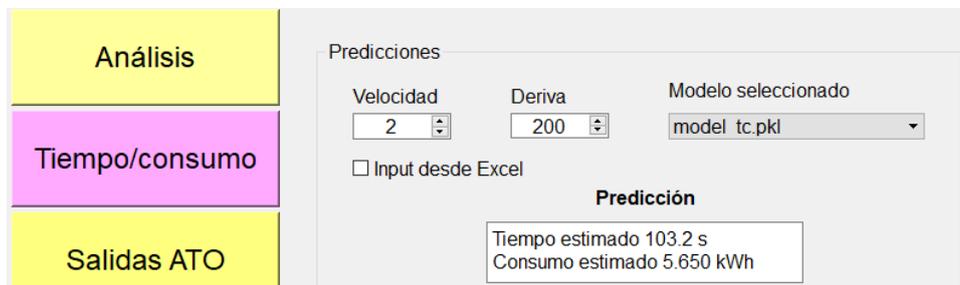


Figura 22: Predicción individual de tiempo y consumo

La predicción de la salida del ATO se hace desde la pestaña de *Cálculo sin simulación* (apartado 5.3.2). En la Figura 13 se incluyen un ejemplo de esta predicción, el valor real correspondiente es de 3443 [10000·m/s<sup>2</sup>].

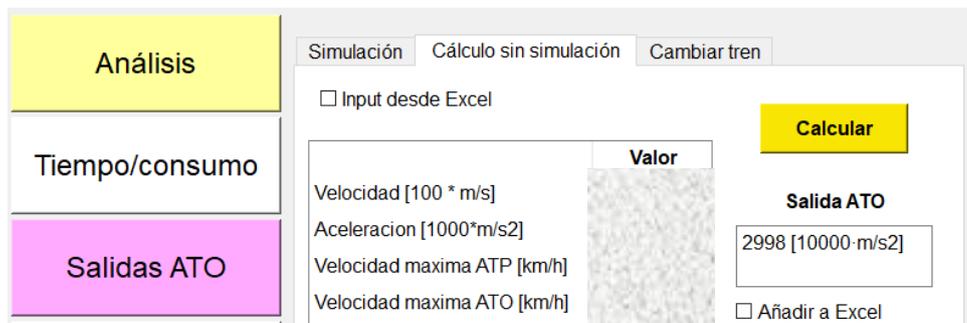


Figura 23: Predicción individual de la salida del ATO

|   |     |  |
|---|-----|--|
| 2 | TC  | Introducir datos a través de fichero Excel |
| 6 | ATO | Introducir datos a través de fichero Excel |

Predicción desde un fichero de prueba, en el que se incluyeron 6 puntos aleatorios. En la Figura 24 se puede ver el Excel de entradas para el caso de tiempo y consumo, y el aspecto tras la predicción con la herramienta. Para el ATO el comportamiento es similar, las predicciones se añaden como una columna extra, y se cambia el formato para resaltar en otro color las salidas.

| Velocidad | Deriva | Velocidad | Deriva | Tiempo   | Consumo  |
|-----------|--------|-----------|--------|----------|----------|
| 2         | 400    | 2         | 400    | 103.1316 | 6.00426  |
| 3         | 200    | 3         | 200    | 104.8116 | 5.452261 |
| 4         | 500    | 4         | 500    | 106.4021 | 5.535566 |
| 2         | 100    | 2         | 100    | 103.4676 | 5.167897 |
| 1         | 200    | 1         | 200    | 101.5039 | 5.836463 |
| 0         | 250    | 0         | 250    | 100.056  | 6.19642  |

Figura 24: Tablas del Excel de entradas para predecir tiempos y consumos

|   |     |   |
|---|-----|---|
| 3 | TC  | Incorporar las predicciones realizadas y entradas a Excel de resultados |
| 7 | ATO | Incorporar las predicciones realizadas y entradas a Excel de resultados |

Para validar este apartado, se hicieron varias predicciones sucesivas con el tick de *Añadir a Excel* activado, comprobando después que en el fichero de resultados aparecían las filas correspondientes. Los valores de las entradas utilizados se cogieron al azar, sin mirar si los puntos concretos existían. La tabla obtenida es igual que la de la Figura 24 en cuanto a formato. Cada vez que se pulsa el botón de *Calcular* se añade una nueva fila al Excel debajo de la anterior.

Adicionalmente se probó la opción de borrar el Excel (*Clear Excel*), eliminando todas las filas y dejando únicamente la cabecera. En este caso, el Excel limpio es la cabecera de la Figura 24, o con las columnas correspondientes al caso de predicción de salidas del ATO. Predicciones sucesivas tras limpiar el Excel se añadieron sin problema como nuevas filas.

|   |     |   |
|---|-----|---|
| 4 | TC  | Cambiar el modelo para predicción utilizado |
| 8 | ATO | Cambiar el modelo para predicción utilizado |

Predicción consecutiva utilizando dos modelos diferentes. Para poder probar este apartado se incluyó en la herramienta un modelo alternativo, pudiendo verse en la Figura 25 los resultados obtenidos para el caso de tiempo y consumo. En este caso el resultado en sí no era importante, por lo que no se compararon en ningún momento los valores predichos con los reales. La prueba en este caso consistió en ver que con un mismo valor de las entradas, las salidas eran diferentes.

Para el requisito 8 se hicieron predicciones con los 3 tipos de modelos que se pueden cargar. El de un modelo único más otro de curva de freno predijo las salidas del modelo único para las entradas introducidas. Sólo en el caso de que no frene el principal se utiliza el de curva de freno. La comprobación de que leía bien ambos modelos y cambiaba de uno a otro tipo se hizo a través del simulador. Simulaciones sucesivas de la misma interestación, usando el modelo único, y el único más curva de freno, dieron resultados distintos, confirmando que el cambio de uno a otro se realizaba correctamente.

