



UNIVERSIDAD DE BUENOS AIRES  
FACULTAD DE CIENCIAS EXACTAS Y NATURALES  
DEPARTAMENTO DE COMPUTACIÓN

**Optimización de Estrategias de  
Comercialización en el Mercado de  
Energía Renovable de California  
mediante Programación Lineal Entera y  
Análisis Predictivo**

Tesis de Licenciatura en Ciencias de la Computación

Facundo Nicolás Decroix

Director: Federico Pousa

Codirector: Agustín Mosteiro

Buenos Aires, 2023

## RESUMEN

La transición hacia las energías renovables no es solo una tendencia, sino una necesidad imperante en el contexto actual. Ante los desafíos medioambientales crecientes y la urgencia por mitigar el cambio climático, las energías limpias emergen como la respuesta sostenible y económicamente viable que el mundo necesita. Dentro de este escenario global, California se destaca como un referente de adaptación y progreso, teniendo un mercado de energía estrictamente regulado por su operador *California Independent System Operator*.

En este trabajo nos centramos en la energía solar, utilizando como caso de estudio el de una empresa que participa actualmente del mercado de energía de California. Esta empresa cuenta con paneles solares que obtienen la energía y baterías donde la almacenan para ser vendida posteriormente.

Nuestro objetivo es adentrarnos en esta problemática y diseñar estrategias para poder vender esta energía maximizando la ganancia en el mercado utilizando técnicas de programación lineal entera.

La primera etapa de este trabajo, descrita en el capítulo 1, consistió en entender el funcionamiento del mercado de energía en California, brindando una panorámica detallada de su estructura y del modo en que las diferentes entidades interactúan.

En el capítulo 2, nos centramos en la recolección y análisis de datos. Recolectamos datos históricos sobre precios y los correlacionamos con factores climáticos como temperatura, nubosidad y radiación solar. Adicionalmente, diseñamos un *dashboard* interactivo que permite una visualización clara y sencilla de la información recabada.

El capítulo 3 nos lleva al desarrollo de un predictor de precios utilizando modelos de series de tiempo, basados en la información previamente reunida. También llevamos a cabo una experimentación donde pusimos a prueba este modelo contra otras alternativas, analizando su precisión y robustez.

El núcleo de nuestra investigación se encuentra en el capítulo 4. Allí describimos cómo, utilizando los precios pronosticados, creamos un modelo de programación lineal entera con el objetivo principal de maximizar las ganancias en el mercado de energía, respetando todas las restricciones del mismo.

En el capítulo 5, desarrollamos un sistema de decisiones completo. Utilizando las predicciones realizadas, planteamos distintas estrategias de programación lineal, evaluándolas para elegir la solución más eficiente.

Para finalizar, comparamos nuestro sistema propuesto con las estrategias de venta actuales de la empresa bajo estudio. Los resultados obtenidos evidencian que nuestra propuesta no solo es viable, sino que supera en rentabilidad a los enfoques tradicionales.

En resumen, con este trabajo brindamos un camino y herramientas concretas para las empresas en el mercado de energía renovable de California, apuntando a una gestión más eficiente y productiva en este sector crucial y en constante evolución.

## ABSTRACT

The transition to renewable energies is not just a trend, but an urgent necessity in today's context. Faced with growing environmental challenges and the urgency to mitigate climate change, clean energies emerge as the sustainable and economically viable answer the world needs. Within this global scenario, California stands out as a beacon of adaptation and progress, having an energy market strictly regulated by the California Independent System Operator.

We delve into this intricate landscape with a specific focus: analyzing a company that actively participates in the California energy market as our case study. Through this lens, we seek to discover and develop more profitable strategies for energy trading.

In Chapter 1, we provide an in-depth examination of the dynamics of the energy market in California, offering a comprehensive overview of its structure and how different entities interact within it.

Chapter 2 focuses on data collection and analysis. We gathered historical data on prices and correlated them with climatic factors such as temperature, cloudiness, and solar radiation. Additionally, we designed an interactive dashboard that allows for a clear and simple visualization of the information collected.

Chapter 3 introduces the development of a price predictor based on the previously collected data. Not only did we develop this predictor, but we also tested it against other alternatives, analyzing its accuracy and robustness.

The core of our research is found in Chapter 4. Here, using the forecasted prices, we created a mixed integer programming model with the primary objective of maximizing profits in the energy market, complying with all its restrictions.

In Chapter 5, we developed a comprehensive decision-making system. Using the predictions made, we proposed various integer programming strategies, evaluating them to select the most efficient solution.

To conclude, we compared our proposed system with the current sales strategies of the company under study. The results obtained demonstrate that our proposal is not only viable but also outperforms traditional approaches in profitability.

In summary, with this research, we provide a path and practical tools for companies in the California renewable energy market, aiming for more efficient and productive management in this crucial and continuously evolving sector.

## AGRADECIMIENTOS

Hay mucha gente a la cual debería agradecer, sólo pude alcanzar el logro de terminar esta carrera gracias a la ayuda que tuve de muchas personas a lo largo de estos más de 10 años.

Quiero agradecer a mi familia y amigos, que me sostuvieron desde el primer momento y no dejaron de hacerlo hasta ahora. Hubo muchos momentos donde dudé de mí y gracias a ellos seguí adelante.

También agradecer a Eryx, la cooperativa donde trabajo y en la que siempre tuve puros incentivos a priorizar la carrera y recibirme. Allí se vive un ambiente donde se fomenta mucho la superación de cada uno de sus integrantes y eso fue clave para no dormirme y conseguir el objetivo.

A mi novia Barbi que en estos últimos años me apoyó muchísimo y fue un sostén muy importante

Muchísimas gracias a mis directores Fede y Agus, que durante todo este proceso me acompañaron y ayudaron en todo lo que necesité, dandome confianza y motivación.

Quiero agradecer a la UBA por haberme dado la posibilidad de estudiar una carrera de primer nivel de forma gratuita, cosa que no es común y no debemos dar por sentado.

## Índice general

1.. Introducción al problema . . . . .	1
1.1. Descripción a alto nivel . . . . .	1
1.2. Nuestro contexto . . . . .	2
1.2.1. El mercado de energía . . . . .	2
1.2.2. Nuestro caso de estudio . . . . .	6
1.3. Descripción de alto nivel de nuestro trabajo . . . . .	7
1.3.1. Obtención y visualización de datos . . . . .	7
1.3.2. Predicción de precios . . . . .	8
1.3.3. Solución de programación lineal entera . . . . .	8
1.3.4. Sistema de decisiones y experimentación final . . . . .	8
2.. Origen de datos . . . . .	10
2.1. OASIS . . . . .	10
2.2. Clima . . . . .	12
2.3. Visualización de datos . . . . .	13
2.4. Conclusion del capítulo . . . . .	19
3.. Modelo predictivo . . . . .	20
3.1. Modelos predictivos para series de tiempo . . . . .	20
3.2. Prophet . . . . .	21
3.2.1. Cómo funciona Prophet . . . . .	21
3.2.2. Primeros pasos . . . . .	22
3.3. Buscando el mejor modelo . . . . .	23
3.3.1. ¿Cómo comparar los modelos? . . . . .	23
3.3.2. Instancia de ejecución de la experimentación . . . . .	24
3.3.3. Búsqueda de hiperparámetros . . . . .	24
3.3.4. Cambios en la cantidad de días . . . . .	28
3.3.5. Temperatura . . . . .	29
3.3.6. Agregar feriados . . . . .	30
3.3.7. Estacionalidad horaria . . . . .	30
3.3.8. Modelo final . . . . .	31
3.4. Experimentación y resultados . . . . .	31
3.5. Conclusión . . . . .	34
4.. Modelo de Programación Lineal Entera . . . . .	35
4.1. Introducción sobre Programación Lineal Entera . . . . .	35
4.2. El modelo . . . . .	36
4.2.1. Discretización del tiempo . . . . .	36
4.2.2. Input . . . . .	36
4.2.3. Variables de decisión . . . . .	37
4.2.4. Función objetivo . . . . .	38
4.2.5. Cotas de las variables . . . . .	38
4.2.6. Restricciones . . . . .	38

4.2.7. Modelo final . . . . .	40
4.3. Implementación del modelo . . . . .	41
4.4. Suite de tests para validar el modelo . . . . .	42
4.4.1. Simplificando los escenarios . . . . .	42
4.4.2. Precios siempre son cero excepto a las 18hs y 19hs. . . . .	42
4.4.3. Precios siempre son cero excepto a las 18hs y a las 12hs. . . . .	44
4.4.4. Precios de un día real. . . . .	45
4.5. Cierre del capítulo . . . . .	47
 5.. Integración y Experimentación . . . . .	48
5.1. Objetivo . . . . .	48
5.2. Heurísticas a comparar . . . . .	48
5.3. Simulaciones . . . . .	49
5.4. ¿Cómo obtener simulaciones realistas? . . . . .	51
5.4.1. Modelo de las baterías y paneles solares . . . . .	51
5.5. El riesgo de trabajar con estimaciones . . . . .	52
5.6. Eliriendo el modelo más adaptable a cambios . . . . .	53
5.7. Framework de simulaciones . . . . .	53
5.7.1. ¿Con qué criterio seleccionamos el mejor MIP? . . . . .	54
5.7.2. ¿Cómo hacemos el muestreo? . . . . .	55
5.7.3. ¿Cómo se componen las estimaciones? . . . . .	55
5.7.4. Validando el framework . . . . .	56
5.7.5. ¿Cómo afectan los sampleos a los MIP? . . . . .	56
5.7.6. ¿Cómo afectan los sampleos a las simulaciones? . . . . .	57
5.7.7. ¿Qué vamos a evaluar en la comparación final? . . . . .	58
5.7.8. Ejemplo del framework . . . . .	59
5.8. Primer experimento . . . . .	60
5.9. Experimentación final . . . . .	64
5.9.1. Ejecución de los experimentos . . . . .	64
5.9.2. Adaptando las heurísticas . . . . .	64
5.9.3. Resultados crudos . . . . .	65
5.9.4. Análisis general . . . . .	69
5.9.5. Verano . . . . .	70
5.9.6. Invierno . . . . .	71
5.9.7. Otoño . . . . .	72
5.9.8. Cierre del capítulo . . . . .	73
 6.. Conclusión . . . . .	74
 7.. Trabajo futuro . . . . .	75

# 1. INTRODUCCIÓN AL PROBLEMA

## 1.1. Descripción a alto nivel

Las energías renovables han tomado un lugar central en el escenario energético global, y su relevancia no deja de crecer. En particular, la energía solar se erige como un actor principal en esta transformación. California, conocida por su liderazgo en sostenibilidad, está adoptando la energía solar con determinación debido a su carácter inagotable y su potencial para reducir las emisiones de carbono.

El principal desafío en el ámbito de la energía solar es su intermitencia. Dado que la generación de energía está ligada exclusivamente a la disponibilidad de luz solar, surge una cuestión crucial: ¿cómo afrontar la ausencia de esta última? La cantidad de energía derivada del sol no se encuentra bajo nuestro control, siendo influenciada por una serie de variables tales como la hora del día y el grado de nubosidad. Estos factores inciden directamente en la magnitud de la radiación solar captada en un momento dado, factores que escapan a nuestro control. Por lo que podría pasar que en algún momento necesitemos contar con más energía que la que el sol nos está brindando.

Una solución a esta problemática es usar un sistema de almacenamiento de energía. De esta forma se puede disponer de la misma en momentos en los cuales no se cuenta con luz solar, o se necesita de más energía que la que el sol nos provee.

Durante el día, cuando disponemos de luz solar, almacenamos parte de esta energía para poder contar con ella en momentos donde no contemos con la energía del sol.

Una forma de almacenar energía es mediante el uso de baterías. Utilizando paneles solares obtenemos la energía del sol y usamos parte de esta energía para cargar baterías. Estas baterías retienen la energía para momentos posteriores en los que queramos utilizarla. Este es el tipo de sistema de almacenamiento en el que vamos a centrarnos.

El caso de estudio que vamos a utilizar a lo largo de esta tesis es el de una planta de paneles y baterías ubicado en Lancaster, California. Este sistema se encuentra integrado a la red de energía de California y se dedica a la venta de energía en el mercado.

El objetivo de esta tesis es comprender el funcionamiento del mercado de energía y el funcionamiento de la planta para luego poder elaborar una estrategia de venta de energía que optimice el uso de los recursos disponibles y maximice la ganancia. Con la finalidad de lograr este objetivo, dividimos el trabajo en tres etapas.

La primera será una etapa de análisis descriptivo, trabajaremos en la obtención de información de precios del mercado de energía. También vamos a recopilar datos del tiempo, como la radiación solar, temperatura y nubosidad. Luego de obtener estos datos el objetivo es poder disponibilizar la información a partir de distintos *dashboards*.

La segunda etapa será de análisis predictivo, donde usaremos los datos anteriormente obtenidos para predecir los precios futuros del mercado.

La tercera y final etapa será de análisis prescriptivo, donde utilizaremos técnicas de programación lineal entera para, a partir de los datos provenientes de las anteriores etapas, formular una estrategia que nos permita utilizar el sistema de almacenamiento de energía enunciado de forma tal que maximicemos la ganancia obtenida. Ya han habido otros trabajos que han utilizado programación lineal entera para atacar la problemática de los

mercados de energía que han obtenido buenos resultados.<sup>1</sup>

También vamos a hacer una comparación de nuestra estrategia resultante con estrategias empleadas actualmente en la empresa utilizada como caso de estudio, para ver si efectivamente contamos con una estrategia superadora.

## 1.2. Nuestro contexto

### 1.2.1. El mercado de energía

En este trabajo nos vamos a centrar en un mercado de energía de Estados Unidos. Puntualmente el mercado de energía del estado de California.

Los mercados de compra y venta de energía en Estados Unidos están regulados por entidades independientes llamadas *Independent System Operators* (ISO). Estas entidades son regionales y coordinan, controlan y monitorean el funcionamiento de las distintas redes eléctricas que hay en el país.

Funcionan como un *marketplace* en el mercado mayorista de energía, uniendo a vendedores y compradores. Al poner a todos los compradores y vendedores bajo la misma red, permiten que haya competencia en el mercado y que funcione de forma más centralizada.

Hoy por hoy hay nueve entidades que regulan distintas regiones del mercado de estados unidos.

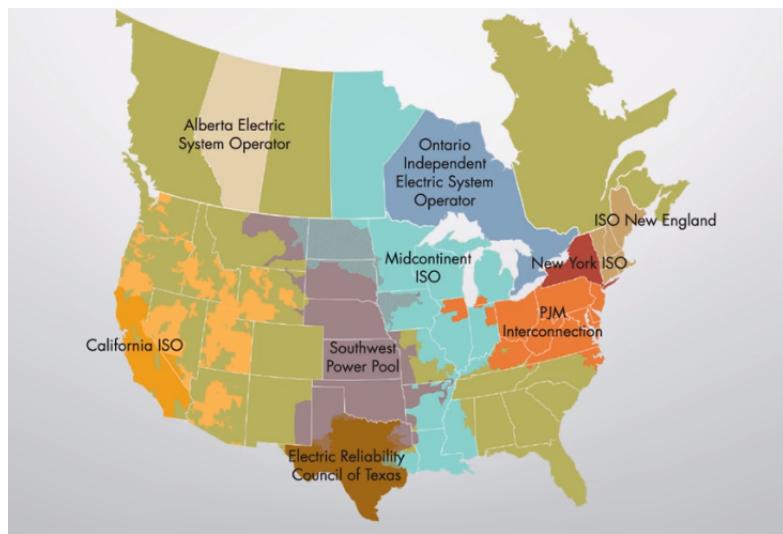


Fig. 1.1: Los distintos ISO de Estados Unidos y Canadá.

Los ISO no poseen ni mantienen activos de transmisión de energía. Solo se dedican a regular el mercado y a asegurarse de que la demanda de energía de la región sea satisfecha, dándole confiabilidad a la red.

Quienes posean generadores de energía y deseen integrarse al sistema deben seguir los procesos del ISO correspondiente a la región en la que se encuentren para poder participar del mercado. A los agentes que están autorizados para participar en el mercado se los conoce como *Scheduling Coordinators* o SC.

---

<sup>1</sup> S. Tong, Handa Yang y W. Torre. En: *Energy Storage System Dispatching Optimization in Stacked Applications for Utility Grid*. 2018. URL: <https://www.semanticscholar.org/paper/Energy-Storage-System-Dispatching-Optimization-in-Tong-Yang/433a90fedbc0a84ad9cb97240d4a08ab8c08b2c1>.

Estos ISO definen las reglas de compra y venta del mercado y quiénes tienen acceso al mismo. Como son independientes, cada ISO tiene sus propias regulaciones. Nosotros puntualmente vamos a centrarnos en el operador perteneciente al estado de California, llamado *California ISO* o CAISO.<sup>2</sup>

En California conviven varios mercados de energía:<sup>3</sup> El *Day-ahead market* o DAM y el *Real-time market* (RTM).

En ambos mercados, la venta de energía se maneja mediante subastas. Los distintos vendedores realizan ofertas de energía donde indican la potencia de energía que van a proveer y el precio que están dispuestos a cobrar por esa energía. Las ofertas son en *megawatt hora* (MWh). El *megawatt hora* es una medida de cantidad de energía, 1 MWh es equivalente a la energía resultante de inyectar energía a una potencia de 1 MW durante toda una hora, equivaldría a 3600M *joules*.

La diferencia entre estos dos mercados es el momento en el cuál se realizan las subastas, cuándo se cierran y el tiempo en el cual uno se compromete a inyectar la energía ofertada en la red.

Un SC que esté integrado al mercado de energía puede participar de cualquiera de estos dos mercados.

En las siguientes secciones detallaremos las particularidades de cada mercado.

#### Day-Ahead Market

Como su nombre lo indica, el *Day-Ahead Market* opera siempre un día por adelantado.

La idea de este mercado es asegurarse de contar con la energía que se va a necesitar con un día de anticipación, basándose en pronósticos de demanda que hace CAISO.

A continuación tenemos una imagen que nos ayudará a ilustrar el proceso de este mercado.

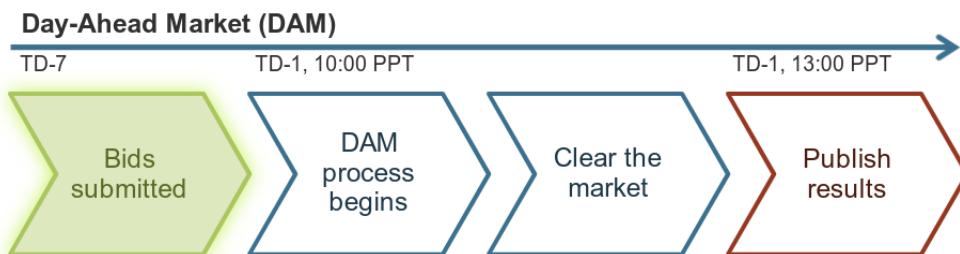


Fig. 1.2: Línea de tiempo del *Day-ahead market*.

El *Day-ahead market* funciona mediante subastas realizadas por los *Scheduling Coordinators*. Estos hacen sus ofertas para hacerle saber al mercado cuánta energía están dispuestos a vender o comprar.

El período para poder hacer estas ofertas es desde siete días antes de la fecha (TD-7) para la cual se está ofertando, hasta el día anterior (TD-1) a las 10 hs.

<sup>2</sup> California ISO. URL: <http://www.caiso.com/Pages/default.aspx>.

<sup>3</sup> California ISO - Markets training. URL: <https://www.caiso.com/participate/Pages/LearningCenter/Markets.aspx>.

En el DAM, el día está dividido en intervalos, cada uno con duración de una hora. Esto significa que en este mercado, las ofertas son para una hora del día en particular. Se pueden hacer ofertas para cada una de las 24 hs. del día.

Las ofertas incluyen una curva que indica un rango de potencias que el SC está dispuesto a vender, o comprar en caso de que sea un comprador. La oferta se realiza en dólares por megawatt por hora. Esto significa que quien realiza una oferta de 10 MWh a 100 \$/MWh, se compromete a brindar 10MW a la red de energía durante esa hora, a cambio de cien dólares por MW.

A las 10 hs. del día anterior a la fecha, se cierra el mercado y no se pueden realizar más ofertas en el DAM. Una vez cerrado el mercado, el ISO realiza varios procesos en los cuales se evalúan las ofertas de compra y venta efectuadas.

Además de las ofertas de venta de energía, CAISO recibe ofertas de compra. También se utiliza información como el mapa de la red de energía, predicciones de clima y faltantes de energía que se hayan informado.

No profundizaremos demasiado en cómo utiliza CAISO esta información, ya que no es relevante para nuestro trabajo. Lo que debemos saber es que, con toda esta información, **CAISO calcula cuál es el precio de venta o *clearing price*** para cada hora del día en el mercado. Esto se realiza, entre otras cosas, viendo cuánta energía se necesita (ofertas de compra) y de cuánta energía dispondrá el mercado (ofertas de venta, predicciones de clima y faltantes de energía). El precio de cierre al cual llega el mercado es único y es el precio que obtendrán todos los participantes del mercado.

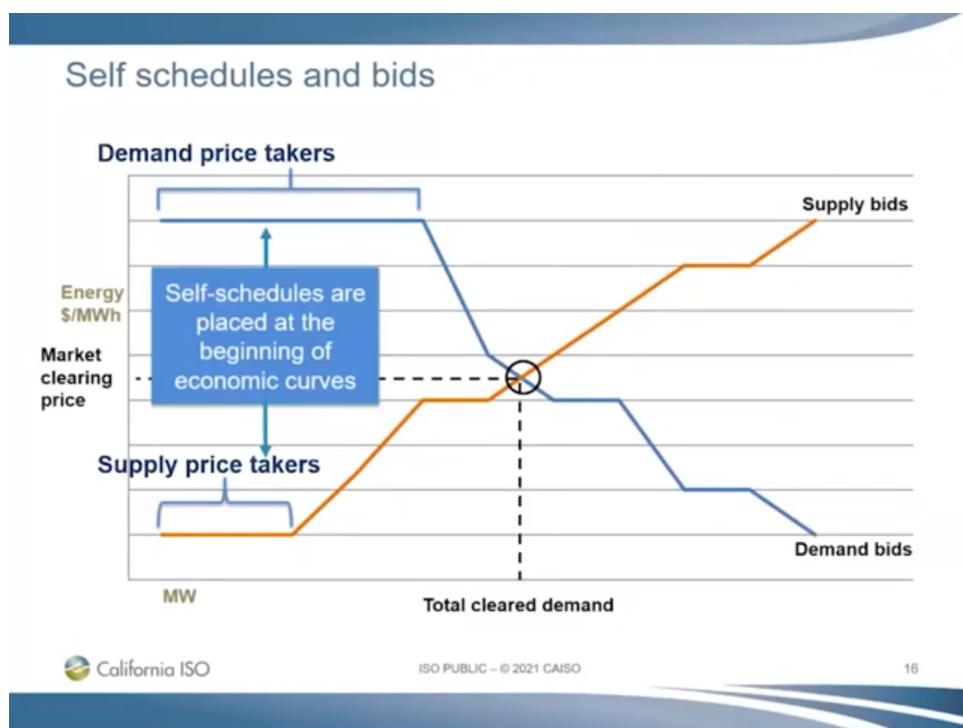


Fig. 1.3: Obtención del clearing price en el Day-ahead market.

En la figura 1.3 se puede visualizar como se obtiene el precio de la energía. En el proceso se busca el *clearing price* viendo cuántas ofertas de compra y venta hay y a qué precio. Se empieza aceptando las ofertas de quienes quieren vender su energía a un precio más

bajo y quienes quieren comprar a un precio más alto. El *clearing price* es el punto donde se intersecan las curvas. Cuando se llega al *clearing price*, inmediatamente se sabe que las ofertas aceptadas son las que quedan a la izquierda en el gráfico y las ofertas rechazadas son las que quedan a la derecha.

Algo que debemos destacar sobre las ofertas en el mercado DAM es que los SC también tienen la opción de hacer una oferta de una determinada potencia sin un precio asociado. Estas ofertas son conocidas como ***self-scheduled o price takers*** y son las primeras que serán aceptadas ya que no tienen una exigencia de precio y venderían la energía a cualquier costo. Este tipo de oferta sirve para que los SC aseguren que podrán vender su energía.

Luego de finalizar este proceso, a las 13 hs. del día anterior a la fecha, se envían instrucciones a los *Scheduling Coordinators* cuyas ofertas de venta hayan sido aceptadas.

Las ofertas aceptadas son las que hayan tenido un precio igual o por debajo al *clearing price*. En las instrucciones están los horarios en los cuales el SC deberá aportar su energía y cuánta potencia deberá aportar.

Todas las ofertas aceptadas se cobran al mismo precio por MWh, sin importar cuál haya sido su oferta. El precio final es el *clearing price* de la hora en el DAM.

Luego de que una oferta haya sido seleccionada, el *scheduling coordinator* se compromete a brindar la energía a la red, respetando la potencia indicada durante la hora de esa oferta.

Por ejemplo, si una oferta de 10 MWh para las 15hs es aceptada, el SC que haya hecho esa oferta va a recibir la instrucción por parte de CAISO de aportar 10 MW de energía desde las 15hs hasta las 16hs de la fecha para la cual se haya hecho la oferta. **De lo contrario, el SC deberá pagar una multa** por incumplir su compromiso. El monto de la multa es por la cantidad de energía que el SC haya fallado en brindar, al valor del mercado RTM.

Al publicarse los resultados y las instrucciones correspondientes a los SCs, el DAM termina y se inicia el *Real-time Market*.

### Real-time Market

El Real-time market, o RTM, es el mercado que complementa al *Day-ahead Market*.

El DAM conforma la base de la actividad del mercado el día de la fecha. Pero al cerrar el día anterior, es un mercado que no cuenta con la información de los imponderables que puedan surgir en el día.

Los resultados del mercado DAM constituyen una foto preliminar del día, pero no necesariamente es lo que efectivamente ocurre luego. Puede pasar que se necesite más energía de lo previsto, o que alguno de los SCs no pueda cumplir con sus compromisos del DAM y haya un faltante de energía. Así como otros SC tal vez puedan aportar más energía a la red de la que habían ofrecido en el DAM. O incluso podría pasar que se necesite menos energía de lo que se pensaba.

Para estos casos es que aparece el RTM. Su función es palear los faltantes o sobrantes de energía que el DAM no haya previsto.

El mercado abre a las 13 hs. del día anterior a la fecha, en el mismo momento en el que el DAM cierra y CAISO publica los resultados. Y cierra 75 minutos antes de cada hora en cuestión.

El RTM usa todos los inputs que utiliza el DAM, con la diferencia de que al estar más cerca de la hora en cuestión es información más precisa, y además usa otros inputs, como los resultados del DAM y las ofertas del RTM.

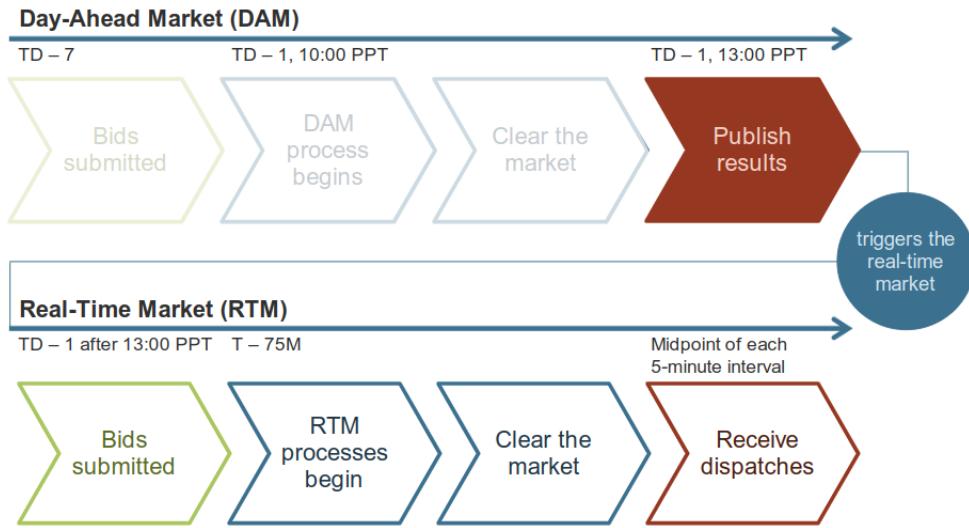


Fig. 1.4: Línea de tiempo del *Real-time market*.

Al igual que el DAM, el mercado RTM funciona con subastas. A diferencia del *-day-ahead market*, **los intervalos de tiempo son de cinco minutos**. Por lo tanto hay 288 intervalos a lo largo del día.

Por ejemplo, supongamos que un SC está enviando 10 MWh de energía a la red, correspondientes al DAM. El SC puede hacer una oferta para enviar más energía a un determinado precio, digamos otros 5 MWh. En caso de que la energía se necesite y el *clearing price* del RTM haya sido igual o mayor al de la oferta, se le enviará al SC una instrucción de aumentar la potencia a 15 MWh. Esto podría ser solo durante un intervalo (cinco minutos) o podría durar varios intervalos.

Como dijimos, el mercado funciona independientemente para cada hora del día. Cierra 75 minutos antes de la hora en cuestión. Luego de que el mercado cierra para esa hora, se emiten las instrucciones a los SC cuyas ofertas hayan sido aceptadas para que brinden la energía requerida.

### 1.2.2. Nuestro caso de estudio

A lo largo de nuestro trabajo tomamos como referencia a una empresa dedicada al almacenamiento y venta de energía que se encuentra en Lancaster, California.

Ellos comenzaron usando paneles solares en el mercado de energía de California, un mercado donde la energía solar tiene una participación alta. En este mercado ellos observaron que los valores de la energía eran mucho más altos durante la tarde o la noche, cuando no hay sol, que durante el día, cuando sí lo hay.

Debido a esto, incursionaron en el almacenamiento, para poder guardar la energía que obtenían con los paneles durante el día, y venderla durante la tarde o la noche.

La operatoria de esta empresa está centrada en una planta que cuenta con paneles solares que reciben energía solar. Para almacenar la energía de estos paneles, utilizan baterías usadas de vehículos eléctricos.

Actualmente, utilizan sus paneles para cargar las baterías durante el día, y a las 18hs. las descargan vendiendo su energía al mercado a través de un *scheduling-coordinator* como

intermediario. También venden la energía recibida por sus paneles cuando las baterías están completas.

Debido a diversos factores como que la problemática de ofertar energía al mercado tiene cierta complejidad, u otros desafíos relacionados a la infraestructura de la planta, optaron por una estrategia de venta de energía bastante simple. Lo que hacen es cargar las baterías una sola vez al día, descargándolas alrededor de las 18hs en el mercado DAM. Durante el día utilizan los paneles para cargar las baterías. Una vez que ya están cargadas, la energía de los paneles es directamente vendida al mercado RTM.

Algo importante a tener en cuenta es que la cantidad de energía que ellos proporcionan al mercado es bastante menor a comparación a los participantes más grandes, por lo que no tienen injerencia a la hora de fijar los precios de la energía. Debido a esto las ofertas que realizan al mercado son de tipo *price taker*, lo cual significa que no tienen una exigencia de precio, y serán aceptadas al precio de cierre del mercado, sea cual sea el mismo.

Esto se traduce a que, en la práctica, todas sus ofertas son aceptadas por CAISO, pudiendo vender toda su energía a los mercados. Es por esto que durante este trabajo vamos a suponer que las ofertas que hagamos al mercado de energía **siempre serán aceptadas**. Esto significa que siempre podremos vender la energía que tengamos al mercado. Esta suposición no afecta al resultado del trabajo ya que es algo que realmente se da en la práctica, y que sirve para simplificar el mismo pudiendo obviar la problemática de las subastas, cuya complejidad es bastante alta y ha sido abordada en otros trabajos<sup>4 5</sup>.

Otra presunción de dominio que haremos es decir que **sólo podemos vender al mercado DAM la energía obtenida por las baterías** y no la de los paneles. Hacemos esto porque la energía que reciben los paneles es muy cambiante y en la realidad no podríamos comprometernos a sostener una potencia constante durante una hora usando los paneles. En cambio sí podemos usar los paneles en el RTM donde los períodos de compromiso son más cortos.

De ahora en más, nos referiremos a esta empresa como **nuestro caso de estudio**.

### 1.3. Descripción de alto nivel de nuestro trabajo

La tesis consiste en analizar los datos que CAISO nos provee para poder construir un sistema que genere estrategias de venta que incremente las ganancias de nuestro caso de estudio. El trabajo está dividido en tres partes, las cuales detallaremos a continuación.

#### 1.3.1. Obtención y visualización de datos

La primera etapa es de **análisis descriptivo**. En esta etapa recolectamos datos históricos sobre los precios que tuvo la energía en los distintos mercados de CAISO. Para esto utilizamos el sistema de información de acceso de California.

<sup>4</sup> Y.-Y. Hong y C.-Y. Hsiao. «Locational marginal price forecasting in deregulated electricity markets using artificial intelligence». En: *IEE Proceedings - Generation, Transmission and Distribution* 149.5 (1 de sep. de 2002), págs. 621-626. ISSN: 1359-7051. DOI: [10.1049/ip-gtd:20020371](https://doi.org/10.1049/ip-gtd:20020371). URL: [https://digital-library.theiet.org/content/journals/10.1049/ip-gtd\\_20020371](https://digital-library.theiet.org/content/journals/10.1049/ip-gtd_20020371).

<sup>5</sup> Anjali Sheffrin. «Empirical Evidence of Strategic Bidding in the California ISO Real-time Market». En: *Electricity Pricing in Transition*. Ed. por Ahmad Faruqui y B. Kelly Eakin. Topics in Regulatory Economics and Policy Series. Boston, MA: Springer US, 2002, págs. 267-281. ISBN: 9781461508335. DOI: [10.1007/978-1-4615-0833-5\\_18](https://doi.org/10.1007/978-1-4615-0833-5_18). URL: [https://doi.org/10.1007/978-1-4615-0833-5\\_18](https://doi.org/10.1007/978-1-4615-0833-5_18).

OASIS, o su nombre completo *California ISO Open Access Same-time Information System*, es un **sistema de libre acceso que contiene todos los datos de CAISO**. Nosotros vamos a centrarnos en la información de precios históricos del mercado.

OASIS nos provee una API a partir de la cual podemos obtener esta información. Durante este trabajo desarrollamos un cliente en Python que consume esta API, parsea la información recibida en formato XML y guarda la información de los precios en una base de datos PostgreSQL.

Por otro lado también **utilizamos Visual Crossing, una API que provee datos del clima** en distintas partes del mundo, para obtener datos del clima en California.

Debimos obtener estos datos por dos razones. Una razón es para poder utilizarlos para detectar patrones en el clima que puedan influir en los precios. La otra razón es para poder tener una estimación de cuánta energía recibirían los paneles solares en un día.

Una vez obtenidos los datos **construimos un módulo que nos permitió visualizar la información** para poder detectar distintos patrones tanto en los precios como en el clima.

### 1.3.2. Predicción de precios

Habiendo ya obtenido y visualizado la información histórica de los precios en los diferentes mercados de CAISO, la segunda etapa consistió en **desarrollar un modelo predictivo capaz de proporcionar estimaciones de los precios futuros**. Nuestro propósito fue emplear estas predicciones como *input* para luego desarrollar estrategias de venta en CAISO. Optamos por emplear modelos de series temporales para este proceso.

Con el objetivo de lograr la máxima eficacia en la predicción, aplicamos el método de validación cruzada o *cross-validation* para determinar los hiperparámetros óptimos para el modelo. Finalmente, realizamos diversos experimentos para evaluar el rendimiento del modelo en distintos puntos temporales y, además, procedimos a compararlo con otras posibles heurísticas de predicción a fin de seleccionar la alternativa más precisa.

### 1.3.3. Solución de programación lineal entera

La etapa final de esta propuesta consistió en construir un modelo de programación lineal entera que brinde una **estrategia que indique qué hacer con los paneles y las baterías** en cada momento, de forma tal que la ganancia sea maximizada.

El modelo recibe como *input* los precios obtenidos del predictor en la segunda etapa. También recibe cuánta energía solar recibirán los paneles y el estado inicial de las baterías. Esta información se extrae de la API de *Visual Crossing* comentada en la sección anterior.

El output del modelo de programación lineal entera representará una estrategia de venta de energía para un participante del mercado de energía en California.

### 1.3.4. Sistema de decisiones y experimentación final

Con el objetivo de llegar a una solución integral y obtener un módulo que, a partir de datos históricos de los precios de mercado y un pronóstico de tiempo, nos brinde una estrategia de venta, diseñamos un sistema que utiliza las 3 partes enumeradas anteriormente.

Primero nuestro módulo utiliza los datos obtenidos de las distintas fuentes y a partir de los mismos realiza una predicción de los precios. Luego se toman muestras de estas

predicciones con determinadas distribuciones, las cuales serán detalladas en el capítulo 5.

Por cada muestra, se ejecuta nuestro modelo de programación lineal entera, el cual genera una solución. Al ejecutar el modelo con distintas muestras obtendremos diferentes estrategias que podríamos utilizar. A partir de una serie de tests, en los cuales probamos el comportamiento de estas estrategias en distintos casos, seleccionamos la que será finalmente utilizada.

Una vez establecido nuestro sistema de toma de decisiones, procedimos a llevar a cabo diversos experimentos para poder comparar nuestro sistema con otras posibles estrategias de venta de energía.

Durante esta experimentación comparamos la solución diseñada por nuestro sistema con las demás estrategias a lo largo de varios días en distintas épocas del año para poder determinar si la estrategia brindada por nuestro sistema es mejor que las otras.

## 2. ORIGEN DE DATOS

La realización de este trabajo requirió la **recolección de una gran cantidad datos de distintas fuentes**. Por un lado debimos recolectar información de los precios del *Day-ahead market* y del *Real-time market*. Necesitamos información histórica de los precios para luego poder predecirlos y también para realizar simulaciones que testeen nuestro modelo de decisiones.

Además, fue necesario adquirir registros del clima en California. Esto está motivado por dos razones. En primer lugar, consideramos que el clima sería un factor importante a la hora de determinar los precios del mercado. La temperatura incide en gran medida en la cantidad de energía consumida por la sociedad. Durante los días más cálidos, el consumo energético se eleva debido al uso intensivo de sistemas de aire acondicionado. Por lo tanto, se estimó que la temperatura constituiría un indicador valioso para anticipar las tendencias de los precios del mercado. Por otro lado, dado que el sistema se basa en energía solar, necesitamos información del clima para tener una estimación de cuánta energía nos brindarán los paneles solares durante el día, lo cual es determinante para posteriormente elaborar una estrategia de venta.

El trabajo de esta etapa no se limitó únicamente a una tarea de recopilación de datos. A medida que avanzábamos, surgía la necesidad de visualizar y comprender los datos reunidos. Por esta razón, **se desarrolló un dashboard para la visualización de datos**.

En las siguientes secciones detallaremos las dos fuentes de datos que consultamos, seguido por una descripción de la herramienta de visualización que implementamos y de las conclusiones que pudimos extraer gracias a su uso.

### 2.1. OASIS

OASIS<sup>1</sup> es el sistema de información de libre acceso en tiempo real (Open Access Same-time Information System) de CAISO.

Este sistema provee una API<sup>2</sup> en la cual se pueden descargar diversos reportes de la actividad en el mercado de energía. Usando esa API, podemos descargar vía HTTP la información de CAISO en formato XML. La información se descarga mediante reportes.

El formato del pedido HTTP para obtener los datos es el siguiente:

`http://oasis.caiso.com/oasisapi/SingleZip?{reporte}&{fecha_desde}&{fecha_hasta}&{mercado}&{version}&{nodo}`

Donde:

- **reporte:** es el nombre del reporte que queremos descargar de OASIS. En nuestro caso queremos los reportes de precios, los cuales son *PRCLMP* para DAM y *PRCINTVLLMP* para RTM.
- **fecha\_desde** y **fecha\_hasta:** para delimitar el rango de fechas de los datos que queremos tener en el reporte.

---

<sup>1</sup> OASIS Prod. URL: <http://oasis.caiso.com/mrioasis/logon.do>.

<sup>2</sup> OASIS User-Interface Application Version 7.0.0: Deployment Scheduled on 2/19/19. URL: <https://www.caiso.com/Documents/OASISUser-InterfaceApplicationVersion700-DeploymentScheduled-021919.html>.

- mercado: Hay que especificarle a OASIS cuál es mercado para el queremos la información. Para *PRCLMP* el parámetro debe ser DAM y para *PRCINTVLLMP* debe ser RTM.
- versión: La versión de la API que usamos.
- nodo: Hace referencia a la zona de California de la cual pedimos la información. En nuestro caso fue Lancaster.