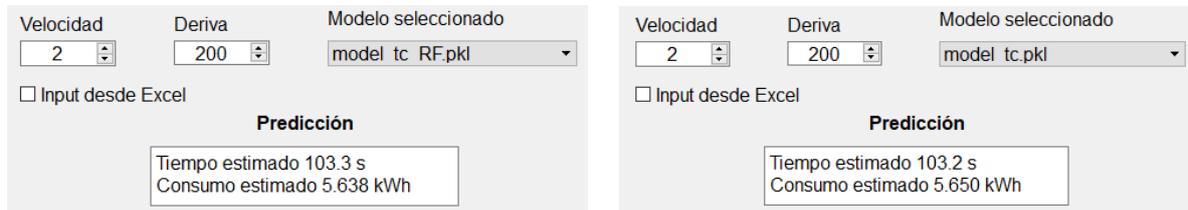


Figura 25: Predicción consecutiva con dos modelos diferentes

|   |     |  |
|---|-----|--|
| 9 | Sim | Cambiar los datos del tren (masa, coefs A B C) |
|---|-----|--|

Cambiar estos parámetros afecta únicamente a la simulación. Las pruebas en este caso consistieron en cambiar los valores por defecto y ver si al abrir de nuevo la herramienta se habían guardado los nuevos valores. También se comprobó que las simulaciones seguían funcionando al cambiar los valores, y que se podía volver a los cargados por defecto.

|    |     |                                |
|----|-----|--------------------------------|
| 10 | Sim | Gráficas básicas de resultados |
|----|-----|--------------------------------|

En este caso la comprobación consistió en la representación de todos los gráficos obtenibles al simular una interestación. En la Figura 26 se puede ver uno de ellos, el denominado '4 en 1', que agrupa los demás. Los gráficos individuales son los mismos, aunque más grandes y con título y leyenda únicos.

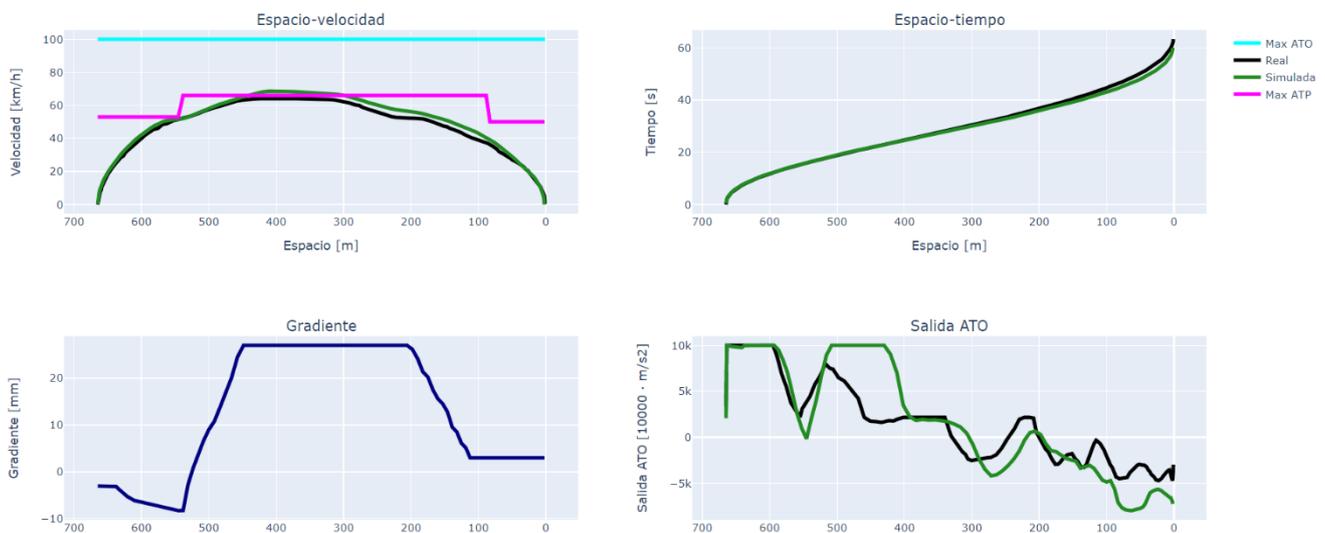


Figura 26: Gráfico combinado de una simulación

|    |     |  |
|----|-----|--|
| 11 | Sim | Sacar Excel con resultado de la simulación |
|----|-----|--|

Dentro de este requisito se distinguen dos partes, como se mencionó en los puntos 4 y 5.3., la validación fue acorde a estas dos posibilidades:

- Obtención de un Excel de todas las salidas
- Obtención de un Excel con las salidas marcadas con ticks en la lista desplegable

Para las pruebas se mantuvo activado el tick de *Obtener Excel*, y consistieron en:

- Obtener un Excel con todas las salidas sin haber marcado ninguna opción en el desplegable, ya que por defecto se obtienen todas las salidas posibles
- Obtener un Excel con las opciones marcadas
- Cuando se activa la opción *Todos*, obtener el Excel completo, aunque haya marcadas otras casillas

Independientemente de la hipótesis probada, en los Excels siempre tiene que aparecer a la izquierda entre qué estaciones es la simulación. En la Figura 27 se incluye un fragmento de una tabla obtenida en una simulación. En las columnas de la izquierda se ven la estación de origen (Present Station Code) y la estación de destino (Next Station Code).

| PSC | NSC | Tiempo [s] | DistEst [m] | v [km/h] | a [m/s <sup>2</sup> ] | Salida ATO [1] |
|-----|-----|------------|-------------|----------|-----------------------|----------------|
| 1   | 2   | 0.5        | 665         | 0.4      | 0.2                   | 2083.94        |
| 1   | 2   | 1          | 665         | 1.4      | 0.6                   | 5542.43        |
| 1   | 2   | 1.5        | 664         | 3.1      | 0.9                   | 9126.99        |
| 1   | 2   | 2          | 664         | 4.9      | 1.0                   | 10000          |
| 1   | 2   | 2.5        | 663         | 6.7      | 1.0                   | 10000          |

Figura 27: Tabla obtenida de resultados de simulación

En la Figura 28 se incluye de nuevo el desplegable para la selección de las salidas a incluir en los resultados.

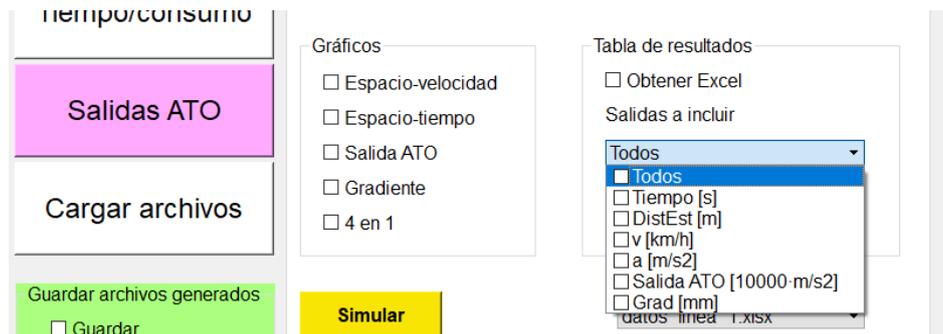


Figura 28: Categorías a incluir en el Excel de resultados de la simulación

|    |     |  |
|----|-----|--|
| 12 | Sim | Simulación de interestación, con o sin datos reales de circulación |
|----|-----|--|

De acuerdo con las dos posibilidades que hay (apartado 5.3.1), se realizaron las siguientes comprobaciones:

- Simulación de una interestación aislada, con datos únicamente de la línea. En este caso, las gráficas que se representan tienen una sola línea, la de resultado de la simulación.
- Si se compara la simulación con un caso real, en los gráficos tienen que aparecer ambas líneas, la de los datos reales y la de las soluciones de la simulación. En la Figura 29 se puede ver una comparación entre la gráfica que se representa si hay datos reales o si no.

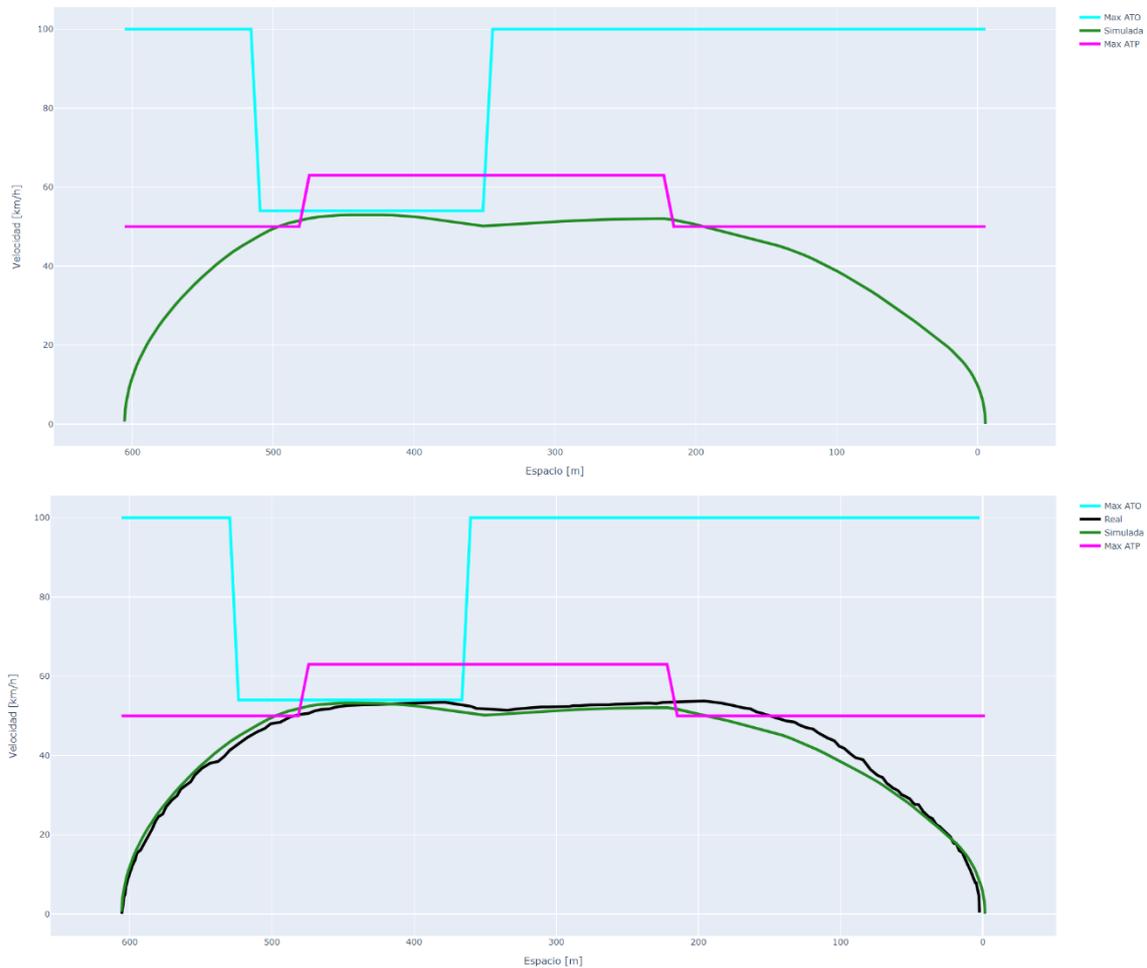


Figura 29: Comparación entre simular teniendo datos reales o solo los datos de la línea

|    |    |   |
|----|----|---|
| 13 | An | Elegir las categorías a incluir en el gráfico |
|----|----|---|

Se comprobó representando dos gráficos seguidos con diferentes ejes. En la Figura 30 se incluye un ejemplo del desplegable en el que se pueden seleccionar las categorías para cada eje. En las pruebas se representaron gráficos en 2D y en 3D, para ver que en ambos se podían cambiar los ejes, y que se podía pasar de uno a otro tipo correctamente.

Aunque no es un requisito explícito, al seleccionar un Excel y una hoja dentro de ese Excel, en estos menús desplegables se tienen que actualizar los nombres que aparecen, para ser las columnas del Excel.