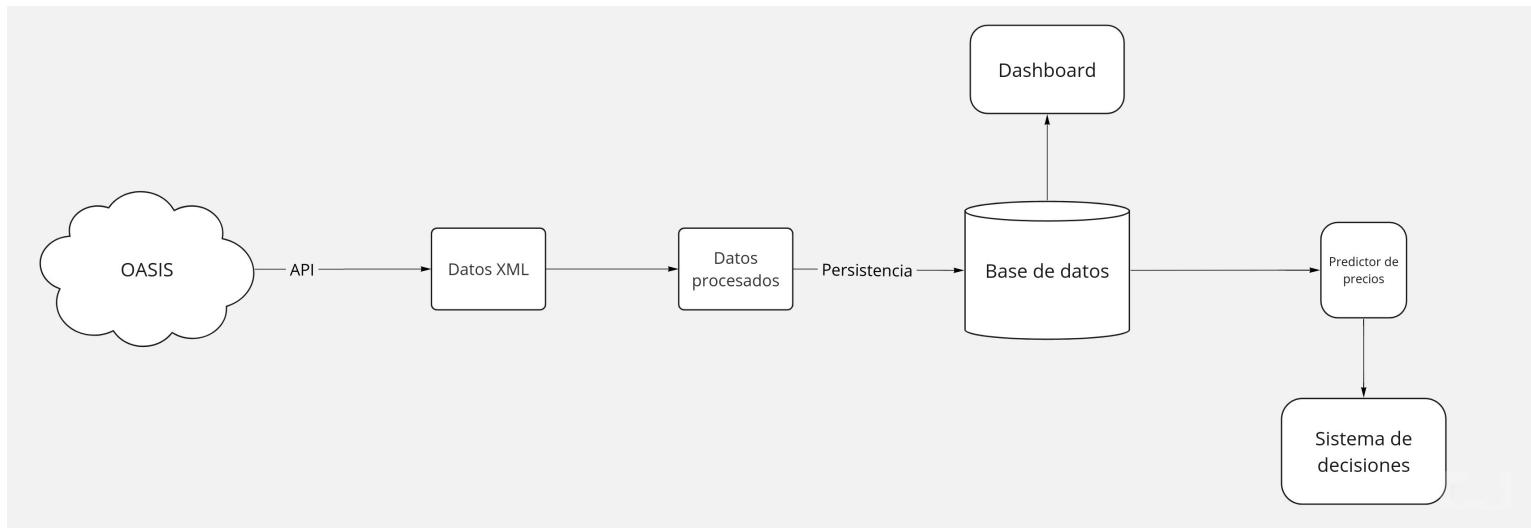


Fig. 2.1: Diagrama de obtención y utilización de datos de OASIS.

Al hacer el pedido a la API de OASIS utilizando esta URL, esta nos provee con un archivo ZIP que contiene los datos en formato XML.

Nosotros parseamos esos datos para quedarnos con una versión final. Finalmente persistimos estos datos en una base de datos. La base de datos consta de una única tabla donde guardamos los precios de cada mercado en cada hora del día para el mercado DAM y cada cinco minutos para el mercado RTM.

La siguiente figura contiene de algunos datos de la tabla para visualizar la forma de los datos:

id	market	price	datetime
441994	RTM	46.80861	2020-10-25 15:05
441995	RTM	41.67369	2020-10-25 15:10
441996	RTM	47.1912	2020-10-25 15:15
441997	RTM	39.22987	2020-10-25 15:20
441998	RTM	34.48952	2020-10-25 15:25
131387	DAM	31.01262	2020-10-25 16:00
131403	DAM	24.54987	2020-10-25 17:00
131402	DAM	20.93661	2020-10-25 18:00
131398	DAM	22.73857	2020-10-25 19:00
131395	DAM	19.64236	2020-10-25 20:00

Como se puede ver, tenemos una entrada en la tabla por cada hora del día para el mercado DAM, y una cada cinco minutos para el mercado RTM. En cada fila se encuentra la fecha y hora, el mercado y el precio de la energía en ese momento.

Como se ve en la figura 2.1, estos datos son posteriormente consumidos por el *dashboard* para su visualización, y también por el predictor de precios y el sistema de decisiones.

## 2.2. Clima

El clima desempeña un papel sumamente relevante en nuestro problema. Hay tres variables que tuvimos en cuenta a lo largo de este trabajo: la **temperatura**, la **radiación solar**, y la **nubosidad**.

La **temperatura tiene un impacto significativo en los patrones de consumo de energía** en la sociedad. En días de temperatura alta, el consumo de energía se eleva en gran medida debido al uso de aires acondicionados. Este incremento en la demanda impacta notablemente en el precio de la energía.

En el siguiente gráfico mostramos un ejemplo de lo que estamos diciendo. Los datos presentados muestran los precios observados en el mercado DAM durante dos períodos de tiempo distintos: del 1 al 30 de marzo y del 1 al 30 de agosto. Cada punto de datos del gráfico representa un día distinto. El valor que se muestra es el promedio de los precios de todas las horas del día.

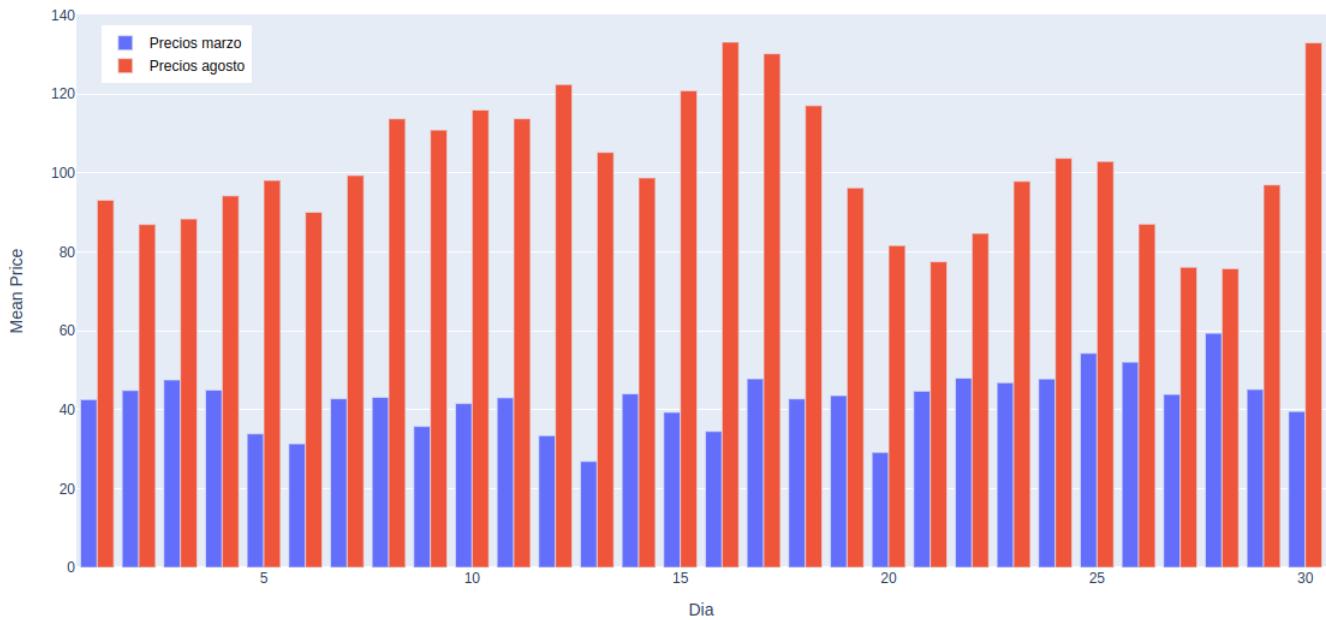


Fig. 2.2: Comparación de precios del DAM en marzo y agosto.

Es notorio que los precios de agosto, época de pleno verano en California, son mayores a los precios de marzo, cuando las temperaturas son más moderadas.

Por estos motivos nos pareció que obtener datos de la temperatura sería importante en la etapa de análisis predictivo, donde necesitaremos desarrollar un predictor de precios. En particular, necesitamos saber la temperatura de Lancaster para cada hora del día.

En cuanto a la **radiación solar y la nubosidad, ambas son fundamentales a la hora de entender cuánta energía recibirán los paneles solares**. Al ser nuestro objetivo final la formulación de estrategias de comercialización de energía, contar con detalles sobre la energía disponible a lo largo del día adquiere una importancia crucial.

Para obtener toda esta información, utilizamos la **API de Visual Crossing<sup>3</sup>**. Esta API nos provee datos del clima, tanto históricos como pronósticos de hasta quince días, desglosados en horas. En este trabajo aprovechamos esta herramienta para obtener tanto la temperatura como la radiación solar y la nubosidad de Lancaster.

En la tabla 2.1 mostramos algunos de los datos del clima obtenidos utilizando *Vissual Crossing*.

datetime	cloudcover (%)	temperature ( ° F )	solar radiation ( W/m <sup>2</sup> )
2023-06-14 06:00:00.000	16	67	21
2023-06-14 05:00:00.000	23	67.9	26
2023-06-14 04:00:00.000	23	69	32
2023-06-14 03:00:00.000	23	70.1	44
2023-06-14 02:00:00.000	23	74.9	183
2023-06-14 01:00:00.000	23	78	375
2023-06-14 00:00:00.000	23	81.9	624
2023-06-13 23:00:00.000	23	83	787
2023-06-13 22:00:00.000	0.8	83.4	714

Tab. 2.1: Datos del clima obtenidos en *Vissual Crossing*.

El flujo de obtención de datos fue muy similar al que realizamos para OASIS, comentado en la sección anterior. Obtuvimos los datos a través de una la API. Luego los procesamos y los persistimos en una base de datos.

Estos datos fueron consumidos posteriormente por nuestro *dashboard*, nuestro modelo predictivo de precios, y nuestro sistema de toma de decisiones.

### 2.3. Visualizacion de datos

Habiendo recopilado datos tanto de los precios de mercado como del clima, la siguiente etapa consistió en poder visualizarlos.

Nuestra intención era analizar día a día el comportamiento de los mercados de una forma fácil y rápida, lo cual nos permitiría observar cómo sus precios cambian a lo largo de los distintos días, encontrar patrones, etc.

También queríamos poder ver datos del clima y cómo se comporta a lo largo del tiempo. Puntualmente nos interesaba ver los datos de la temperatura, ya que según nuestra hipótesis, es la variable que tiene una mayor influencia en los precios.

Por último, quisimos entender cómo se relaciona la temperatura con los precios de la energía, por lo que también necesitamos poder ver ambas cosas en un mismo gráfico.

Debido a esto nuestro siguiente paso fue armar un *dashboard* en el que podamos visualizar toda esta información. A la hora de armar este *dashboard* encontramos una herramienta que nos fue muy útil, llamada *Streamlit*.<sup>4</sup> Esta biblioteca de Python facilita la manipula-

<sup>3</sup> Weather Data & Weather API | Visual Crossing. URL: <https://www.visualcrossing.com/>.

<sup>4</sup> Streamlit Docs. URL: <https://docs.streamlit.io/>.

ción, filtrado y presentación de datos, y simplifica en gran medida la creación de interfaces gráficas para una visualización dinámica.

Utilizando esta herramienta armamos un *dashboard* con el que podíamos observar para cada día, los precios de los distintos mercados. Todos los gráficos que mostremos en esta sección habrán sido generados por nuestra herramienta de visualización.



Fig. 2.3: Precios del 04/04/2022 en el mercado DAM.

En la imagen se puede ver por un lado cómo uno puede configurar el *dashboard* utilizando los campos con los que podemos determinar las fechas, mercados y qué información queremos observar. En este caso presentamos **todos los precios de un día puntual en el mercado DAM**. Para el DAM podemos ver el precio del mercado en cada hora, y para RTM cada 5 minutos.

También fue de nuestro interés poder, a partir de un rango de fechas, agrupar los datos por hora y verlos en promedio. De esta manera **pudimos ver como se comportan los mercados en diferentes horas del día a lo largo del tiempo**. Esto se puede observar en la figura 2.4.

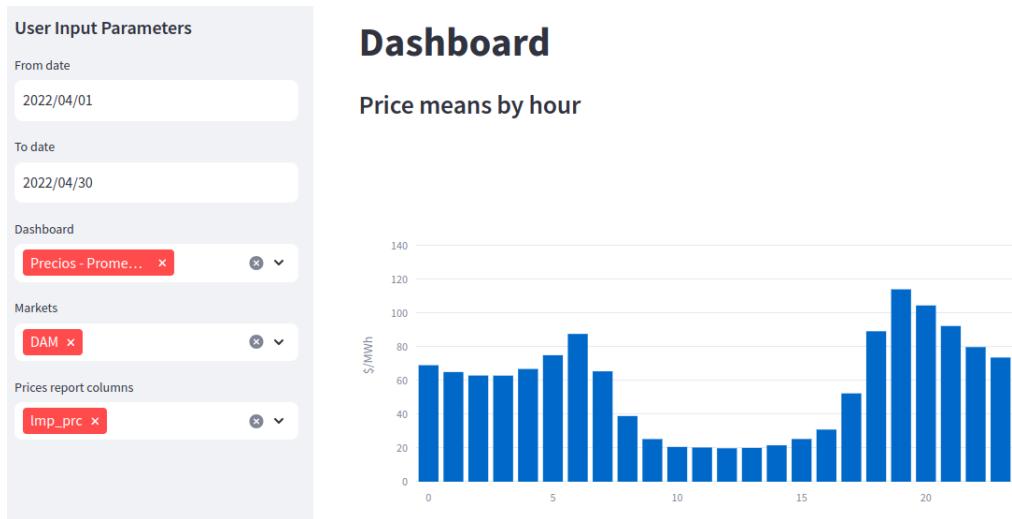


Fig. 2.4: Precios promedio del mes de abril del 2022 en el mercado DAM.

En este caso, a pesar de que la curva está un poco más plana, podemos ver que lo observado en la figura 2.3 es un patrón que se dio en todo el mes de abril.

Observando estos gráficos, detectamos un patrón. Los precios de la energía son mayores por las mañanas y las tardes/noches. Mientras que al mediodía los precios disminuyen. Este es un fenómeno conocido en el mercado de energía como la *duck curve*<sup>5</sup>.

La *duck curve*, o en español la curva del pato, ilustra que, debido a la mayor presencia de energía solar, los precios del mercado disminuyen en el mediodía y por la tarde. Esto se debe a una mayor oferta de energía en esa franja horaria, donde la actividad solar está en su pico. A su vez, los precios tienen punto máximo alrededor de las 18hs. cuando el sol empieza a caer, pero aún hay demanda de energía.

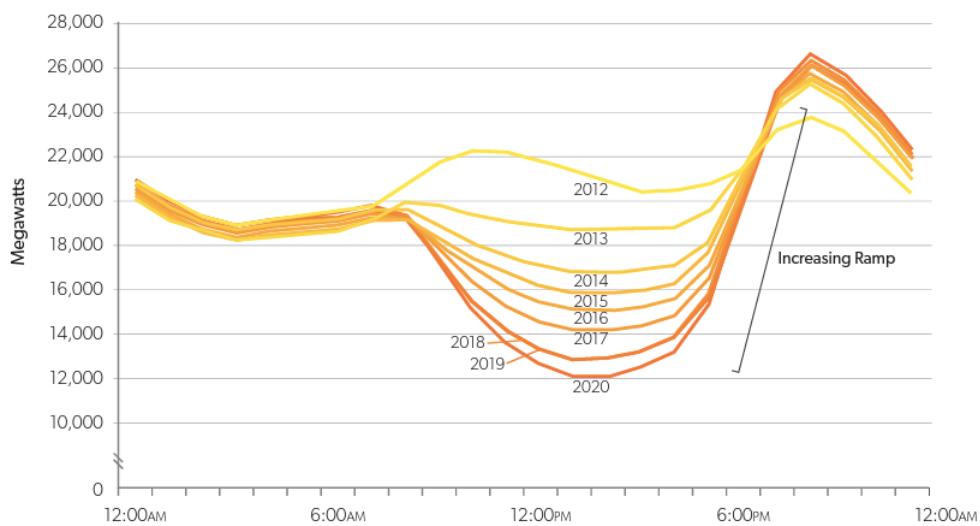


Fig. 2.5: Curva del pato.

<sup>5</sup> [https://www.caiso.com/documents/flexibleresourceshelprenewables\\_fastfacts.pdf](https://www.caiso.com/documents/flexibleresourceshelprenewables_fastfacts.pdf)

En la figura 2.5 podemos ver la demanda de energía convencional, no renovable, a lo largo de los años del 2012 al 2020. Podemos ver cómo la demanda baja considerablemente luego de las 6 hs. y sube hacia las 18 hs. Esto se debe a que durante el día se suma a la oferta de energía convencional, una gran oferta de energía solar. Esta curva de demanda es acompañada por los precios, debido a que los precios suben cuando la oferta es menor.

Notar como de esta forma pudimos verificar visualmente y de una forma rápida que los datos obtenidos utilizando OASIS se comportan de una manera ya reconocida por quienes estudian el mercado de energía. Algo que hubiese sido más difícil de lograr de ser observado en una planilla.

Continuando con las funcionalidades del *dashboard*, nuestro siguiente paso fue darle la posibilidad de **comparar ambos mercados en un mismo gráfico**. Esto se puede ver en la figura 2.6, la cual muestra los precios del mercado DAM y RTM a lo largo del mes de abril del 2022. La figura permite ver que el comportamiento de ambos mercados es similar en cuanto a picos y valles. Precios altos durante la noche, comenzando a bajar a las 7hs y subiendo a partir de las 17hs. Ambos se comportan siguiendo el fenómeno *duck curve* que describimos anteriormente.

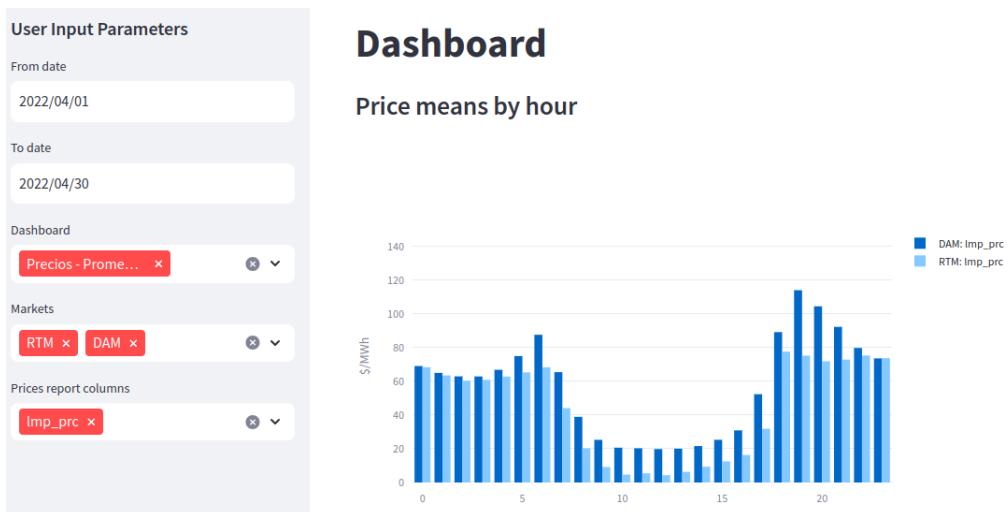


Fig. 2.6: Comparación del DAM y el RTM en el mes de abril del 2022.

Notar que para hacer este gráfico, los precios del RTM pertenecientes a una misma hora del día (hay 12 precios, pertenecientes a 12 intervalos distintos por cada hora), fueron agrupados y promediados.

Vale la pena destacar que por más que los mercados tengan un comportamiento similar en la forma general, sí hay una diferencia entre los mercados. **El RTM es mucho más variable que el DAM**, lo que puede ser visto en la figura 2.7.



Fig. 2.7: Precios del RTM el 6 de abril del 2022.

En la figura 2.7 podemos ver todos los precios del RTM en un día. Notar que en el caso de este mercado ya no hay una barra por hora, sino que hay una cada cinco minutos. Podemos ver como los saltos son mucho mas bruscos que en el DAM, aunque en este día particularmente seguimos observando los patrones de la *duck curve*. Sin embargo, hay otros días, como podemos ver en la figura 2.8 donde estos patrones se pierden. En este caso vemos un pico muy grande antes de las 15hs. y otro antes de las 21hs. mientras que el resto del día la curva se mantiene plana. **Este tipo de patrones en el comportamiento del RTM nos mostró que es un mercado mucho más variable**. Esto es razonable dado que los precios del RTM se deciden en tiempo real a partir de imprevistos que surjan en cada momento, lo cual le da al mercado esta variabilidad.



Fig. 2.8: Precios del RTM el 19 de julio del 2022.

Una vez lograda la visualización de precios. Pasamos a **visualizar la temperatura** de la misma forma. Podemos ver tanto la temperatura de cada hora del día para cualquier

día, como ver el promedio de la temperatura en cada hora del día para un rango de fechas.



Fig. 2.9: Temperatura del 1ro de abril del 2022 en Lancaster.