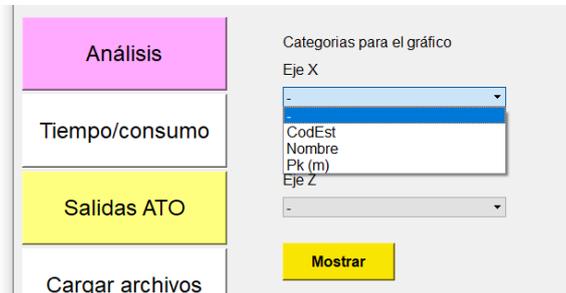


Figura 30: Ejemplo del desplegable para seleccionar los ejes del gráfico

|    |     |                                       |
|----|-----|---------------------------------------|
| 14 | Sim | Cambiar el Excel de datos de la línea |
|----|-----|---------------------------------------|

Para comprobar este requisito se tenían cargados en la herramienta dos Exceles diferentes, y se hicieron simulaciones consecutivas con ellos. Dado que lo que se quería era comprobar que leía bien el nuevo fichero, sólo se representaron las gráficas de espacio-velocidad para ver que eran las de las interestaciones simuladas.

|    |    |                           |
|----|----|---------------------------|
| 15 | An | Cambiar el Excel de datos |
|----|----|---------------------------|

En este caso, al igual que en el requisito anterior, se representaron gráficas de forma consecutiva con ficheros Excel diferentes. Las categorías seleccionadas para los ejes fueron las mismas para los dos Exceles, comparando las gráficas para ver que eran diferentes y que los puntos representados encajaban con los rangos de valores de los Exceles de datos.

|    |      |  |
|----|------|--|
| 16 | Misc | Seleccionar donde guardar los Exceles/gráficos |
|----|------|--|

La solución para realizar el guardado se presentó en el apartado 5.4. Las pruebas se llevaron a cabo con la pestaña de simulación, para comprobar también que al activar la opción de no guardar los gráficos funcionaba.

Se realizaron 2 tipos de pruebas, en las que se mantuvo activado el tick de guardar:

- Que al no seleccionar ninguna carpeta se guarde en el escritorio
- Que al seleccionar una carpeta o escritorio se guarde en ella

Para ambas opciones se corrieron simulaciones para comprobar que guardaba los archivos bien, tanto las gráficas como el Excel de resultados. En paralelo se comprobó también que funcionaba la opción de no guardar las gráficas. Cuando está seleccionada, las gráficas únicamente se tienen que abrir en el explorador. Si no se selecciona, las gráficas se tienen que guardar junto al resto de archivos en la carpeta indicada. El guardado se realiza en formato *.html*, para que las figuras mantengan las opciones de Plotly, como la de rotar los gráficos 3D.

|    |    |                               |
|----|----|-------------------------------|
| 17 | TC | Mostrar el cluster de óptimas |
|----|----|-------------------------------|

Las pruebas realizadas dependen de desde dónde se haga la optimización. Como se comentó en el apartado 5.2, se pueden hacer las predicciones desde un Excel y todo seguido la optimización. O existe también la posibilidad de hacer la optimización sobre un fichero existente en el que ya estén todos los valores.

❖ Caso 1: predicción + optimización

La validación partió de un fichero de prueba, que es el mismo con el que se validó el requisito 2. Al activar el tick de obtener gráficas y Excel en el apartado de optimización (ver en la Figura 9 la caja denominada 'Pareto') se ha de obtener una gráfica como la presentada en la Figura 10, y obtener un Excel de resultados, en el que haya una columna extra al lado de los resultados, indicando si el punto pertenece al *Pareto front* o no. Adicionalmente se ha de obtener en el mismo Excel una hoja en la que sólo aparezcan los puntos que forman parte del *Pareto front*. En la Figura 31 se pueden ver las tablas del Excel con la columna adicional. La tabla completa se muestra a la izquierda y la hoja que sólo tiene las soluciones óptimas a la derecha.

❖ Caso 2: optimización de un fichero independiente

| Velocidad | Deriva | Tiempo   | Consumo  | IsPareto  |
|-----------|--------|----------|----------|-----------|
| 2         | 400    | 103.1316 | 6.00426  | FALSO     |
| 3         | 200    | 104.8116 | 5.452261 | FALSO     |
| 4         | 500    | 106.4021 | 5.535566 | FALSO     |
| 2         | 100    | 103.4676 | 5.167897 | VERDADERO |
| 1         | 200    | 101.5039 | 5.836463 | VERDADERO |
| 0         | 250    | 100.056  | 6.19642  | VERDADERO |

| Velocidad | Deriva | Tiempo   | Consumo  | IsPareto  |
|-----------|--------|----------|----------|-----------|
| 2         | 100    | 103.4676 | 5.167897 | VERDADERO |
| 1         | 200    | 101.5039 | 5.836463 | VERDADERO |
| 0         | 250    | 100.056  | 6.19642  | VERDADERO |

Figura 31: Excels obtenidos de la optimización

Para este caso se borró la columna de *IsPareto* del Excel de la Figura 31, y la hoja en la que están las soluciones óptimas segregadas. Después se seleccionó este fichero en la parte inferior, solo de optimización, que se puede ver en la Figura 9. Como en el apartado anterior también se eligió que hiciera la gráfica y el Excel, obteniéndose los mismos resultados que en el Caso 1.

|    |     |  |
|----|-----|--|
| 18 | Sim | Elegir las salidas a incluir en el gráfico |
|----|-----|--|

Este requisito buscaba complementar las gráficas básicas que se ofrecen en la herramienta. Dado que se dispone de un apartado de representación, se comprobó que desde este apartado de *Análisis* se podía acceder a los Excels de resultados de las simulaciones generados. En la Figura 32 se pueden ver las categorías del desplegable al seleccionar el Excel de resultados de la simulación.

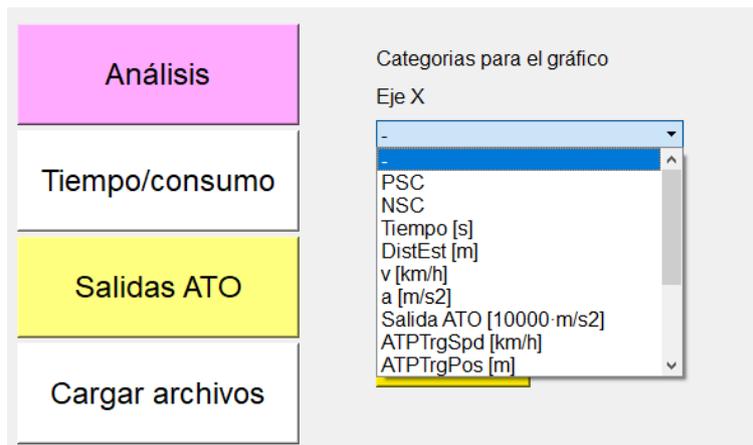


Figura 32: Uso del apartado de 'Análisis' para representar los resultados de la simulación

|    |     |   |
|----|-----|---|
| 19 | Sim | Simulación de varias interestaciones seguidas y gráficas de todas |
|----|-----|---|

Se comprobó poniendo números no consecutivos a los códigos de las estaciones en el apartado de simulación. Como abre gráficas para cada interestación, se seleccionó que sólo representara la gráfica de espacio-velocidad. Para los Excels de resultados no se seleccionó ninguna opción

(obteniéndose por defecto todas las categorías). La prueba consistió en la comprobación de que se generaban las gráficas y los Excels de cada una de las interestaciones, con los valores de longitud y velocidades adecuados.

|    |      |                        |
|----|------|------------------------|
| 20 | Misc | Explorador de archivos |
|----|------|------------------------|

Este requisito se comprobó al tiempo que otros que lo utilizan, como la optimización de tiempos y consumos. Se desarrolló a parte antes de integrarlo en la herramienta, y en las pruebas preliminares se probó con un código que imprimía la ruta del archivo seleccionado. Al utilizarlo en los diferentes apartados fue necesario un ajuste extra para adaptarlo a la herramienta.

|    |    |                       |
|----|----|-----------------------|
| 21 | TC | Gráfica de resultados |
|----|----|-----------------------|

Para este requisito también se aprovechó la pestaña de *Análisis*, comprobando que dejaba acceder a los Excels de resultados. En la Figura 33 se muestra una de las gráficas obtenidas. Se utilizó el mismo fichero de prueba que para el requisito 2, obteniéndose únicamente los 6 puntos aleatorios que contiene.

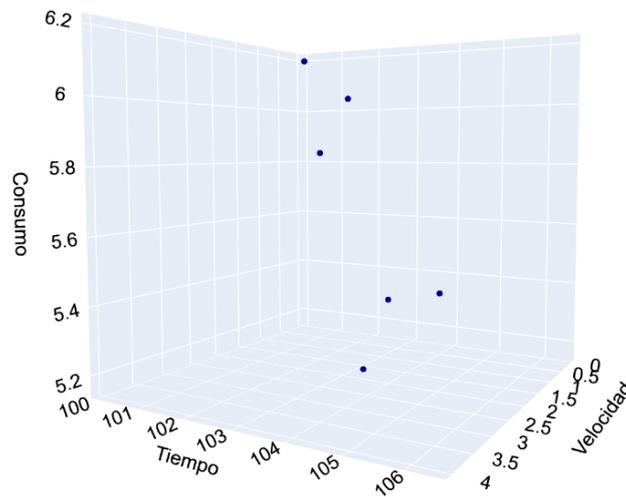


Figura 33: Gráfica de la predicción de tiempo y consumo desde la pestaña de 'Análisis'

## 7. Caso de ejemplo

En este apartado se presenta un caso de ejemplo, en el que se utilizará la herramienta para analizar una marcha, pasando por las etapas que tendría este proceso. Ya que se trata de un caso base, no se utilizarán algunas opciones, como la de guardar, o la de análisis de datos.

Se parte de una serie de posibles marchas. En este caso se utilizará el Excel de las pruebas modificado para contener más puntos. En la Tabla 4 se pueden ver las consignas elegidas para predecir sus tiempos y consumos y seleccionar una.

Tabla 4: Marchas originales en el caso de ejemplo

| Velocidad | Deriva |
|-----------|--------|
| 2         | 400    |
| 3         | 200    |
| 4         | 500    |
| 2         | 100    |
| 1         | 200    |
| 0         | 200    |
| 0         | 300    |
| 0         | 400    |
| 0         | 500    |

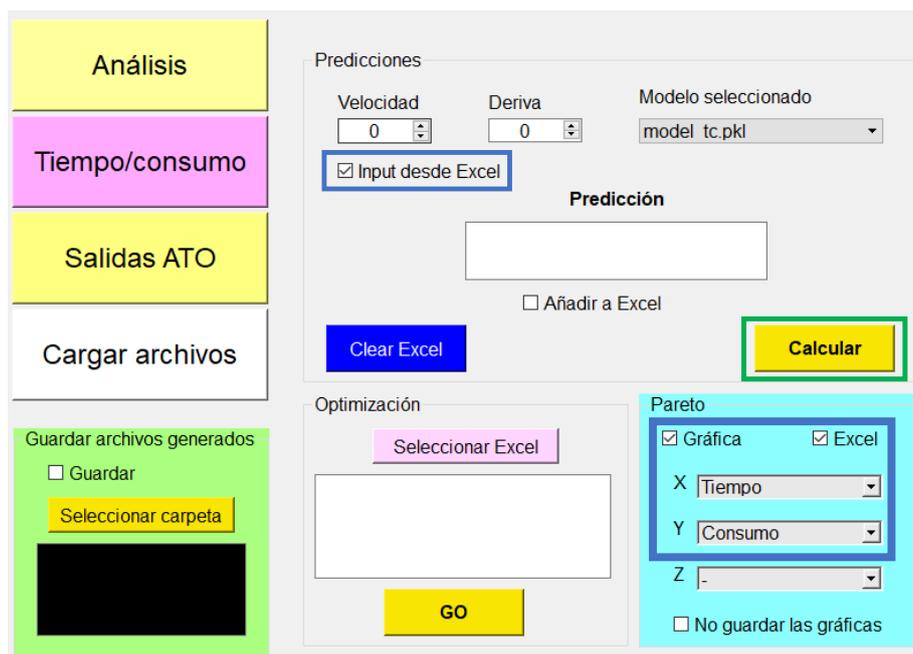


Figura 34: Configuración de la pestaña de 'Tiempo/consumo' para simular el caso de ejemplo

Al tener las entradas en un Excel, en el apartado de *Tiempo/consumo* se selecciona la opción de *Input desde Excel* (Figura 34). Como se quiere que a continuación haga una optimización, se seleccionan también en el cuadro azul las opciones de *Gráfica* y *Excel*. Para los ejes del gráfico se especifican el tiempo y el consumo.