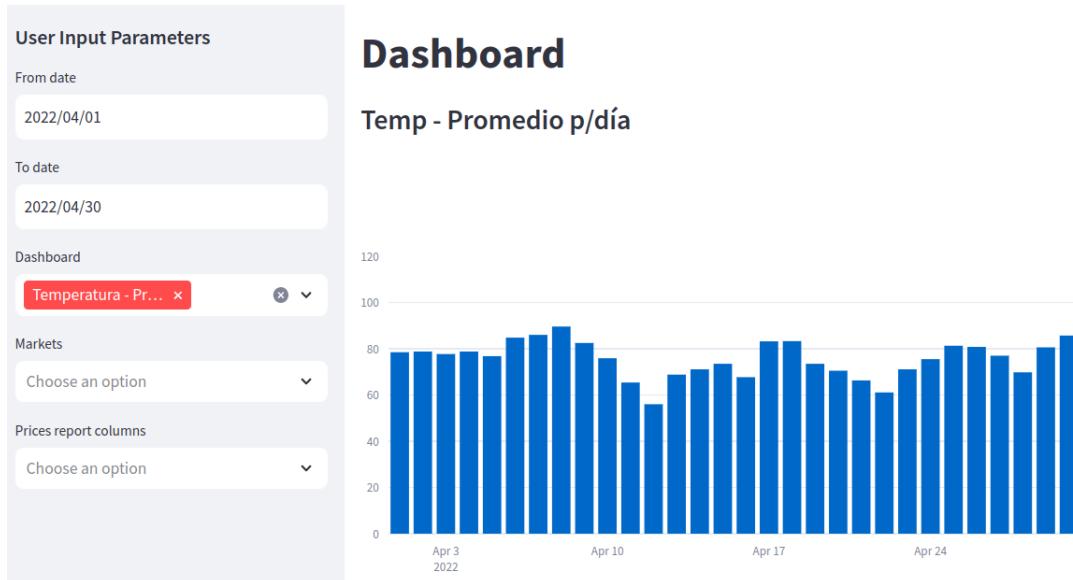


Fig. 2.10: Temperatura promedio del mes de abril del 2022 en Lancaster.

En la figura 2.9 se puede ver la temperatura del día 4 de abril del 2022, hora por hora, mientras que la figura 2.10, muestra la temperatura de todo el mes de abril, día por día, con el promedio de la temperatura de cada día.

Finalmente con el objetivo de ver si los precios y la temperatura se relacionan de algún modo, generaremos un gráfico que muestra los precios del mercado junto con la temperatura de ese día.

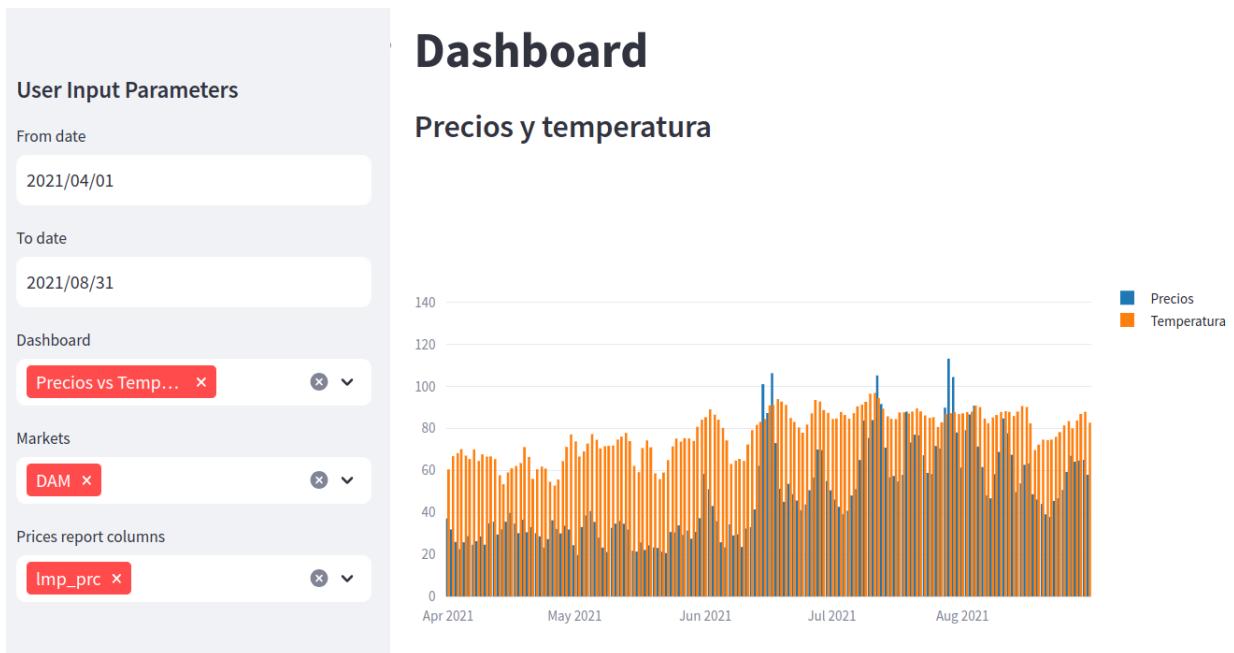


Fig. 2.11: Precios del DAM y temperatura desde abril hasta agosto del 2021.

Este gráfico nos permitió observar que los precios se elevan en los meses de julio y agosto, en estos meses también se puede observar que la temperatura también aumenta.

## 2.4. Conclusion del capítulo

La generación de este *dashboard* fue el resultado de la investigación y utilización de datos de distintas fuentes como lo fueron OASIS y *Visual Crossing*. Utilizamos estas APIs para obtener información desde el 2021 hasta la actualidad.

Logramos construir una **herramienta muy simple** que permite, como observamos a lo largo del capítulo, visualizar todos estos datos fácilmente. La herramienta nos fue **de mucha utilidad a lo largo de todo el trabajo**, pudiendo comprender los datos, detectar patrones y verificar hipótesis rápidamente, lo cual hubiera sido imposible de otra manera. Además, fue programada de manera que sea flexible y que sea **muy fácil extenderla** agregando más gráficos o datos de otras fuentes, por lo que pudimos ir extendiéndola a medida que surgían distintas necesidades.

También es importante destacar que esta herramienta **podría ser utilizada por distintos actores** en el mercado de energía **sin la necesidad de conocimientos técnicos**. Podría ser usada por ejemplo por directivos de una empresa que quieran poder entender tendencias del mercado en las últimas semanas, o en un período del año en particular.

### 3. MODELO PREDICTIVO

En este momento del trabajo ya disponemos de datos históricos de precios y de temperatura en Lancaster, California.

El **siguiente paso es poder predecir los precios del mercado** en un día en particular. Esto es crucial, ya que estas predicciones serán luego utilizadas como input para nuestro modelo de programación lineal entera, que determinará la estrategia de venta de energía que maximice la ganancia.

Para que nuestro modelo nos brinde una estrategia realmente buena, tenemos que abastecerlo con precios que sean similares a los precios reales. Si los precios que le proveemos a nuestro modelo distan mucho de la realidad, no vamos a poder garantizar la obtención de una buena estrategia.

En este capítulo vamos a hablar del modelo predictivo que usamos en este trabajo y vamos a compararlo con otras alternativas que podríamos utilizar.

Como ya vimos en el capítulo pasado, puntualmente en la sección 2.3, el mercado RTM es muy variable y depende mayormente de imprevistos que surjan en el día. El hecho de que el RTM se base mayormente en imprevistos hace que no tenga mucho sentido intentar predecirlos en base a información histórica. Lo que haya pasado hace una semana no va a influir en el precio del RTM hoy, mucho menos lo que haya ocurrido hace un mes o un año.

Es por este motivo que **nuestro objetivo será predecir los precios del mercado DAM y no del mercado RTM**.

#### 3.1. Modelos predictivos para series de tiempo

Los modelos de series de tiempo trabajan con datos que están secuenciados en un intervalo de tiempo. Estos modelos pueden observar cómo los datos van cambiando a lo largo de ese intervalo y a partir de eso detectar patrones que lo ayuden a inferir qué ocurrirá en el futuro. Generalmente requieren una gran cantidad de datos para poder ser confiables y funcionan bien cuando los datos se comportan de forma cíclica.

Nosotros disponemos de los datos de precios del mercado a lo largo del tiempo. Tanto del DAM como del RTM, tenemos el precio para cada hora de cada día.

Además, sabemos que los datos se comportan de forma cíclica ya que como vimos en el capítulo anterior, siguen el patrón de la *duck curve*.

Dadas estas características, supusimos que un modelo basado en series de tiempo sería bueno para predecir estos precios.

También estuvieron en consideración otros métodos más generales como por ejemplo utilizar redes neuronales recurrentes para la predicción de precios. Pero éste es un método cuya utilización requiere una configuración más específica, la cual llevaría un mayor tiempo y esfuerzo. Por este motivo y porque este no es un trabajo cuyo foco esté en utilizar técnicas de *machine learning*, optamos por un método de utilización más rápida como lo es series de tiempo. Además de esto, debido a estudios realizados sobre este conjunto de datos por ingenieros que trabajan en la empresa que utilizamos como caso de estudio, es de nuestro conocimiento que el uso de redes neuronales recurrentes no otorgó grandes resultados para la predicción de los precios.

### 3.2. Prophet

Para implementar nuestro modelo de series de tiempo utilizamos *Prophet*,<sup>1</sup> una herramienta de predicción de series de tiempo desarrollada por *Facebook*. *Prophet* fue pensado para hacer predicciones a gran escala. Ellos se refieren a escala en dos aspectos:

- Un gran número de personas haciendo predicciones, muchas de ellas sin conocimiento de métodos de series de tiempo.
- Una gran variedad de problemas con distintas características para los que se necesitan predicciones.

Esto significa que *Prophet* es una **herramienta de fácil uso**, que se adapta a una gran variedad de problemas y que se puede configurar y utilizar fácilmente.

Esta herramienta también se adapta fácilmente a estacionalidad en los datos - tanto diaria, semanal, mensual o anual - y a fechas especiales o feriados (navidad, superbowl, etc).

Además esta herramienta tiene la cualidad de que uno puede por un lado contar con un modelo predictor muy rápidamente, con tan solo darle los datos de entrenamiento. Pero a su vez es **flexible y permite configurarlo** para que se adapte mejor a **cada caso**.

#### 3.2.1. Cómo funciona Prophet

El modelo de *Prophet*<sup>2</sup> consta de 3 componentes principales: **las tendencias, los feriados y la estacionalidad**. Estos tres componentes son combinados en la siguiente ecuación:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

En esta ecuación,  $g$  es la función de tendencias, que modela los cambios no periódicos en los datos. La función  $s$  es la función de estacionalidad, que representa los cambios periódicos en los datos. Y  $h$  representa el efecto de los feriados o fechas especiales.

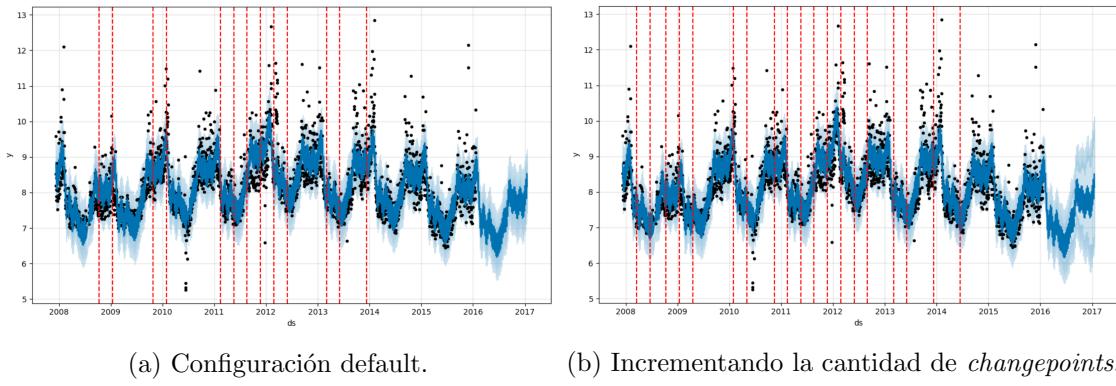
En cuanto a las tendencias, *Prophet* analiza los datos de entrenamiento y detecta los cambios que hay en los mismos. En nuestro caso, detecta cuándo suben los precios y cuándo bajan. Se conoce como **changepoints a los puntos** donde *Prophet* detecta un **cambio en las tendencias**.

Para poder ilustrar qué son los *changepoints*, incluiremos un ejemplo tomado de la documentación de *Prophet*. Los datos utilizados están constituidos por el historial de visitas de la página de wikipedia de Peyton Manning. A continuación mostramos dos imágenes de los *changepoints* que el modelo detecta a partir de estos datos.

Para evitar caer en *overfitting*, *Prophet* no utiliza todos los *changepoints* que encuentra, sino que se queda con algunos. Por default, utiliza un 5 % de los *changepoints* encontrados. Podemos ver eso en la figura 3.1a, donde vemos que por ejemplo entre el 2010 y el 2011 hay una caída en los datos y sin embargo no se registran *changepoints*. Se puede configurar el modelo para aumentar o disminuir la cantidad de *changepoints* utilizados, como se ve

<sup>1</sup> *Prophet*. Prophet. URL: <http://facebook.github.io/prophet/>.

<sup>2</sup> Sean J. Taylor y Benjamin Letham. *Forecasting at scale*. e3190v2. PeerJ Preprints, 27 de sep. de 2017. DOI: [10.7287/peerj.preprints.3190v2](https://doi.org/10.7287/peerj.preprints.3190v2). URL: <https://peerj.com/preprints/3190>.

Fig. 3.1: Ejemplos de *changepoints* en Prophet

en la figura 3.1b, donde utilizamos el 25 % de los *changepoints*. También se pueden poner *changepoints* manualmente.

Podemos observar además que en ambos modelos no se ubican *changepoints* en el segmento final de los datos, esto también puede ser configurado, como veremos más adelante en el capítulo.

*Prophet* también permite configurar fechas especiales, como feriados o días donde ocurrirá un evento importante. De esta forma, el modelo puede saber que esos días son distintos al resto y tratarlos de forma particular y no como cualquier otro día. Se pueden configurar los feriados de un país en particular en caso de que los datos sean de ese país, y también se pueden configurar manualmente. También se puede configurar algún evento que tome varios días seguidos, como la semana de navidad.

Finalmente, *Prophet* permite configurar estacionalidad. Con esto nos referimos a que la herramienta divide el tiempo en estaciones y utiliza esa información a la hora de hacer predicciones. Por ejemplo, para datos que cambian mucho año a año, será conveniente agregarle al modelo una estacionalidad anual.

Por default, la herramienta utiliza estacionalidad anual, semanal y diaria para el modelo. También se pueden definir estacionalidades particulares que uno quiera. Además se puede configurar para cada estacionalidad, qué tanta relevancia se le da.

### 3.2.2. Primeros pasos

Como dijimos en la sección anterior, una ventaja de *Prophet* es que al utilizarlo uno es capaz de obtener un predictor inicial tan solo con brindarle los datos de entrenamiento, sin tener que configurar nada. Este predictor no será tan bueno como el que podemos obtener configurando algunos parámetros de manera tal que el modelo se aadecue mejor a nuestros datos, pero nos permite poder probar la herramienta rápidamente.

A continuación mostramos el **resultado de entrenar un modelo básico**, con datos de precios del 1 de enero del 2020 al primero de julio del 2020, o sea 180 días de datos de entrenamiento. A ese modelo le pedimos que prediga los precios del 2 de julio del 2020.

En el resultado, visible en la figura 3.2, podemos observar que, si bien el modelo no captura del todo bien la magnitud del pico entre las 18 y las 20hs, sí **captura bastante bien la forma general en la que los precios** se comportan durante el día. Con un pequeño pico a la mañana, luego una baja al mediodía y otro pico por la tarde/noche.

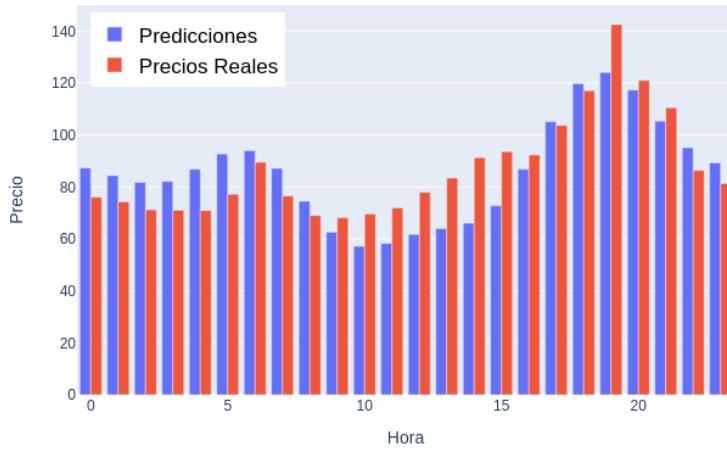


Fig. 3.2: Precios del DAM del predictor vs precios reales.

### 3.3. Buscando el mejor modelo

Como dijimos anteriormente, *Prophet* nos da muchas formas de mejorar el predictor inicial que mostramos en la sección anterior. En esta sección vamos a hablar de las **distintas estrategias que fuimos empleando para mejorarlo**.

Inicialmente el modelo sólo toma los datos de entrenamiento, pero hay muchas cosas que podemos hacer además de eso.

Para cada cambio aplicado al modelo, realizamos una comparación para comprobar si fue una mejora, si el modelo empeoró o si mantuvo resultados similares.

Finalmente **nos quedamos con un modelo final que obtenga los mejores resultados posibles** e hicimos una **comparación con otras posibles heurísticas** de predicción de precios.

#### 3.3.1. ¿Cómo comparar los modelos?

La forma que elegimos para comparar los distintos modelos que fuimos generando fue la **raíz del error cuadrático medio (a partir de ahora *RMSE*)**.

La raíz del error cuadrático medio es una medida muy común a la hora de comparar conjuntos de datos. Sea  $P$  el vector que representa el conjunto de datos resultante de una predicción y  $O$  el conjunto de datos reales observados, la formula para calcular esta métrica es la siguiente:

$$\sqrt{\frac{\sum_1^n (P_i - O_i)^2}{n}} \quad (3.1)$$

Para comparar distintos modelos, los evaluamos con datos del pasado en distintos días. Para cada día se predicen los precios, comparamos con la realidad y calculamos el *RMSE* para ese día.

Ejecutamos los modelos a partir del dia 1 de abril de 2022, y comparamos en los siguientes 180 días consecutivos. O sea, hasta el 27 de septiembre del mismo año.

En cada día comparamos el nuevo modelo contra el anterior y luego vimos en cuántos días era mejor uno y otro, además **calculamos la media y la mediana del RMSE de ambos modelos a lo largo de los 180 días.**

### 3.3.2. Instancia de ejecución de la experimentación

De ahora en adelante vamos a hablar de los distintos experimentos que hicimos con el fin de encontrar el mejor predictor posible utilizando *Prophet* y comparándolo con otras alternativas.

Todos estos experimentos fueron ejecutados en una computadora modelo *Dell Latitude 7400*, con un procesador modelo *i7-8665U* y 16 GB de memoria y utilizando *Python 3.9*.

### 3.3.3. Búsqueda de hiperparámetros

Nuestro primer intento para mejorar el modelo fue una búsqueda de hiperparámetros de *Prophet*.

A continuación enumeramos los hiperparámetros de *prophet* que usamos:

- ***changepoint\_prior\_scale*** Como vimos antes, *Prophet* no usa todos los *changepoints* que encuentra en los datos. Entre **más alto es este parámetro, el modelo usa más *changepoints***. Esto hace que se pueda adaptar más a los datos de entrenamiento. Al aumentar este número, aumenta el **riesgo de overfitting** cuando se usan demasiados *changepoints*, ya que se ajustaría demasiado a los datos.
- ***changepoint\_range*** Como ya observamos en las figuras [3.1a](#) y [3.1b](#) de la sección anterior, **por default *Prophet* evita usar *changepoints* en el final de la serie de tiempo**, dejando afuera los *changepoints* ubicados en el último 20% de los datos. Elevando el valor de este parámetro, **se puede hacer que el modelo use *changepoints* en toda la serie de tiempo**. Esto sirve si uno tiene series cuyos valores varían en forma muy paulatina y por lo tanto los datos a predecir sean probablemente similares a los últimos datos de la serie.
- ***seasonality\_prior\_scale*** entre más alto sea este valor, el modelo le dará más relevancia a los efectos de la estacionalidad, ya sea semanal, mensual o anual.
- ***weekly\_seasonality*** Entre más alto sea este parámetro, el modelo le dará mayor relevancia a la estacionalidad semanal.
- ***yearly\_seasonality*** Entre más alto sea este parámetro, el modelo le dará mayor relevancia a la estacionalidad anual.