Al hacer click en *Calcular*, se obtiene una gráfica mostrada en la Figura 35, en la que se resaltan las marchas óptimas. En la Tabla 5 se puede ver como se modifica el Excel de entradas.

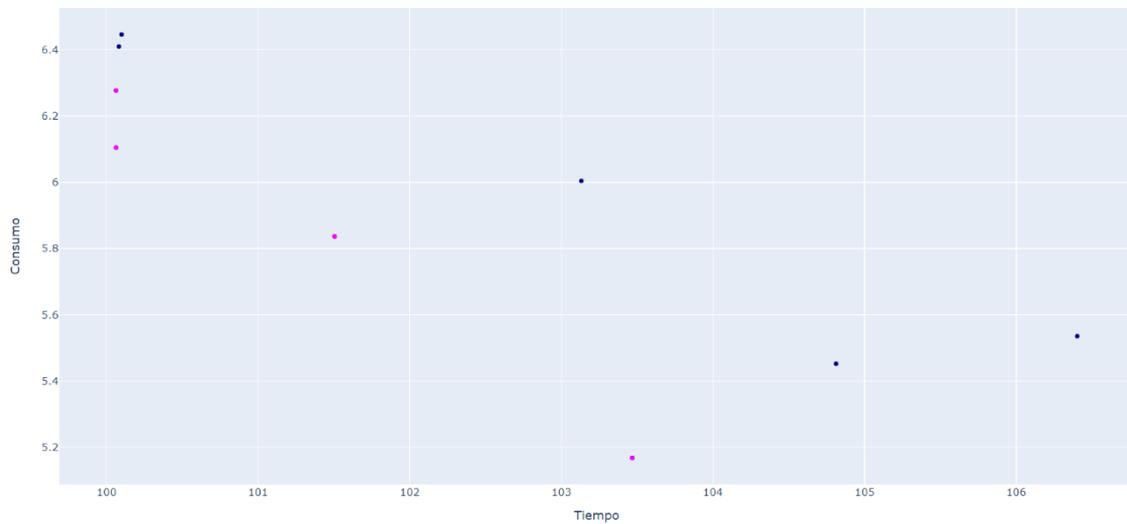


Figura 35: Marchas óptimas del caso de ejemplo

Tabla 5: Soluciones a la predicción de tiempos y consumos para el caso de ejemplo

| Velocidad | Deriva | Tiempo   | Consumo  | IsPareto  |
|-----------|--------|----------|----------|-----------|
| 2         | 400    | 103.1316 | 6.00426  | FALSO     |
| 3         | 200    | 104.8116 | 5.452261 | FALSO     |
| 4         | 500    | 106.4021 | 5.535566 | FALSO     |
| 2         | 100    | 103.4676 | 5.167897 | VERDADERO |
| 1         | 200    | 101.5039 | 5.836463 | VERDADERO |
| 0         | 200    | 100.0634 | 6.104643 | VERDADERO |
| 0         | 300    | 100.0626 | 6.277013 | VERDADERO |
| 0         | 400    | 100.0815 | 6.409804 | FALSO     |
| 0         | 500    | 100.0996 | 6.446154 | FALSO     |

La marcha seleccionada para ser estudiada a continuación en el simulador es la de

- ❖ Velocidad = 0
- ❖ Deriva = 200

La velocidad a la que conducirá será la máxima que tenga la interestación, ya que la reducción de velocidad es 0. Como se menciona en el apartado 3.1, este valor indica los km/h por encima o por debajo de la velocidad máxima a los que puede conducir.

También es posible obtener el tiempo y consumo de esta marcha de forma individual, en la Figura 36 se puede ver esta predicción. Se podría haber partido de aquí para estudiar la marcha.

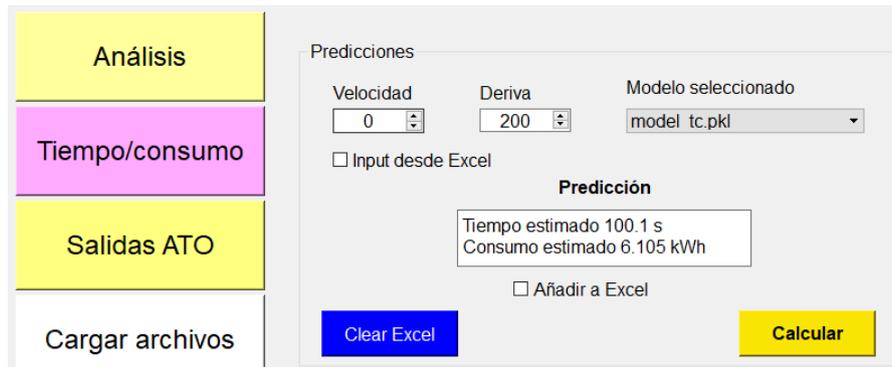


Figura 36: Predicción individual de la marcha del caso de ejemplo

Una vez seleccionada la marcha, se pasa a la simulación. Se simulará una interestación disponible entre las cargadas en la herramienta. Se pasa a la pestaña de *Salidas ATO*, y se rellenan los campos de la simulación (Figura 37). En los *Datos de la línea* se selecciona el Excel que contiene los valores de esta interestación (*datos\_linea\_1*), y en los *Valores reales* el que contiene los valores registrados. En las estaciones de inicio y fin se escriben los códigos utilizados, que en este caso son 1 y 2, y en la distancia de deriva, la distancia a la estación de destino. En el tipo de modelo se ha seleccionado el que tenía menos error de los presentados en la Tabla 1.

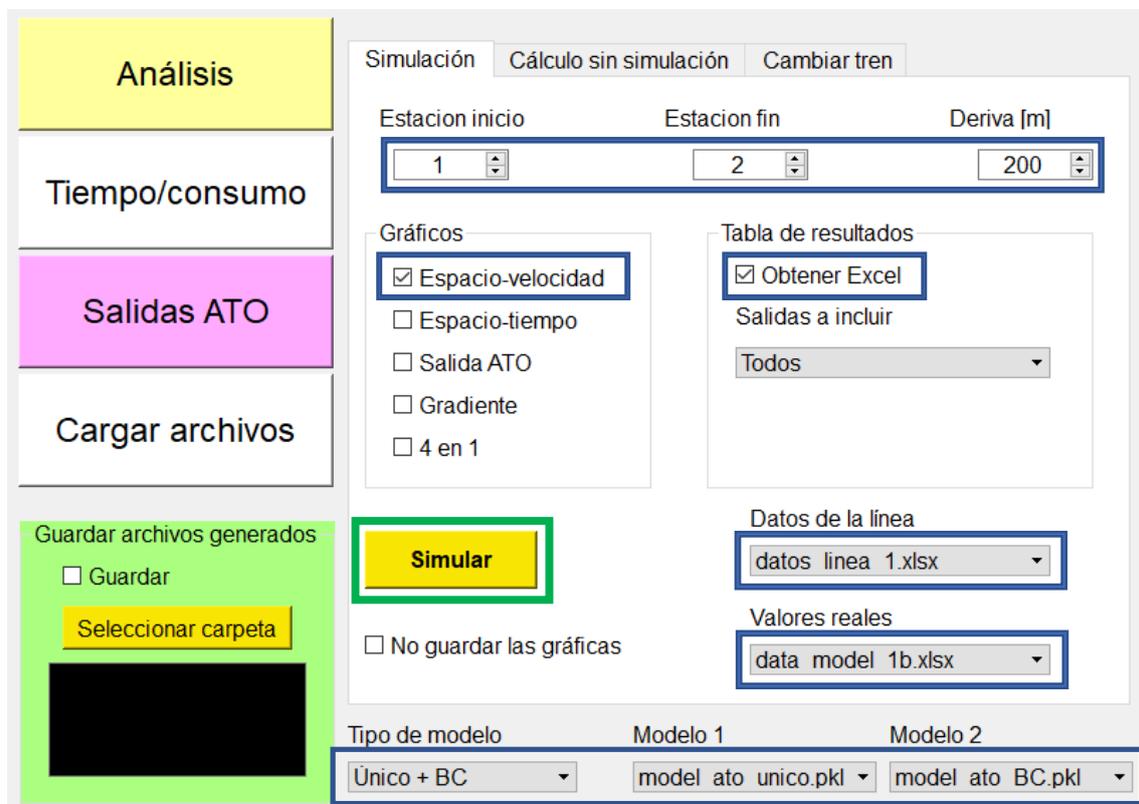


Figura 37: Configuración de la pestaña de 'Salidas ATO' para simular el caso de ejemplo

Sólo se obtendrá una de las gráficas (*Espacio-velocidad*), y un Excel de resultados en el que aparecerán todas las columnas posibles, al no haber seleccionado nada en el desplegable de

*Salidas a incluir.* Al pulsar en *Simular* se genera la gráfica, que se puede ver en la Figura 38, y el Excel de resultados, que se incluye de forma abreviada en la Tabla 6. El tren simulado tarda 97 s, ligeramente inferior al tiempo predicho por el modelo de tiempo y consumo (100.1 s).



Figura 38: Simulación del caso de ejemplo

En la Tabla 6 se ve una distancia a la estación (ParEst) negativa. Esto quiere decir que el tren se pasa 13.2 m del punto de parada.

Tabla 6: Resultados de la simulación del caso de ejemplo

| PSC | NSC | Tiempo [s] | DistEst [m] | v [km/h] | a [m/s <sup>2</sup> ] | Salida ATO [1 |
|-----|-----|------------|-------------|----------|-----------------------|---------------|
| 1   | 2   | 0.5        | 1157.02     | 0.70     | 0.39                  | 4035.82       |
| 1   | 2   | 1          | 1156.83     | 2.06     | 0.76                  | 7718.91       |
| 1   | 2   | 1.5        | 1156.42     | 3.83     | 0.98                  | 10000         |
| 1   | 2   | 2          | 1155.77     | 5.60     | 0.98                  | 10000         |
| 1   | 2   | 2.5        | 1154.87     | 7.36     | 0.98                  | 10000         |
| ... |     |            |             |          |                       |               |
| 1   | 2   | 95.5       | -12.69      | 2.93     | -0.60                 | -6165.58      |
| 1   | 2   | 96         | -13.02      | 1.85     | -0.60                 | -6261.45      |
| 1   | 2   | 96.5       | -13.20      | 0.75     | -0.61                 | -6341.15      |
| 1   | 2   | 97         | -13.24      | 0        | -0.62                 | -6394.09      |

Esta variación en el tiempo observado conduciría también a un consumo ligeramente diferente. En la Tabla 7 se muestra una comparación de los resultados obtenidos. Se ha incluido el tiempo y consumo que tiene la marcha original. Para los casos siguientes se muestran los valores obtenidos junto con el error esperable que tiene cada modelo. Adicionalmente se ha incluido el resultado de la simulación si se utiliza el modelo único. En este caso el tiempo que tarda es más próximo al real, en cambio se desvía más del punto de parada.

*Tabla 7: Resumen de resultados del caso de ejemplo*

|                                  | <b>Tiempo [s]</b> | <b>Consumo [kWh]</b> | <b>Distancia al punto de parada [m]</b> |
|----------------------------------|-------------------|----------------------|---|
| Marcha original                  | 100.1             | 6.084                | -                                       |
| Predicción modelo tiempo/consumo | 100.1±0.04        | 6.105±0.02           | -                                       |
| Simulaciones de la interestación |                   |                      |   |
| Simulación con modelo único + BC | 97±0.1            | -                    | 13.2±0.1                                |
| Simulación con modelo único      | 102.5±4.8         | -                    | 60.7                                    |

## 8. Conclusiones y aportaciones

La herramienta desarrollada se puede utilizar para hacer predicciones: de tiempos de recorrido y consumos de un tren equipado con ATO, y de las salidas que daría el equipo de ATO durante el recorrido. Estas predicciones con técnicas de machine learning tienen la ventaja frente a otros modelos y simuladores de que son más rápidas. Las predicciones de tiempos y consumos, sumado a la optimización, permite encontrar el espacio de soluciones más eficientes a la hora de diseñar las marchas de un equipo ATO.

Las simulaciones de marcha realizadas necesitan unos 15 segundos por interestación, frente a los 3.5 segundos que tarda el primer modelo en predecir 10000 posibles marchas. La simulación supone un tiempo considerablemente mayor que la predicción del tiempo y consumo. Con este primer modelo se pueden evaluar y seleccionar las mejores marchas. Estas soluciones se pueden estudiar posteriormente con más detalle en el apartado de simulación.