En la tabla [3.1](#) podemos observar los distintos valores que probamos para cada parámetro. En esa tabla podemos ver con un asterisco al valor por default que usa *Prophet* en cada uno.

Campo	Valores						
changepoint_prior_scale	0.005	0.01	<b>0.05*</b>	0.1	0.5	0.9	0.99
seasonality_prior_scale	0.1	1.0	<b>10.0*</b>	20.0	25.0		
changepoint_range	0.7	<b>0.8*</b>	0.9	0.95	0.99		
weekly_seasonality	1	<b>3*</b>	5	10	20		
yearly_seasonality	1	<b>3*</b>	5	10	20		
daily_seasonality	1	<b>3*</b>	5	10	20		

Los valores indicados con un \* son los valores default que tiene cada parámetro

Tab. 3.1: Hiperparámetros para el modelo de Prophet

Para hacer la búsqueda hicimos un producto cartesiano de todos estos valores para todos estos parámetros. Esto nos dio en total 17500 posibilidades diferentes.

Lo que tuvimos que hacer fue evaluar cada posible combinación de parametros para ver cuál era la mejor. Para esto utilizamos la técnica de validación cruzada o *cross-validation*.

### Cross validation para series de tiempo

*Cross-validation* es un método para evaluar el funcionamiento un modelo predictivo. Esencialmente consiste en **usar una parte de los datos a disposición para entrenar un modelo**, y apartar otra para luego **validar el modelo con estos últimos datos**.

Es común que al hacer *cross-validation* no se elija un solo conjunto de validación, sino que se elijan varios. Se eligen  $k$  conjuntos, o en inglés *folds*, y para cada uno se entrena el modelo con el resto de los datos y se valida posteriormente utilizando los datos del *fold*. De esta forma se puede validar el modelo varias veces y no quedarse con una única evaluación, obteniendo una validación más robusta. A este método se lo conoce como  **$k$ -fold cross-validation**.

Sin embargo, a la hora de utilizar el método *cross-validation* para series de tiempo, se debe tener en cuenta la particularidad de que **no se pueden tomar conjuntos aleatorios para usar para validación**, porque los datos tienen un orden particular. Uno tiene que validar con datos posteriores a los datos que se usan para entrenar. Entonces métodos clásicos, como *k-fold cross-validation* no funcionan.

La forma de hacer *cross-validation* con series de tiempo que utilizamos consiste en elegir varios puntos de corte para nuestros datos. Cada punto de corte es una fecha particular.

Para cada punto de corte, utilizamos una cierta cantidad de datos desde el mismo hacia atrás y entrenamos el modelo con esos datos. Una vez entrenado, evaluamos el modelo para una cantidad de datos determinada a partir del punto de corte. A esta porción de datos la llamamos *horizonte*. En nuestro caso como nuestro objetivo es tener un predictor que funcione para un día individualmente, utilizaremos *horizonte* con una longitud de un día.

En caso de tener  $k$  puntos de corte, al terminar tendremos las predicciones para los  $k$  días. Luego de tener las predicciones hay que medir que tan buenas son. La forma en que hicimos esto fue calcular el error cuadrático medio de las predicciones con respecto a los precios reales de esos días.

## Cross validation para series de tiempo

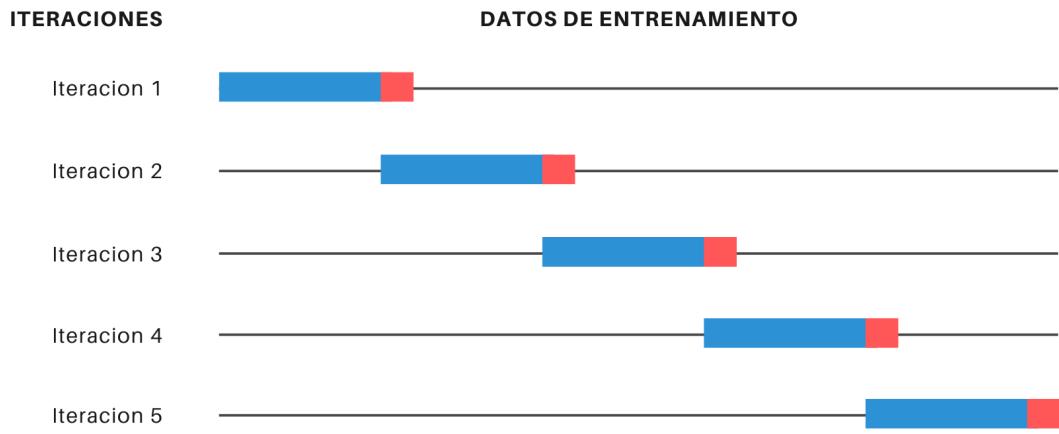


Fig. 3.3: Funcionamiento de *cross-validation* para series de tiempo.

```

predicciones ← []
datos_reales ← []
for cutoff in cutoffs do
    fecha_inicio ← cutoff
    fecha_fin ← cutoff + horizonte
    inicio_fold ← fecha_inicio - longitud_fold
    datos_entrenamiento ← precios[inicio_fold...fecha_inicio]
    datos_validacion ← precios[fecha_inicio...fecha_fin]
    predicciones.agregar(modelo.predecir(fecha_inicio...fecha_fin))
    datos_reales.agregar(datos_validacion)
end for
calcular_rmse(predicciones, datos_reales)

```

Fig. 3.4: Pseudocódigo de cómo calcular el RMSE para un parámetro en cross-validation

En la figura 3.3 tenemos una representación visual de cómo funciona este proceso, con los datos de entrenamiento para cada *fold* en azul y los datos que se utilizan para testear cada *fold* en rojo. Mientras que en la figura 3.4 podemos ver un pseudo código de cómo funciona *cross-validation* para un parámetro en particular.

Para nuestra búsqueda de hiperparámetros, utilizamos todos los precios del mercado del año 2021. Para cada combinación de parámetros, generamos 22 *folds*. El primero a los 40 días del año, o sea el 9 de febrero. Y luego un *fold* nuevo cada 15 días hasta terminar el año.

De esta forma, luego de computar cada *fold*, obtenemos para cada uno, un vector de 24 predicciones, uno para cada hora del día a predecir en ese *fold*.

Luego de tener las predicciones de todos los *folds* podemos calcular el puntaje del parámetro en el que estemos. Formamos un solo vector (que tendrá  $24 * 22$  predicciones) y calculamos el *RMSE* este vector contra el que tiene los datos reales del mercado.

Una vez hecho esto con cada parámetro, podemos quedarnos con el que obtenga un *RMSE* más bajo.

Luego de hacer la búsqueda de parámetros usando *cross-validation*, la combinación que obtuvo un RMSE más bajo fue la siguiente:

Parámetro	Valor
<i>changepoint_prior_scale</i>	0.005
<i>seasonality_prior_scale</i>	0.1
<i>changepoint_range</i>	0.99
<i>weekly_seasonality</i>	1
<i>yearly_seasonality</i>	3
<i>daily_seasonality</i>	20

Esta combinación de parámetros obtuvo un RMSE de 20.8714. La diferencia entre los distintos parámetros fue considerable. El conjunto de parámetros con peor resultado obtuvo un RMSE de 425.5620 y la mediana es de 127.12. Mientras que el conjunto de parámetros por defecto que usa *Prophet*, fue de 50.2439.

Algunas primeras impresiones del resultado final de este *cross-validation*, son que el mejor modelo que encontramos es un modelo que utiliza pocos *changepoints* (*changepoint\_prior\_scale* es muy bajo) y que le da mucha importancia a lo ocurrido en la recta final de los datos que le brindamos (*changepoint\_range*) es muy alto. También podemos observar que el modelo final le da poca importancia a la estacionalidad, debido al valor bajo que vemos en el parámetro *seasonality\_prior\_scale*.

Una vez obtenidos los mejores parámetros, procedimos a hacer una **comparación entre el modelo inicial de *Prophet* y un nuevo modelo** que utilice estos parámetros. Para comparar los modelos, hicimos pruebas, como dijimos en la sección 3.3, desde el 1 de abril de 2022 hasta el 27 de septiembre del mismo año (180 días).

En cada jornada utilizamos ambos predictores para determinar los precios de ese día. Luego calculamos el RMSE de ambos predictores al compararlo con los datos reales, y damos como **ganador del día al predictor que tenga un RMSE que sea menor a los otros al menos por un 5 %**. Tomamos esta decisión para descartar empates y sólo quedarnos con días donde un predictor supera al otro por una diferencia significativa.

Los resultados de comparar ambos modelos fueron los siguientes:

Parámetros	Días ganados	Media	Mediana
default	43	36.5442	21.7042
cross-validation	61	35.9327	22.2958

Fig. 3.5: Resultados de comprarar un modelo default de *Prophet* contra el modelo resultante de *cross-validation*.

Podemos observar que no hubo una gran diferencia entre ambos modelos. Una gran parte de los días resultaron en empate entre los mismos.

Sin embargo, el modelo que utiliza los parámetros obtenidos mediante *cross-validation* resultó algo mejor que el modelo que los parámetros por defecto de *Prophet*, teniendo

una mediana menor y ganando en 61 días de los 180, mientras que el predictor por defecto solo ganó en 43. Los 73 días restantes fueron empates, donde la diferencia entre ambos modelos fue menor al 5%.

Por este motivo, decidimos utilizar los parámetros obtenidos a partir de *cross-validation* para nuestro predictor.

### 3.3.4. Cambios en la cantidad de días

La cantidad de datos de entrenamiento que le damos al modelo es en sí una variable que afecta esencialmente al modelo y con la que decidimos experimentar.

Investigamos qué ocurre al darle una mayor cantidad de datos al modelo, para que tenga mayor conocimiento histórico de los precios. También observamos qué efecto ocasiona darle una menor cantidad de datos, para que le de mayor importancia al pasado inmediato y no tanto a lo histórico.

Nuestra primera prueba fue usando 180 días de datos de entrenamiento. Si queríamos predecir los precios del primero de julio del 2022, le brindábamos al modelo los datos desde el primero de enero hasta el 31 de mayo de ese año.

En este experimento comparamos **tres modelos diferentes**. Uno que recibe **180 días de datos**, cómo veníamos haciendo, otro modelo que recibe sólo **90 días**, y el último **365 días**.

Nuevamente hicimos la comparación probando ambos modelos durante 180 días a partir del 1 de abril del 2022. Recordemos que para los tres modelos ya usamos los parámetros obtenidos mediante *cross-validation* en la etapa anterior, y no los parámetros default que nos daba *Prophet*.

Cantidad de datos	Días ganados	Media	Mediana
90 días	106	33.63	20.53
180 días	41	35.93	22.30

Fig. 3.6: Resultados de comparar un modelo usando 180 días de datos vs 90 días.

Observando los resultados en la imagen 3.6, podemos notar que **disminuir la cantidad de días a 90 días mejora el predictor** con respecto a usar 180 días de datos.

De 180 días usados para la comparación, el predictor con menor cantidad de datos de entrenamiento ganó en 106. El predictor con 6 meses de datos ganó solo en 41 días.

El predictor con 90 días de datos tuvo un RMSE con un promedio de 33.63 y una mediana de 20.53, mientras que el predictor con 6 meses de datos tuvo un promedio de 35.93 y una mediana de 22.30.

Cantidad de datos	Días ganados	Media	Mediana
365 días	61	35.842	22.02
180 días	48	35.93	22.30

Fig. 3.7: Resultados de comprarar un modelo usando 365 días de datos vs 180 días.

Al ver la imagen 3.7, vemos que si bien **aumentar la cantidad mejora un poco el predictor, no lo mejora tanto** como disminuirla. La cantidad de días ganados es algo mayor para el modelo que utiliza más datos. La media y la mediana son bastante similares.

Luego de ver estos resultados, quisimos ver qué pasaba al disminuir aún más la cantidad de datos utilizados, por lo que **hicimos una prueba utilizando únicamente los últimos 30 días**.

Cantidad de datos	Días ganados	Media	Mediana
30 días	117	32.1529	16.5778
90 días	44	33.6283	20.5302

Fig. 3.8: Resultados de comprarar un modelo usando 90 días de datos vs 30 días.

Los resultados mostraron que bajar los datos a los **últimos 30 días bajó la media y la mediana** con respecto a los predictores anteriores. También la cantidad de días donde el predictor gana es de casi el 80 %.

Al ver estos resultados, **pensamos en seguir bajando** la cantidad de días y **usar sólo los últimos 15 días**. Comparamos el modelo resultante con el modelo que usa 30 días de datos y los resultados fueron los siguientes:

Cantidad de datos	Días ganados	Media	Mediana
15 días	81	32.2081	16.6659
30 días	74	32.1529	16.5778

Fig. 3.9: Resultados de comprarar un modelo usando 30 días de datos vs 15 días.

Los resultados muestran que al bajar la cantidad de datos a los últimos 15 días el modelo gana en algunos días más, pero la media y la mediana se mantuvieron muy similares e incluso subiendo un poco, por lo que **decidimos dejar la cantidad de datos en los últimos 30 días**.

Luego de esta experimentación decidimos que nuestro predictor de precios utilice los datos de los últimos 30 días. Entendemos que el hecho de que reducir la cantidad de días mejore el rendimiento del predictor, es coincidente con que los precios del mercado DAM cambian de forma bastante suave. Los precios no tienen saltos abruptos constantemente sino que son cambios paulatinos que se dan con el correr de varios días. Por esto, los días más cercanos al día que queramos predecir, son los que más información le aportan al modelo.

### 3.3.5. Temperatura

Sabemos que la **temperatura afecta de forma bastante directa al uso de energía** en la sociedad. Ya abordamos el tema en el capítulo anterior donde mostramos que los precios aumentaban en meses donde las temperaturas eran más elevadas.

Como también vimos en el capítulo anterior, contamos con los datos históricos de la temperatura en Lancaster, California.

Sabiendo esto, nuestro siguiente experimento fue ver que ocurre al **agregar esta información a nuestro modelo** predictivo.

Una ventaja que nos brinda *Prophet*, es la posibilidad de usar información extra para enriquecer los datos de entrenamiento. En particular, nos pareció interesante agregar información de la temperatura a nuestro modelo.

Hicimos la misma prueba que antes. Probamos el modelo que tenemos hasta ahora (usando 30 días de entrenamiento y los parámetros de *cross-validation*) contra el mismo

modelo pero agregándole datos de la temperatura en cada hora de cada día.

Predictor	Días ganados	Media	Mediana
Sin datos de la temperatura	59	32.1529	16.5778
Con datos de la temperatura	102	29.4738	14.6628

Fig. 3.10: Resultados obtenidos al agregar datos de la temperatura al predictor.

En los resultados podemos observar que **agregar datos de la temperatura mejora el modelo**. El predictor que cuenta con los datos de la temperatura fue superior en casi el 60 % de los días. La media y mediana también bajan considerablemente.

### 3.3.6. Agregar feriados

Como dijimos, Prophet tiene la posibilidad de agregar feriados al modelo.

Sin embargo, **al utilizar solo 30 días de datos, no creemos que tener los feriados en cuenta sirva de mucho al modelo**. Para que esta información fuese útil, el modelo necesitaría tener datos de al menos un año para ver como estas fechas repercutieron en los precios.

Por esto decidimos no explorar esta posibilidad.

### 3.3.7. Estacionalidad horaria

Por defecto, Prophet nos brinda estacionalidad anual, semanal y diaria, por lo que sabemos que nos sirve para detectar patrones en distintos años, semanas y días. La configuración de estas estacionalidades ya fue abordada en la sección 3.3.3, donde hicimos una búsqueda de hiperparámetros de la herramienta.

Sin embargo, **Prophet no nos da una estacionalidad mensual ni horaria**. Como estamos utilizando solo los últimos 30 días de datos, no vamos a agregar una estacionalidad mensual y **sólo nos concentraremos en la horaria**.

Somos conscientes de que los precios del mercado tienen un comportamiento muy definido según la hora en la que nos situamos (recordar la *duck curve* mencionada en la sección 2.3). Por lo que nos pareció que agregar una estacionalidad por hora podría ser beneficioso para el modelo.

**Prophet nos da la posibilidad de agregar estacionalidades nuevas**, especificando su duración y qué tan significativa uno quiere que sea la misma. Esto se hace a partir de dos parámetros, el *periodo* y el *orden*.

La duración del período se especifica diciendo cuantos días dura la estación. Si hablamos de estacionalidad horaria, queremos un período de 1/24. Y el orden se utiliza de la misma forma que el parámetro *seasonality\_prior\_scale*.

Para encontrar el mejor parámetro de orden para la estacionalidad horaria, utilizamos nuevamente la técnica de *cross-validation*. Le pedimos que busque en un rango de [1, 5, 10, 20]. El resultado fue que el mejor orden para la estacionalidad horaria es de 1.

Luego de esto comparamos al modelo habiéndole agregado esta estacionalidad contra el modelo anterior. Desde el primero de abril de 2022 y por 180 días. Obteniendo los siguientes resultados:

Predictor	Días ganados	Media	Mediana
Sin estacionalidad horaria	89	29.4738	14.6628
Con estacionalidad horaria	70	29.4608	15.3795

Fig. 3.11: Resultados de comparar nuestro modelo actual contra el mismo modelo agregando estacionalidad horaria.

Los resultados nos muestran que agregar estacionalidad horaria no benefició nuestro modelo. De hecho obtuvimos resultados peores que en el modelo anterior. Por lo tanto no utilizaremos estacionalidad horaria en el resto del trabajo.

### 3.3.8. Modelo final

Analizando los resultados obtenidos en secciones anteriores, llegamos a la conclusión de que el mejor modelo que obtuvimos utilizando *Prophet* es el modelo que:

- Utiliza los parámetros obtenidos por *cross-validation*
- Utiliza datos de entrenamiento de los últimos 30 días
- Utiliza los datos de la temperatura de cada día.

Con estas mejoras, logramos que el predictor básico de *Prophet* pase de tener resultados con una media de error de 36.5442, con una mediana de 21.7042, a tener una media de 29.4738 con una mediana de 14.6628.

## 3.4. Experimentación y resultados

Una vez obtenido el modelo final a utilizar, pasamos a evaluarlo contra otras alternativas.

Una alternativa que elegimos es **usar precios del día anterior** y suponer que los precios van a ser los mismos al día siguiente. Sabemos que utilizar los datos del día anterior es una buena heurística ya que los precios del mercado generalmente no varían tanto de un día a otro, sino que los cambios se dan más paulatinamente a lo largo de varios días. Puede pasar que de un día a otro cambien mucho los precios por algún motivo, como alguna falla en el suministro de energía, pero no es la regla general.

La otra alternativa que probamos, es **suponer que el precio va a ser el promedio de los precios de las últimas 4 semanas**. Por ejemplo, si estamos en un día lunes, vamos a usar el precio promedio de los últimos 4 lunes. Creímos que este modelo podría dar buenos resultados en caso de que los precios tengan un comportamiento distinto según el día de la semana en que uno se encuentre. Por ejemplo, la energía que se consume un día de semana no es la misma que se consume un fin de semana. Esto podría afectar a los precios.

Los experimentos consistieron en **comparar el predictor construido durante esta sección con estas dos alternativas** en distintos días. Para esto, nuevamente utilizamos el error cuadrático medio como medida para comparar los modelos.

Inicialmente comparamos los 3 modelos en días particulares.

Sabemos que el modelo que utiliza los datos del día anterior va a ser mejor cuando los precios no cambien mucho de un día al otro. A continuación mostramos un ejemplo de esto:

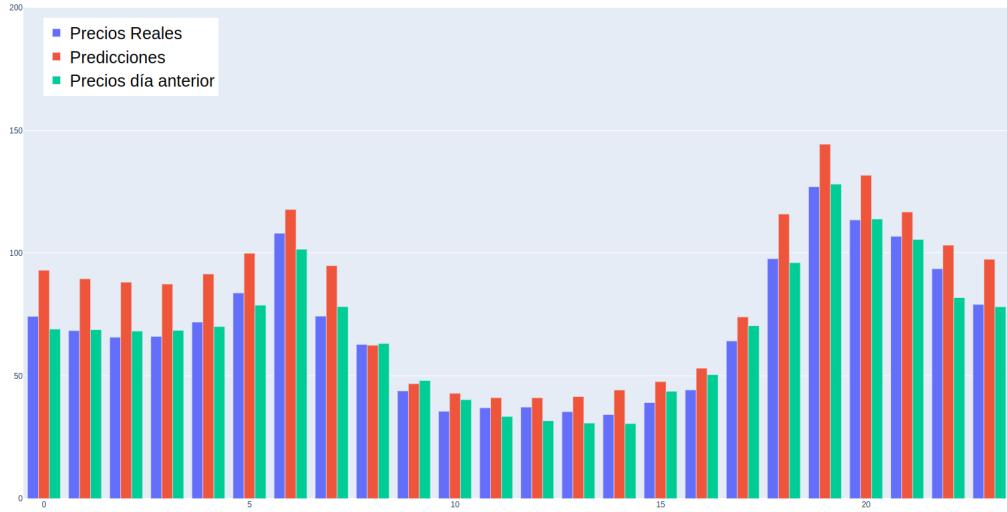


Fig. 3.12: Comparación de nuestro modelo para un día donde los precios del día anterior son muy similares al día siguiente.