El modelo de ATO que integra el simulador de marcha es una red neuronal. Ha sido entrenado con datos de registros para que aprendiera a comportarse como un ATO. Visto como caja negra, ambos funcionan igual. Sin embargo, internamente son totalmente diferentes. La red neuronal no conoce el funcionamiento interno del ATO, sólo la relación entre entradas y salidas. Esto la convierte en un modelo de ATO independiente de fabricantes, pudiendo usarse para replicar el comportamiento de un ATO en cualquier situación.

Tener ambas partes integradas en una misma herramienta supone que se puede pasar de una a otra de forma consecutiva. De la optimización de marchas se pueden obtener las más eficientes, y con esas pasar a la pestaña de simular y seguir estudiándolas.

Desde el punto de vista del espacio ocupado en disco, también es más ventajoso que las dos partes estén juntas en una misma herramienta. El *.exe* se genera para que la aplicación se pueda usar en cualquier ordenador. Esto supone que ha de tener en su carpeta todos los paquetes necesarios. Aunque algunos son específicos para cada parte, muchos son comunes. Separar la evaluación de marchas de la simulación supondría tener que duplicar estos paquetes, ocupando casi el doble.

El apartado adicional de *Análisis* complementa a los otros dos, permitiendo obtener más gráficas que las presentes por defecto. Sin embargo, hay que tener en cuenta que no se puede utilizar en sustitución de la parte correspondiente en los otros apartados. Las gráficas con la pestaña de *Análisis* tienen un único color, y sólo permiten representar una categoría por eje. Esto excluye las gráficas que muestran en colores diferentes los óptimos de tiempo y consumo, y las de simulación en las que hay varias velocidades o tiempos en un mismo eje. Estos gráficos sólo se pueden hacer desde la pestaña a la que pertenecen.

Este apartado busca una mayor capacidad de visualización de los datos. Teniendo en cuenta las dos condiciones mencionadas en el párrafo anterior, es posible obtener gráficos de cualquier Excel. Como se comentó en el apartado 5.1, se puede elegir la hoja del Excel que se quiere utilizar para la representación.

La herramienta, en todos los apartados, presenta la posibilidad de intercambiar los ficheros utilizados. Para cambiar de archivo no es necesario cerrar la herramienta y sustituir el fichero con otro del mismo nombre. Al haber un desplegable con una lista de Excels, o modelos, se puede elegir el que usar en cada momento, agilizando y flexibilizando la operación. Cambiar los Excels utilizados supone una mayor capacidad de análisis y simulación. En esta simulación, además de poder cambiar la línea o interestación, se pueden cambiar las características del tren, buscando que se acerque lo más posible al caso que se quiere estudiar.

El poder cambiar los modelos de predicción deja la herramienta abierta a mejoras futuras de los mismos, si se consiguen otros con menos error. Esta característica permite también evaluar el comportamiento de modelos diferentes a la hora de simular.

## 9. Bibliografía

- [1] D-RAIL, «D-RAIL - Predictive railway maintenance - Real time infrastructure monitoring,» 2021. [En línea]. Available: <https://www.d-rail.com/>.
- [2] Thales, «Artificial intelligence for the autonomous trams,» 2019. [En línea]. Available: <https://www.thalesgroup.com/en/germany/news/artificial-intelligence-autonomous-trams>.
- [3] Rail Vision, SBB, «Rail Vision Shunting Yard Solution,» <https://vimeo.com/294567425>, <https://railvision.io/>, 2019.
- [4] UIC - Rail System Department, «Artificial intelligence - Case of the railway sector - State of play and perspectives,» Marzo 2021.
- [5] M. Niestadt, A. Debyser, D. Scordamaglia y M. Pape, «Artificial intelligence in transport - Current and future developments, opportunities and challenges,» *EPRS - European Parliamentary Research Service*, 3 2019.
- [6] Comisión Europea, «Annexes to the Communication: Fostering a European approach to Artificial Intelligence (Coordinated Plan on Artificial Intelligence 2021 Review),» 2021.
- [7] Gobierno de España. MINECO, «ENIA - Estrategia Nacional de Inteligencia Artificial,» 2020.
- [8] UITP, «World report on metro automation,» 2018.
- [9] European Union Agency for Railways, «Single Programming Document 2021-2023,» 25-11-2020.
- [10] Shift2Rail, «Annual Work Plan and Budget 2021 - Annex to GB decision no 01/2021,» 12-05-2021.
- [11] P. Cucala y A. Fernández, «Energy efficiency in high-speed train operation,» *Global Railway Review*, 4 Agosto 2010.
- [12] UNIFE, «UNIFE Vision Paper on Digitalisation: Digital Trends in the Rail Sector,» 2019.
- [13] UNIFE, «UNIFE Vision Paper on Digitalisation - Second Chapter: Rail fit for digital age,» 2020.
- [14] M. Casteel, «Simulation of train driver behaviour using a data-driven agent,» 2016.
- [15] Z. Wang, X. Liao, C. Wang, D. Oswald, G. Wu, K. Boriboonsomsin, M. Barth, K. Han, B. Kim y P. Tiwari, «Driver Behavior Modeling Using Game Engine and Real Vehicle: A

Learning-Based Approach,» *IEEE Transactions on Intelligent Vehicles*, vol. 5, pp. 738-749, 11 2020.

[16] P. Martínez, «Estudio y optimización del consumo energético del ferrocarril mediante redes neuronales y algoritmos heurísticos,» 2019.

[17] L. Breiman, «Random Forests,» *Machine Learning*, vol. 45, pp. 5-32, 10 2001.

[18] S. Rodríguez, *Modelado de la marcha y del consumo de trenes metropolitanos de conducción automática mediante técnicas de machine learning*, TFM 2º MII - Universidad Pontificia Comillas, 2021.

[19] Qt, *Qt Group*, <https://www.qt.io/>.

[20] Plotly, «Plotly Technologies Inc. Collaborative data science,» 2015. [En línea]. Available: <https://plot.ly>.

[21] European Railway Agency - Arcadis, «Feasibility study reference system ERTMS - Digitalisation of CCS and migration to ERTMS,» 2018.