En este gráfico tomamos nuestro predictor al querer predecir los precios del 20 de abril del 2022 y lo comparamos contra las otras alternativas. Podemos ver que los precios del día anterior son muy similares a los precios reales.

Sin embargo hay días donde los precios cambian muy abruptamente de un día a otro, en esos casos nuestro modelo se adapta mejor. Por ejemplo en el siguiente gráfico:

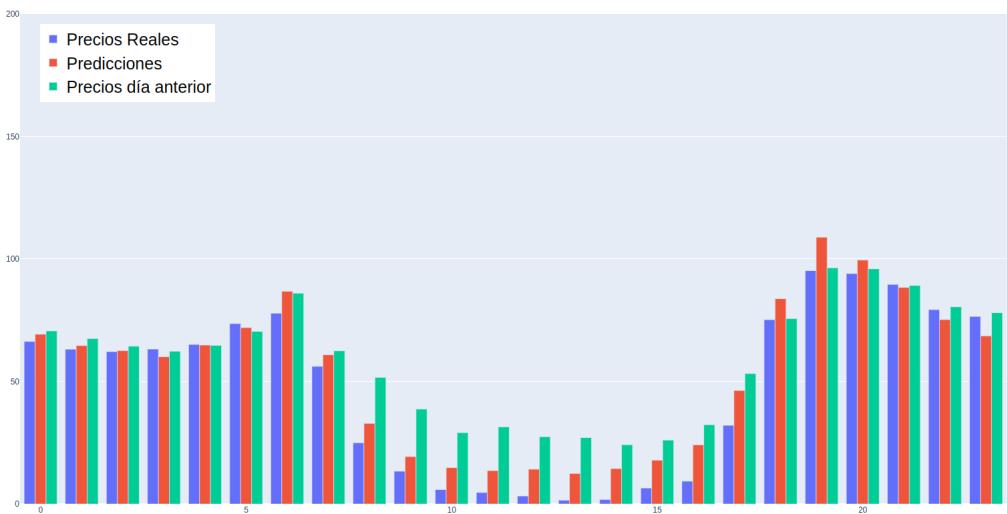


Fig. 3.13: Comparación de nuestro modelo para un día donde los precios del día anterior cambian mucho al día siguiente.

En este caso podemos ver que el día anterior los precios fueron bastante más altos, y en

un día bajaron drásticamente. En este caso, nuestro predictor obtiene mejores resultados que la alternativa de usar los precios del día anterior.

Vale la pena destacar que **en ambos casos**, por más que a un modelo le haya ido peor y a otro mejor, los tres **son bastante efectivos al calcular la curva general de los precios de cada día**.

#### Resultados generales

Para tener una visión más general, realizamos una experimentación similar a la llevada a cabo en secciones anteriores, que es comparar las predicciones de nuestro modelo contra usar los datos del día anterior o un promedio de las últimas 4 semanas para el primero de abril de 2022 y 180 días en adelante.

Primero hicimos una comparación del modelo de Prophet contra calcular el promedio de los días.

Predictor	Días ganados	Media	Mediana
Promedio ult. 4 semanas	60	34.2630	18.7161
Modelo de series de tiempo	112	29.4366	14.4859

Fig. 3.14: Resultados de comprarar un modelo de series de tiempo usando Prophet con utilizar el promedio de las últimas 4 semanas.

Los resultados nos muestran que **nuestro modelo es superior a la alternativa de utilizar el promedio de las últimas 4 semanas**. Gana en más del 60 % de los días y la media y mediana del error cuadrático medio son menores.

Luego hicimos la misma comparación entre nuestro predictor y la alternativa de utilizar los precios del día anterior, obteniendo los siguientes resultados:

Predictor	Días ganados	Media	Mediana
Día anterior	111	18.2243	10.8276
Modelo de series de tiempo	60	29.4366	14.4859

Fig. 3.15: Resultados de comprarar un modelo de series de tiempo usando Prophet con utilizar los precios del día anterior.

En este caso nuestro predictor tuvo peores resultados. Ganó en el 33 % de los días y tuvo una media y una mediana superiores.

Contrastamos nuestros resultados con ingenieros que trabajan en la planta de nuestro caso de estudio, quienes también construyeron un predictor basado en *Prophet* para predecir los precios. Ellos llegaron a la misma conclusión, utilizando los precios del día anterior como predicción, se obtiene un resultado que en general es más preciso que utilizando un modelo basado en *Prophet*.

Su análisis es que el mercado es muy volátil a lo largo del tiempo y a su vez no es muy cambiante en períodos cortos. La estacionalidad no influye mucho porque el mercado es muy cambiante a lo largo de los días. Lo que haya ocurrido en los últimos días es muchísimo más relevante que datos históricos que uno pueda llegar a tener. Factores como la inflación (que fue mucha en el 2022 en Estados Unidos), o cambios en cómo la gente usa la energía, afectan y hacen que el modelo de *Prophet* sea menos preciso.

### 3.5. Conclusión

Luego de esta experimentación pudimos comprobar que **logramos tener un modelo predictivo sin tener que meternos en tantos detalles teóricos**, y pudimos mejorar este modelo agregando y modificando distintos factores como cambiar la cantidad de días y agregar datos de la temperatura.

Lamentablemente **no se logró superar la alternativa de utilizar los precios del día anterior a la fecha**. Esto, como ya vimos en la sección anterior, se debe a que el mercado de energía es muy volátil a lo largo del tiempo y lo que haya pasado los últimos días tiene mucha más relevancia que los datos históricos.

Sin embargo, no se debe perder el objetivo final de este trabajo, que es generar una estrategia del uso de la energía para maximizar la ganancia en el mercado. En ese sentido, **lo más importante es que nuestro modelo detecte los momentos del día en el que la energía es más cara**, más que acertar el precio en sí. Lo importante es que el modelo pueda predecir bien la curva de precios, eso ya nos posibilitará tener una buena estrategia. Esto es logrado tanto por el modelo que usa los precios del día anterior como por el modelo implementado utilizando *Prophet*.

Además, una estrategia como **utilizar los precios del día anterior sería poco flexible si nuestras necesidades cambiaran**. Podría pasar, por ejemplo, que en vez de planificar la estrategia solo para el próximo día, queríamos planificar una estrategia para toda la semana. En ese caso, si usáramos la estrategia de usar los precios del día anterior, deberíamos suponer que esos precios se mantendrían durante toda la semana lo cual sería muy poco flexible ante cualquier posible cambio. En ese caso tener un modelo predictor sería más adecuado.

Más allá de la flexibilidad temporal, nuestro modelo también es flexible ante la aparición de nuevos factores que podrían tener importancia al predecir los precios. Ya vimos cómo pudimos agregar fácilmente datos de la temperatura al predictor, de esta misma manera se podrían agregar más datos que representen otras variables.

Otro aspecto a destacar es que por más que el RMSE promedio de nuestro predictor sea mayor al modelo que usa el día anterior, **nuestro predictor gana en un poco más de un tercio de los días en los que fue probado**. Recordemos que son días donde nuestro predictor obtuvo un RMSE un 5% menor, por lo que esa diferencia es significativa. En caso de tener un oráculo que nos dijera que días son más idóneos para nuestro predictor y cuáles lo son para el modelo de precios del día anterior, podríamos obtener una alternativa híbrida superior a utilizar este último modelo.

Por estos motivos, creemos que vale la pena tener un predictor que no sea tan rígido y que sea útil para distintos escenarios, y en el que **eventualmente se podría seguir trabajando para mejorar su precisión, lo cual no es posible con las otras alternativas**.

Más allá de esto, como en el escenario puntual de este trabajo lo que necesitamos es tener una predicción de los precios un día en adelante, **vamos a seguir el resto de la experimentación utilizando** la alternativa que mejor resultados nos dio, que es la de utilizar los **precios del día anterior**.

## 4. MODELO DE PROGRAMACIÓN LINEAL ENTERA

En los capítulos anteriores describimos las primeras dos etapas del trabajo. En la primera etapa, realizamos un análisis descriptivo de los datos históricos de precios de CAISO y de la información climática en Lancaster. En la segunda etapa, de análisis predictivo, utilizamos esta información para poder predecir los precios de los mercados de energía. En este capítulo describiremos cómo utilizamos estas predicciones para hacer un análisis prescriptivo. A partir de las predicciones de precios generada en la etapa dos, y de información extra como características de los paneles, de las baterías y predicciones climáticas que nos provee *Visual Crossing*, **diseñamos un modelo de programación lineal entera** cuya solución nos brinda una estrategia de venta de energía que maximiza nuestras ganancias.

### 4.1. Introducción sobre Programación Lineal Entera

Para atacar el problema de toma de decisión en la planta de producción de energía optamos por realizar un modelo de Programación Lineal Entera. El problema a resolver es un desafío de optimización (dado que queremos maximizar las ganancias), sujeto a una serie de restricciones operativas y de negocio de la planta. Si bien hay diversas técnicas para lidiar con este tipo de problemas, la Programación Lineal Entera resulta muy idónea por diversos motivos:

- En su versión *entera* permite la **simple modelización de toma de decisiones binarias** (es decir, hago esto o hago aquello).
- La forma de resolver el problema será **descriptiva y no procedural**. Es decir, indicaremos qué solución estamos buscando y no cómo la buscamos. Esto es muy útil para un problema real ya que es muy posible que el relevamiento preliminar del problema diste de lo que realmente se termine necesitando en la realidad (además de que la misma es cambiante), entonces un modelado con Programación Lineal Entera **permite ir adaptando la solución** sin la necesidad de introducir grandes cambios.
- En caso de toparse con instancias complejas del problema, existen muchos recursos teóricos y prácticos para poder lidiar con el mismo, tales como reformulaciones, estudio poliedral, etc. De todas maneras este punto no fue utilizado en este trabajo dado que al tener que resolver un día de producción los tiempos de ejecución no fueron un bloqueante.
- Existen diversas piezas de software complejas que nos permiten resolver instancias grandes de problemas modelados con Programación Lineal Entera con mucha rapidez.

En este trabajo solo nos centraremos en la formulación del modelo mediante Programación Lineal Entera y su posterior resolución directa mediante un *solver* de propósito general, por lo que no ahondaremos en otros conceptos teóricos y prácticos sobre esta técnica. En cualquier caso, en<sup>1</sup> se puede encontrar una muy buena bibliografía de introducción a esta técnica de resolución de problemas.

---

<sup>1</sup> Laurence Wolsey. *Integer Programming*. 1.<sup>a</sup> ed. Wiley, 20 de oct. de 2020. doi: [10.1002/9781119606475](https://doi.org/10.1002/9781119606475).

## 4.2. El modelo

En esta sección vamos a explicar el modelo de nuestro problema. Empezamos por nuestras variables de input, luego vamos a explicar las variables de decisión, la función objetivo a optimizar, las cotas y las restricciones de nuestras variables.

### 4.2.1. Discretización del tiempo

Antes de poder describir nuestro modelo, debemos aclarar cómo discretizar el tiempo del día. Para poder saber esto debemos tener en cuenta cómo funcionan los mercados. En el mercado DAM, las ofertas se realizan de a una por hora, mientras que en el RTM, se puede ofertar cada cinco minutos.

Por este motivo, podemos **discretizar el tiempo del día en intervalos de cinco minutos**. En cada intervalo podremos tomar decisiones sobre el mercado RTM, y cada 12 intervalos, sobre el mercado DAM. De esta manera **nuestro modelo puede dividir el tiempo del día en 288 intervalos**.

### 4.2.2. Input

Los valores de input se pueden resumir en:

- Las restricciones físicas de los paneles y las baterías.
- La potencia de energía de la que van a disponer los paneles.
- Los precios del mercado de energía.

El primer valor de input que tenemos en cuenta es la **potencia que van recibir los paneles en cada momento del día**. Al ser paneles solares, la potencia que estos tengan va a depender del clima, particularmente de la radiación solar que reciban. Para referirnos a esto en esta sección, podemos pensar en que tenemos una función que nos dice la potencia que van a recibir los paneles en un momento del día  $t$ . Llaremos a esa función *pot*:

$$pot : t \in [0, 287] \rightarrow \mathbb{R} \quad (4.1)$$

Como discretizamos el tiempo en intervalos de cinco minutos, si el problema se piensa en cómo utilizar los paneles a lo largo de un día, en total serían  $24 * 60 / 5$ , o sea 288 intervalos, por eso es que  $t$  va de 0 a 287.

El resultado de esta función será la potencia que recibirán los paneles y estará expresado en MW.

En el próximo capítulo explicaremos cómo, a partir del pronóstico del tiempo obtenido en *Visual Crossing*, se calculan los valores representados por la función *pot*.

Los siguientes parámetros de input constituyen las **características de las baterías**. Por un lado, tenemos un **tope para la potencia de carga y otro para la de descarga**. Estos son valores fijos, que llamaremos  $C_{top}$  y  $D_{top}$  respectivamente. Estas variables también estarán expresadas en MW.

Otros dos valores de input para las baterías serán por un lado la **cantidad máxima de energía que pueden almacenar** y por otro lado la **cantidad mínima de energía**

**que la batería siempre deberán mantener.** Llamemos a estas variables  $E_{max}$  y  $E_{min}$  respectivamente. Estas variables estarán expresadas en MWh.

También vamos a recibir el **estado inicial de las baterías**. Esto significa cuánta energía vamos a tener almacenada en el instante cero. Esta variable se llamará  $E_{ini}$ , también expresada en MWh.

La última variable de input que hace referencia a las baterías es la **cantidad máxima de intervalos de tiempo en los que se pueden estar cargando** a lo largo del día. Llamamos a esta variable  $K$ . Limitar esto es importante para que el modelo sea realista, ya que las baterías no pueden ser cargadas durante todo el día.

Y finalmente, el modelo debe tener la **información de los precios de los mercados** en cada momento. Esto también puede ser expresado en forma de función:

$$precio : t \in [0, 287], i \in \{DAM, RTM\} \rightarrow \mathbb{R} \quad (4.2)$$

Dijimos que el precio de la energía en los mercados es por USD/MWh. Sin embargo, nosotros estamos dividiendo el tiempo en intervalos de cinco minutos. Esto significa que para armar esta función tendremos que tomar los precios de los mercados y dividirlos por 12. Entonces, si el precio de un mercado es de \$120 en una hora, tendremos  $precio(t) = \$10$  para un  $t$  que este dentro de dicha hora.

#### 4.2.3. Variables de decisión

Las variables de decisión van a ser las que representen **qué vamos a hacer con los paneles y con las baterías en cada instante del tiempo**.

En cuanto a los paneles, el modelo debe decidir entre dirigir la potencia de carga a alguno de los dos mercados o a las baterías. Notar que esto no es exclusivo, la energía podría dividirse en partes y dirigirse a los mercados y a las baterías en el mismo intervalo. Más allá de que decidimos que no vamos a vender energía del panel al DAM, mantenemos variables de decisión para esto, para poder contemplarlo en el modelo.

Las variables de decisión serían las siguientes:

$$P_{i,t} \quad \forall t \in [0, 287], i \in \{BAT, DAM, RTM\} \quad (4.3)$$

Donde  $BAT$  representaría que la potencia de los paneles van dirigidas a la batería y  $RTM$  o  $DAM$  representaría que la energía es vendida a esos respectivos mercados.

Estas variables representan la potencia que uno dirige desde los paneles hacia la batería o hacia los mercados. Al igual que la función  $pot$ , estas variables están expresadas en MWh.

Para modelar el comportamiento de las baterías, necesitamos variables que representen la potencia de descarga de las mismas en cada intervalo de tiempo. Cuando las variables sean mayores a cero, significa que estamos vendiendo la energía de las baterías (o sea, descargando las baterías) al mercado correspondiente en ese momento. Las variables son:

$$B_{i,t} \quad \forall t \in [0, 287], i \in \{DAM, RTM\} \quad (4.4)$$

Al igual que las variables  $P$ , las variables  $B$  estarán expresadas en MW.

Finalmente, vamos a utilizar variables auxiliares que nos indiquen si las baterías están cargándose o no en un instante de tiempo determinado. Estas variables serán binarias y se llamarán  $O_t$ :

$$O_t = \begin{cases} 1 & \text{si las baterías están cargándose en el instante } t \\ 0 & \text{si no} \end{cases} \quad \forall t = 0 \dots 287 \quad (4.5)$$

En total el modelo cuenta con 1728 variables de decisión.

#### 4.2.4. Función objetivo

Teniendo ya las variables y el input, podemos plantear la función objetivo del modelo:

$$\text{Max} \sum_{t=0}^{288} \sum_{i \in \{RTM, DAM\}} \text{precio}_{it} * (P_{it} + B_{it}) \quad (4.6)$$

La función objetivo se puede pensar como la **suma de la cantidad de ingresos en cada intervalo de tiempo**. Los ingresos en un intervalo de tiempo es lo que se obtiene por vender al DAM sumado a lo que se obtiene por vender al RTM en cada dicho intervalo. Esto se obtiene en base al precio del mercado (valor que tenemos como input) y a la potencia enviada al mercado en ese momento, tanto por los paneles como por las baterías.

#### 4.2.5. Cotas de las variables

Las cotas que tiene este modelo serán para las variables  $P$ , y servirán para indicar que los paneles no pueden enviar a las baterías una potencia mayor al tope de carga de las mismas. Tendremos una cota por cada instante  $t$ :

$$0 \leq P_{BAT,t} \leq C_{top} \quad \forall t \in [0, 287] \quad (4.7)$$

#### 4.2.6. Restricciones

Una restricción que tenemos que modelar es la del **estado de carga de la batería**, el cual debe ser mayor o igual  $E_{min}$  y menor o igual a  $E_{max}$ . Para esta restricción definimos una variable auxiliar, que es el estado de carga de la batería en un momento determinado:

$$E_{min} \leq C_t \leq E_{max} \quad \forall t \in [0, 287] \quad (4.8)$$

Y definimos  $C_t$  de esta manera:

$$C_0 = E_{ini} \quad (4.9)$$

$$C_t = C_{t-1} + (P_{BAT,t} - \sum_{i \in \{DAM, RTM\}} B_{i,t}) * 300/3600 \quad \forall t \in [1, 287] \quad (4.10)$$

Esta restricción está planteada de forma recursiva, el estado de carga en el instante 0 es el estado inicial de la batería que recibimos como input. Luego para cada momento  $t$  el estado es igual al estado anterior sumado a la energía que el panel le transmite a las baterías menos la energía vendida al mercado. Como tanto  $P$  como  $B$  son medidas de potencia expresadas en MW, se debe multiplicar por la longitud del intervalo en segundos (que es de 300 seg) y dividir por 3600 para llegar a la cantidad de energía cargada en el intervalo expresada en MWh.

También vamos a restringir que **al final del ciclo, las baterías deberán tener la misma energía que al principio**. Queremos esto para poder garantizar que luego de un ciclo van a estar dadas las condiciones para resolver el mismo problema. La restricción sería:

$$C_{287} = E_{ini} \quad (4.11)$$

Ahora seguimos con las restricciones de los paneles. La primera restricción nos sirve para que la **energía de los paneles no sea utilizada para vender al mercado DAM**.

$$P_{DAM,t} = 0 \quad \forall t \in [0, 287] \quad (4.12)$$

Necesitamos esta restricción porque, por lo visto en el capítulo dos, cuando uno vende energía al mercado DAM se compromete a vender una potencia determinada durante una hora. La energía que reciben los paneles es variable, por lo que no es viable vender energía de los paneles al DAM ya que no podríamos asegurar vender una potencia fija durante un periodo prolongado de tiempo.

La siguiente restricción será para **representar que la potencia que los paneles pueden brindar en su conjunto está limitada por la cantidad de energía que están recibiendo** en el momento:

$$\sum_{i \in \{BAT, DAM, RTM\}} P_{i,t} \leq pot(t) \quad \forall t \in [0, 287] \quad (4.13)$$

También debemos restringir que la potencia en la que las baterías son descargadas para vender su energía a los mercados no puede ser mayor a su tope de descarga:

$$0 \leq \sum_{i \in \{DAM, RTM\}} B_{i,t} \leq D_{top} \quad \forall t \in [0, 287] \quad (4.14)$$

Debemos aclarar que la potencia con la que las baterías son descargadas puede ser regulada. No hay una potencia fija de descarga, sino que hay un máximo y uno puede moverse entre 0 y ese máximo de potencia. En caso de que esto no fuera así, lo que habría que hacer es restringir que la suma de la potencia de descarga de la batería en los distintos mercados sume siempre 0 o el máximo de descarga en cada momento  $t$ .

La siguiente restricción será para forzar que las variables binarias  $O_t$  sean 1 siempre que las baterías se estén cargando.

$$P_{BAT,t} \leq C_{top} * O_t \quad \forall t \in [0, 287] \quad (4.15)$$

De esta forma nos aseguramos que si  $P_{i,t}$  es mayor a cero, entonces  $O_t$  es mayor a cero.

Estas restricciones **servirán para regular la cantidad de cargas de la batería**.

Lo que queremos es limitar la cantidad de intervalos de tiempo en el que las baterías se cargan. Esto lo hacemos para prolongar la vida útil de las baterías, dado que no es recomendable que las mismas carguen y descarguen muchas veces.

Lo que vamos a restringir es que la batería no se cargue durante más intervalos que la cantidad definida como variable de input  $K$  a lo largo del día. La restricción quedará de esta forma:

$$\sum_{t \in [0 \dots 287]} O_t \leq K \quad (4.16)$$

Como  $O_t$  siempre va a ser igual a 1 cuando las baterías se estén cargando, entonces con estas restricciones no dejamos que las baterías se cargen más de  $K$  intervalos.

El último punto a tener en cuenta es lo siguiente: Como vimos al explicar los mercados, **cuando uno hace una oferta para el mercado DAM**, se compromete a enviar una misma potencia durante toda la hora para la cual se hizo la oferta.

Esto significa que si nosotros enviamos una determinada potencia durante un intervalo  $t$ , debemos enviar la misma potencia en el resto de los intervalos que entren dentro de la misma hora.

Asumiendo que nuestro ciclo va a empezar en el minuto cero de una determinada hora, podemos asumir que los primeros 12 intervalos conformarán una hora, y luego cada 12 intervalos pasamos a la hora siguiente. Por lo que nuestras últimas restricciones serían:

$$P_{DAM,t} + B_{DAM,t} = P_{DAM,t+1} + B_{DAM,t+1} \quad \forall t \in [0, 287], \lfloor t/12 \rfloor = \lfloor (t+1)/12 \rfloor \quad (4.17)$$

#### 4.2.7. Modelo final

Finalmente, el modelo queda de esta manera:

$$\text{Max} \sum_{t=0}^{288} \sum_{i \in \{RTM, DAM\}} precio_{it} * (P_{i,t} + B_{i,t})$$

s.a.

$$E_{min} \leq C_t \leq E_{max} \quad \forall t \in [0, 287]$$

$$C_0 = E_{ini}$$

$$C_t = C_{t-1} + (P_{BAT,t} - \sum_{i \in \{DAM, RTM\}} B_{i,t}) * 300/3600 \quad \forall t \in [1, 287]$$

$$C_{287} = E_{ini}$$

$$P_{DAM,t} = 0 \quad \forall t \in [0, 287]$$

$$\sum_{i \in \{BAT, DAM, RTM\}} P_{i,t} = pot(t) \quad \forall t \in [0, 287], i \in \{BAT, DAM, RTM\}$$

$$0 \leq \sum_{i \in \{DAM, RTM\}} B_{i,t} \leq D_{top} \quad \forall t \in [0, 287]$$

$$P_{BAT,t} \leq C_{top} * O_t \quad \forall t \in [0, 287]$$

$$\sum_{t \in [0 \dots 287]} O_t \leq K$$

$$P_{DAM,t} + B_{DAM,t} = P_{DAM,t+1} + B_{DAM,t+1} \quad \forall t \in [0, 287], \lfloor t/12 \rfloor = \lfloor (t+1)/12 \rfloor$$

$$P_{BAT,t} \leq C_{top} \quad \forall t \in [0, 287]$$

### 4.3. Implementación del modelo

Este modelo fue **implementado utilizando CPLEX**<sup>2</sup>, un software desarrollado por IBM que se utiliza para resolver problemas de programación matemática, específicamente problemas de optimización lineal, como el que nosotros queremos resolver. Es una de las herramientas más reconocidas y populares a la hora de resolver este tipo de problemas, ampliamente utilizado tanto en la academia como la industria.

La implementación de nuestro problema y su resolución fue **programada utilizando el lenguaje Python**, aprovechando la API que CPLEX provee para este lenguaje.<sup>3</sup>

---

<sup>2</sup> CPLEX User's Manual. 25 de nov. de 2021. URL: <https://www.ibm.com/docs/en/icos/20.1.0?topic=cplex-users-manual>.

<sup>3</sup> CPLEX for Python users. 25 de nov. de 2021. URL: <https://www.ibm.com/docs/en/icos/20.1.0?topic=apis-cplex-python-users>.

#### 4.4. Suite de tests para validar el modelo

Luego de tener nuestro modelo de Programación Lineal Entera ya programado, quisimos validararlo haciendo algunas pruebas sobre casos básicos y revisando que la estrategia brindada por el modelo sea realmente la óptima.

Estas **pruebas fueron fundamentales para tener confianza en nuestro modelo**, ya que ante cada cambio que le hacíamos, podíamos ejecutarlos y verificar que el comportamiento del mismo fuera correcto.

##### 4.4.1. Simplificando los escenarios

Los casos se diseñaron experimentando con distintos comportamientos en los precios del DAM y el RTM. **El propósito fue analizar casos muy simples, por eso también simplificamos el comportamiento de las baterías y los paneles**, el cual mantuvimos igual para todos los casos de *test* a menos que se diga lo contrario en el caso de prueba particular particular.

Las baterías cuentan con un límite de 48 intervalos de carga, lo que indica que su tiempo máximo de carga es de cuatro horas. Tienen una potencia de carga y descarga de 7.5 MW. Con una capacidad total de 15 MWh, las baterías pueden ser descargadas en su totalidad en un período de dos horas.

La potencia absorbida por los paneles fue pensada de manera tal que sea suficiente para que, en su máxima potencia, sea posible cargar las baterías en un lapso de 4 horas. Es decir, se debe generar 15 MWh en dicho tiempo, estableciendo así una potencia máxima de 3.75 MW.

Desde las 6hs. hasta las 17hs. las baterías se cargan a una potencia de entre el 80% y el 100% de su capacidad, la idea fue simular un día soleado. Luego de las 17hs. hasta las 18hs. las baterías comienzan a cargarse al 60%, suponiendo que el sol comienza a bajar. Y de las 19hs en adelante, al igual que desde las 0hs. a las 5hs. las baterías no se cargan debido a la falta de sol.

##### 4.4.2. Precios siempre son cero excepto a las 18hs y 19hs.

El primer caso a probar es muy simple. Los precios son siempre cero para ambos mercados a toda hora, excepto a las 18hs y a las 19hs. donde el precio del RTM es de 5000 USD/MWh.

En este caso resulta muy evidente que la mejor estrategia es tener las baterías cargadas en su totalidad para las 18hs. y vender toda la energía de las baterías y la de los paneles en las siguientes dos horas.

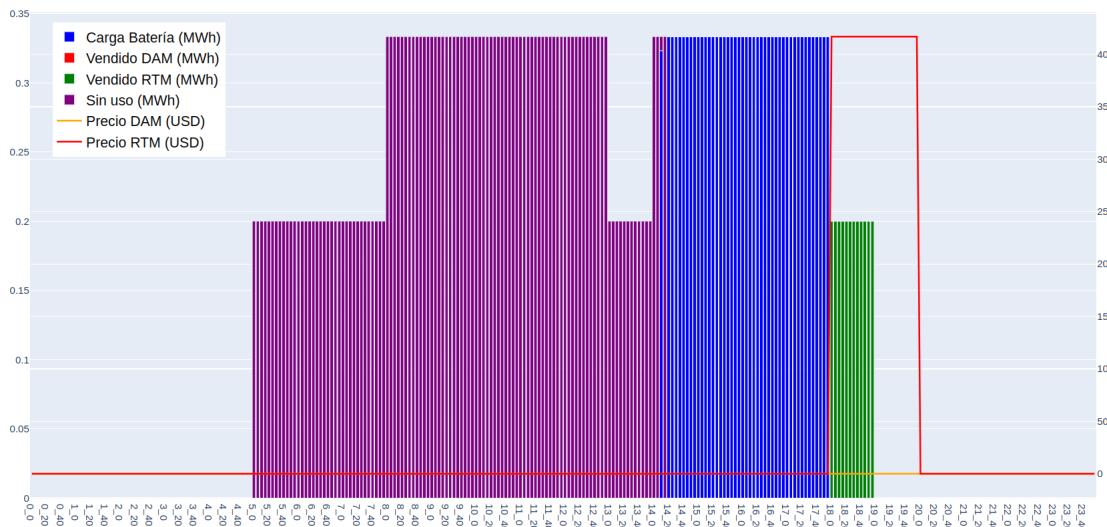
Para poder visualizar cuál fue la estrategia resultante del modelo, hicimos un gráfico que muestre la actividad de la batería que la estrategia prescribió.

En el gráfico de actividad de las baterías **podemos observar por un lado, la energía que tiene almacenada la batería en cada momento del día**. Mostramos esto mediante un gráfico de líneas usando el eje que se encuentra a la derecha como referencia. Por otro lado **podemos ver en forma de un gráfico de barras cuánta energía vende la batería a cada mercado en cada momento del día**. Para interpretar esto debemos usar el eje que se encuentra a la izquierda.

En el gráfico de actividad de paneles, podemos ver para qué se utilizaron los paneles en cada momento del día. Los paneles pueden estar siendo:



(a) Actividad de la batería cuando el precio es cero excepto a las 18hs.



(b) Actividad de los paneles cuando el precio es cero excepto a las 18hs y a las 19hs.

- Inutilizados.
- Utilizados para cargar la batería.
- Utilizados para vender al DAM.
- Utilizados para vender al RTM.

El gráfico nos muestra que la batería no se carga hasta las 14:10hs. luego se carga por completo las siguientes 4hs. A las 18hs. se vende toda la energía hasta agotarse a las 20hs. Las barras nos muestran que la energía fue vendida al mercado RTM.

Por otro lado el gráfico de actividad de paneles nos muestra que los mismos estuvieron inutilizados por completo hasta las 14hs (los cambios en la potencia inutilizada son porque

en algunas horas hay menos sol). Luego, de 14.10hs. hasta las 18hs. se usan al 100 % para cargar las baterías. A las 18hs, la energía que llega, se usa para vender al mercado RTM. Notar que a las 18hs. ya llega menos energía a los paneles, porque el sol baja, y de las 19hs. en adelante ya no llega energía.

En este caso, verificamos que la estrategia brindada por nuestro modelo fue la correcta. Ahora podemos ver casos un poco más complejos.

#### 4.4.3. Precios siempre son cero excepto a las 18hs y a las 12hs.

En este caso, el precio es siempre cero excepto en el mercado DAM de 12hs a 14hs. y de 18hs a 20hs.

Si este fuera el único cambio, la estrategia debería ser claramente descargar la batería en alguna de esas dos horas. Sería el mismo comportamiento que en la prueba anterior, pero podría ser en dos momentos del día.

Para que la prueba sea un poco más enriquecedora, **hicimos que en este caso la batería pueda cargarse durante más tiempo**. Le dimos un máximo de tiempo de carga de 72 intervalos, lo cual serían **6 horas**. De esta forma la batería puede cargarse un ciclo y un poco más.

Además hicimos que **el precio de las 18hs. y 19hs. sea algo superior** al de las 12hs. y 13hs. Los precios serán de 6000 USD/MWh y 4500 USD/MWh respectivamente.

Lo que esperamos de la estrategia es que tenga las baterías al 100 % para las 18hs. para descargarlas en su totalidad, y que use esas dos horas extra para cargarla antes de las 12hs. y vender la energía en esas dos horas.

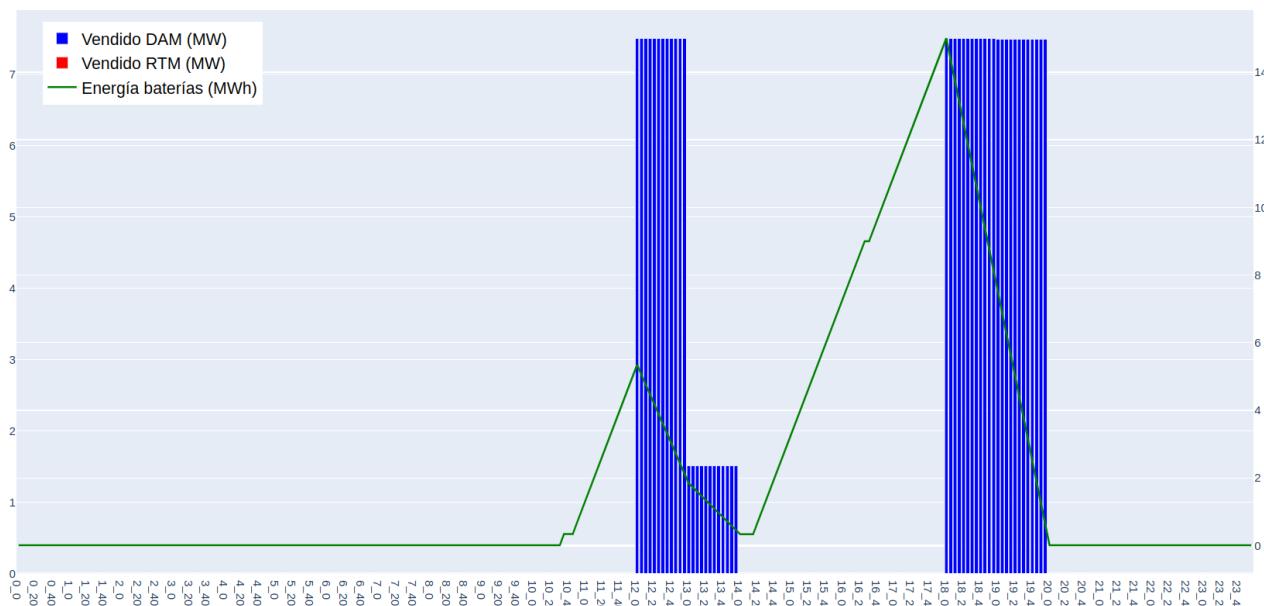


Fig. 4.2: Actividad de la batería cuando el precio es cero excepto a las 18-20hs y a las 12-14hs.

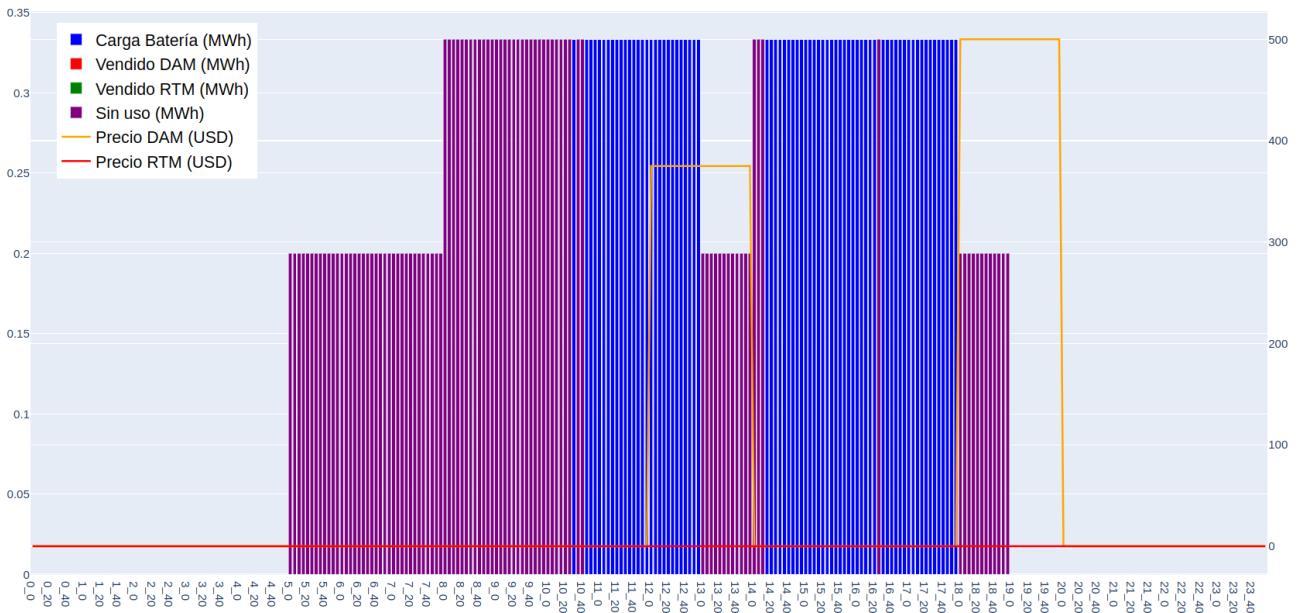


Fig. 4.3: Actividad de los paneles con precio cero excepto a las 18-20hs. y a las 12-14hs.

En el gráfico de actividad de las baterías, podemos ver que efectivamente las mismas son cargadas parcialmente para vender su energía a las 12hs. Como no puede vender el máximo de 7.5 MWh durante las dos horas, vende 7.5 MWh en la primera hora, y 1.51 MWh en la segunda. Luego, desde las 14hs en adelante, las baterías se cargan en su totalidad para ser vendidas a las 18hs. donde el precio es mayor.

En el gráfico de actividad de los paneles podemos ver que el tiempo de inutilización fue algo menor esta vez, ya que se pudo usar más energía para cargar la batería, un poco antes de las 12hs. y luego antes de las 18hs. De las 12hs a las 14hs. Notar que al tratarse del mercado DAM, la energía de los paneles no puede ser vendida a los mercados, por lo que solo es usada para cargar las baterías.

#### 4.4.4. Precios de un día real.

En el siguiente y último caso, queremos probar qué ocurre cuando usamos los precios de los mercados de un día real. En este caso seguiríamos simplificando el comportamiento de la batería y los paneles, pero ya tendríamos un comportamiento de los precios realista y no uno tan simple como los que venimos probando.

Vamos a probar utilizando los precios del mercado de energía del 25 de julio del 2022.

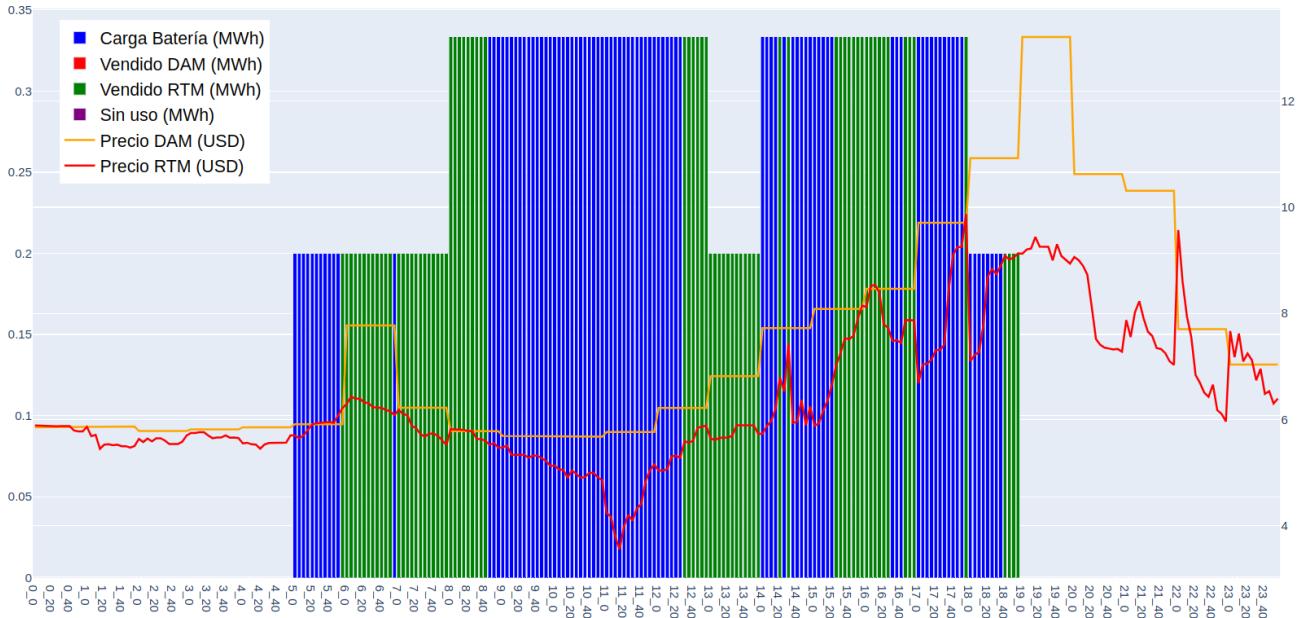


Fig. 4.4: Actividad de los paneles utilizando los precios del 25 de mayo del 2022.

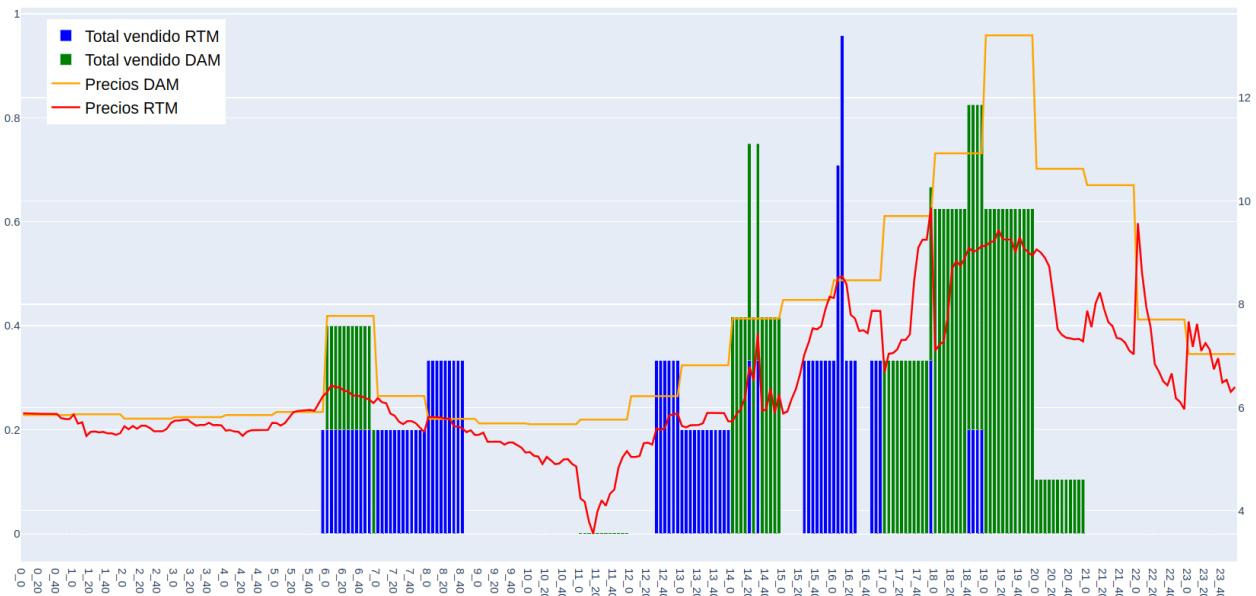


Fig. 4.5: Energía total vendida utilizando los precios del 25 de mayo del 2022.

Lo primero que debemos observar son los precios de la energía. Podemos ver esto en la figura tanto en la figura 4.4 como en la 4.5. Vemos que el DAM tiene un leve pico a las 6hs. seguido una meseta en la mañana y el mediodía para luego aumentar y llegar a un pico a las 19hs. El mercado RTM en general se sitúa por debajo del DAM en cuanto a los precios.

Ahora podemos analizar cuál es la estrategia que el modelo nos prescribe.

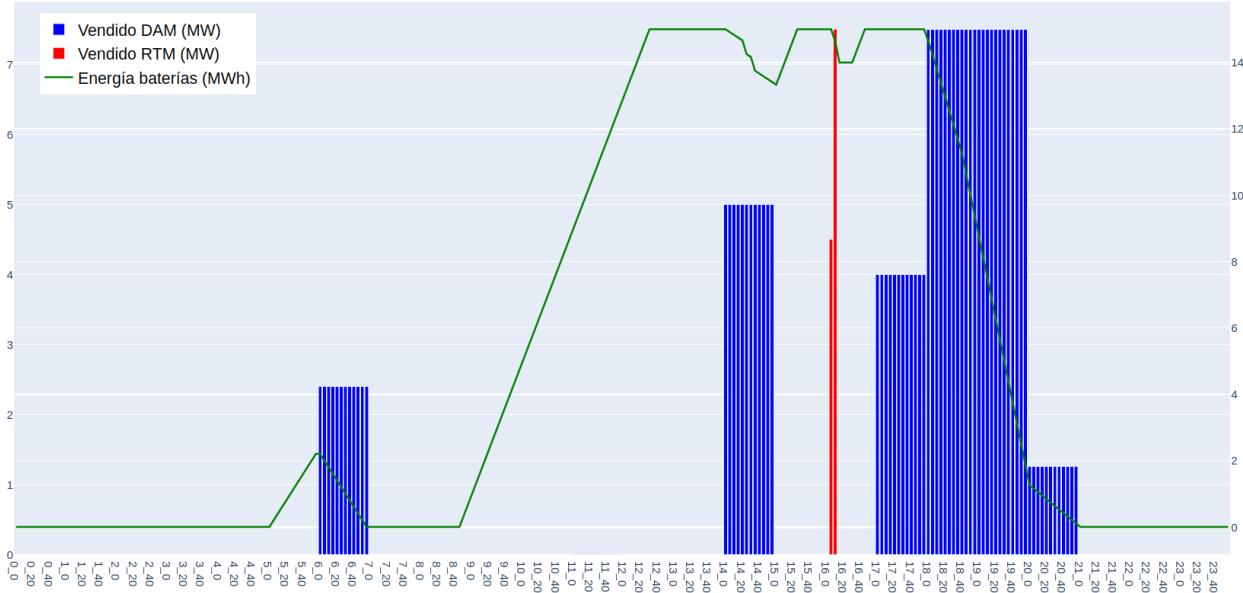


Fig. 4.6: Actividad de la batería utilizando los precios del 25 de mayo del 2022.

En cuanto a los paneles, podemos ver que la estrategia se divide entre cargar la batería y vender al RTM. Los momentos donde los paneles cargan la batería coinciden con los momentos donde los precios son más bajos, por ejemplo de las 9hs a las 12:30hs. De esta manera se aprovecha esta meseta en los precios para vender la energía de los paneles cuando el precio es mayor.

La batería es usada para vender su energía a la máxima potencia a las 18 hs. y las 19 hs. momento donde los precios del DAM son más altos. Y también aprovecha el pico que vimos a las 6 hs. para vender algo de su energía en ese momento.

Además se vende una mínima parte de la energía de las baterías entre las 16 y las 16:10. Momento donde el RTM se incrementa levemente.

#### 4.5. Cierre del capítulo

En este capítulo **mostramos el modelo de programación lineal entera generando**, comentando cómo fueron pensadas las variables de decisión, la función objetivo y las restricciones del problema. Posteriormente, el **modelo fue implementado en CPLEX**.

Una vez implementado, el **modelo fue probado en casos simples** para revisar su comportamiento. Mediante esas pruebas pudimos ganar confianza tanto en el modelo como su implementación. Esta confianza nos permitió poder hacer pruebas en el modelo, verificando que su comportamiento sea correcto ante cada cambio.

En el próximo capítulo, utilizaremos este modelo generado junto a nuestro predictor de precios descrito en el capítulo 3 para generar un sistema de decisiones y hacer pruebas sobre el mismo utilizando escenarios reales.

## 5. INTEGRACIÓN Y EXPERIMENTACIÓN

### 5.1. Objetivo

En este trabajo propusimos una forma de brindar una estrategia de compra y venta de energía en el mercado de California utilizando un modelo de programación lineal entera. Hicimos diversas pruebas tanto para la etapa de predicción de precios como para el modelo de programación lineal. Lo que resta es hacer pruebas integrales de todo el *pipeline*. Queremos **averiguar qué tan bueno es este modelo comparado con otras posibles estrategias**. Ese es el objetivo de esta etapa.

Para poder comparar, vamos a hacer **simulaciones de las distintas estrategias con datos reales de distintos días**. Usaremos datos de los precios del mercado obtenidos utilizando OASIS, para saber cuánta plata hubiésemos ganado con cada estrategia. Y también utilizaremos datos del clima para poder estimar cuanta energía podríamos haber vendido ese día. De esta manera, podemos comparar cómo le hubiera ido a cada estrategia en cada día.

### 5.2. Heurísticas a comparar

Una de las estrategias con la que vamos a comparar a nuestro modelo es **la utilizada en el parque de baterías que estamos usando como objeto de estudio** en este trabajo.

Esto significa que tendremos la posibilidad de probar nuestra solución contra una estrategia que está siendo realmente utilizada hoy en día por un participante del mercado de energía de California. Esto le da mayor valor al experimento ya que podremos saber si nuestra estrategia le hubiera servido a alguien que está efectivamente participando del mercado.

Lo que ellos hacen es **cargar las baterías durante el día** (generalmente al mediodía) utilizando los paneles. **A las 18hs, descargan las baterías en su totalidad**, vendiendo su energía al mercado DAM. La energía de los paneles es vendida a toda hora al RTM, **excepto cuando se están utilizando para cargar las baterías**. Esta estrategia se basa en que, como vimos anteriormente al mencionar la *duck curve*, la oferta de energía cae por la tarde cuando el sol empieza a descender y ya no se cuenta con la energía solar, provocando que los precios sean más altos. De ahora en adelante llamaremos a esta heurística “**H1**”.

Otra heurística similar que usamos es **cargar las baterías más de una vez en el día**. Vamos a mantener el hecho de vender la energía de las baterías en su totalidad a las 18hs, pero además vamos a **vender una fracción de la misma a las 10hs**. Estimamos que para las 10hs. podemos llegar a tener un cuarto de la batería cargada, por lo que vamos a vender al DAM la **totalidad de las baterías a las 18hs y además un cuarto de las baterías a las 10hs**. Al igual que en la anterior heurística, la energía de los paneles es vendida al RTM en todo momento donde no se estén cargando las baterías. De ahora en adelante llamaremos a esta heurística “**H2**” .

### 5.3. Simulaciones

Como comentamos anteriormente, para poder comparar estrategias debemos poder realizar simulaciones de las mismas a lo largo de un día. En esta sección explicaremos cómo realizamos dichas simulaciones.

Una estrategia puede verse como un **conjunto de decisiones** sobre qué hacer con nuestros paneles y nuestras baterías a lo largo del día. Como nuestro modelo de optimización lineal divide el tiempo en intervalos de cinco minutos para armar la estrategia, en nuestro caso **cada cinco minutos tomamos una decisión** indicada por la misma.

Cada decisión consiste en decir qué hacer con la energía recibida por los paneles y qué hacer con la energía disponible en las baterías.

Lo que vamos a hacer en la simulación es arrancar en un estado inicial, e ir evaluando intervalo por intervalo cada decisión que la estrategia nos diga y registrando en qué estado quedamos luego de cada decisión hasta llegar al final del día.

En la simulación ya tenemos toda la información sobre los precios y el clima del día a simular, por lo que podemos saber exactamente qué hubiera pasado en caso de utilizar una estrategia u otra.

Arrancamos en el intervalo 0, que representa las 00hs. de la fecha que estemos simulando, y terminamos en el intervalo 287, que representa el intervalo donde tomo la última decisión del día, a las 23:55hs. Luego de computar todas las decisiones para todos los intervalos del día, obtendremos cuáles hubieran sido las ganancias al ejecutar la estrategia en esa fecha.

#### Input de las simulaciones

La simulación recibe un *input*, constituido por la fecha que queremos simular, la estrategia a seguir, los precios de ese período, y la radiación solar que hay ese día (recordemos que a partir de este dato podemos calcular la potencia de los paneles). También tomamos el estado inicial de las baterías al iniciar la simulación.

t						Estrategia			
	Radiación solar	Nubosidad	Precio DAM	Precio RTM	Panel a batería	Panel a RTM	Batería a DAM	Batería a RTM	
105	408	0	17.75	7.45	0	0.09	0	0	
106	408	0	17.75	11.38	0	0.09	0	0	
107	408	0	17.75	6.16	0.45	0.45	0	0	
108	450	10	20.20	10.24	0	0.49	0.62	0	

Tab. 5.1: Ejemplo del input que reciben las simulaciones en distintos intervalos t

En la tabla 5.1 figura un ejemplo de 4 instantes distintos, que serían desde las 8:45 am a las 9:05 am. En el ejemplo vemos cómo para cada intervalo, tenemos información de los precios y clima. Además podemos ver cómo la estrategia nos dice en cada momento qué hacer con las baterías y los paneles (las columnas de estrategia están todas en MWh).

#### Factibilidad de las decisiones

Algo importante a tener en cuenta en la simulación, es que la misma es ejecutada utilizando los datos reales de los precios y el clima, mientras que las estrategias son pensadas

en base a estimaciones. Podría decirse que **la estrategia** (ya sea una heurística o el MIP), **proviene de un plano táctico, mientras que las simulaciones son ejecutadas en el mundo operativo**. En caso de las estimaciones sean muy diferentes a lo que termine pasando en la realidad, estas estrategias, pensadas para el plano táctico podrían **no ser factibles** en el aspecto operativo.

Por lo tanto, en caso de que la decisión sea factible, se seguirá al pie de la letra lo que dicte la estrategia. En caso de que la decisión no sea factible, se debe adaptar la estrategia en el momento. A continuación explicamos esto con más detalle.

### ¿Cómo se computa un intervalo?

Supongamos que estamos en el intervalo  $t$  de la simulación. Lo **primero** que hacemos es **ver cuánta potencia reciben los paneles** en dicho instante, esto será explicado más adelante pero podemos saberlo a partir de la radiación solar y la nubosidad que tuvimos en ese tiempo.

Luego de saber esto, vemos **qué tenemos que hacer con los paneles**. Esto estará determinado por la estrategia, que nos va a decir que carguemos una determinada cantidad de energía en las baterías, y que vendamos otra cantidad al RTM. Recordemos que los paneles no pueden vender al mercado DAM.

**Si la cantidad de energía recibida por el panel es mayor o igual a la prevista por la estrategia, hacemos lo que la misma nos diga** al pie de la letra. Si el panel recibiera más energía que la que predijo la estrategia, ese sobrante se envía al RTM siempre y cuando el precio sea mayor a cero. Si el precio del RTM fuera cero, la energía sobrante se envía a la batería.

En caso de que haya menos energía y no podamos cumplir con lo que la estrategia nos dice, la prioridad es **primero cumplir con la carga de las baterías, y luego vender el sobrante el RTM**. Hacemos esto para evitar multas por incumplir con el mercado DAM.

Una vez que sabemos qué se hizo con los paneles, pasamos a ver qué hacen las baterías. Lo primero que hacemos es **sumarle la energía que cargaron los paneles** a la energía que tenían las baterías en el instante  $t - 1$ . Luego de hacer esto, sabemos con qué energía cuenta la batería en este punto. Y en este momento vemos qué es lo que nos dice la estrategia que tenemos que hacer con las baterías.

La estrategia nos va a decir que vendamos una cantidad de energía al DAM y otra al RTM. Nuevamente, esto **podría no ser factible** según el contexto. Por ejemplo, si durante el día los paneles recibieron menos energía de la prevista, las baterías podrían no tener la energía necesaria para cumplir con la estrategia. Otra vez, en caso de que no sea factible, debemos adaptar la simulación.

En caso de que la energía alcance para todo, hacemos exactamente lo que la estrategia diga. Notar que en este caso, a diferencia de los paneles, **las baterías no van a tener energía sobrante** no prevista por la estrategia, ya que toda la energía que tienen las baterías son producto de cargas que la estrategia nos indicó anteriormente. Únicamente podría haber energía extra si en algún momento los paneles reciben energía de más y el precio del RTM es cero en el momento. Al ser muy poco probable, no contemplaremos ese caso en nuestras simulaciones.

Si ocurriese que la energía no alcance para todo, **se priorizará primero mandar energía al DAM, y luego al RTM**. En caso de no poder vender al DAM, también

debemos guardar esta información, ya que tendríamos una multa por incumplir con esto y debemos contemplar esa multa al calcular nuestros ingresos finales.

Al terminar esta parte de la simulación ya sabremos el estado de la batería luego de que el intervalo  $t$  haya finalizado. Y al saber cuanto se vendió al DAM y al RTM podemos saber cuanto dinero tendremos por estas ventas.

Recordar que para saber el ingreso total **tenemos que contemplar las multas**. En caso de no haber podido cumplir con que debíamos vender al DAM según la estrategia, tendremos que restar de nuestros ingresos la multa correspondiente, que es la energía que nos haya faltado cobrada al precio del RTM.

### Calculando las ganancias

Al finalizar la simulación para el último intervalo, tendremos la cantidad de dinero que habremos ganado ese día si hubieramos usado la estrategia elegida. Para resumir, la cuenta de ganancias sería:

$$\sum_{t=0}^{288} \left( \sum_{i \in \{RTM, DAM\}} \text{precio}_{it} * P_{it} + \text{precio}_{it} - \sum_{t=0}^{288} \text{precio}_{RTM,t} * P'_{DAM,t} \right) \quad (5.1)$$

Donde  $P$  sería una función que nos dice cuánta energía vendimos realmente a cada mercado, y  $P'$  nos dice cuanta energía fallamos en vender al mercado DAM.

## 5.4. ¿Cómo obtener simulaciones realistas?

Para que los resultados de las simulaciones sean útiles a la hora de hacer un análisis sobre nuestra solución en comparación con otras heurísticas, necesitamos poder ejecutar nuestras simulaciones a partir de un *input* que sea realista.

En este punto del trabajo ya contamos con información real de los precios del mercado y de las condiciones climáticas (temperatura, nubosidad y radiación solar) de Lancaster. Pero aún nos faltaba información.

### 5.4.1. Modelo de las baterías y paneles solares

Uno de los puntos faltantes para poder ejecutar simulaciones realistas es tener información de como funcionan las baterías en el parque que usamos como caso de estudio. A continuación detallamos esta información.

Las baterías tienen una **capacidad de carga de 15 MWh**. Y su potencia de **carga y descarga es de 7.5 MW**. Esto significa que si la batería es cargada a su máxima potencia, llegaría al máximo de su capacidad en dos horas. De la misma forma, la batería puede vaciarse en dos horas si se la descarga al máximo. En cuanto a la **carga inicial** de la batería, la misma será de **0.01 MWh**.

En cuanto a los paneles, su capacidad máxima de carga depende de la radiación solar que reciba y también de la nubosidad en cada momento. Este es un cálculo muy complejo porque en la realidad la potencia que recibe un panel solar no depende únicamente de estos dos factores sino que depende de algunos más, como por ejemplo el ángulo en el que está ubicado el panel con respecto al sol en cada momento del día.

Para no exceder el alcance de este trabajo decidimos simplificar este cálculo. Primero obtuvimos datos reales de la potencia generada con los paneles en distintos días en la planta que utilizamos como caso de estudio. De esta manera tuvimos una **referencia real de la potencia que reciben los paneles en la planta**. Una vez obtenida esta información, nuestro objetivo fue **obtener una función** que use los dos factores que sí disponemos, cuyos resultados sean **lo más similares posibles a los datos reales** obtenidos.

Como vimos en el capítulo 2, nosotros contamos con información de la radiación solar y la nubosidad en California. Esta información fue obtenida a través de la API de *Visual Crossing* y está desglosada en horas. Para la radiación solar tenemos la información en  $W/m^2$ , mientras que para la nubosidad es un porcentaje.

Nuestro primer acercamiento fue simplemente usar el valor de radiación solar obtenido y multiplicarlo por un factor de nubosidad.

Representamos el factor de nubosidad con la siguiente función, que recibe como parámetro un porcentaje de nubosidad:

$$factor\_nubosidad(nub) = \begin{cases} 1 & \text{si } 0 \leq nub < 10 \\ 0,8 & \text{si } 10 \leq nub < 20 \\ 0,65 & \text{si } 20 \leq nub < 30 \\ 0,3 & \text{si } 30 \leq nub < 40 \\ 0,2 & \text{si } 40 \leq nub < 50 \\ 0,1 & \text{si } 50 \leq nub < 60 \\ 0,05 & \text{si } 60 \leq nub \end{cases} \quad (5.2)$$

Lo que nos lleva a la siguiente formula:

$$pot(t) = factor\_nubosidad(nubosidad(t)) * radiacion\_solar(t) \quad (5.3)$$

Esto nos dio una curva con una forma similar a los datos que nos fueron proporcionados, pero con números bastante menores.

Para acercar los números a los números de nuestro caso de estudio, multiplicamos por una constante que igualara nuestra función con estos datos.

$$pot(t) = factor\_nubosidad(nubosidad(t)) * radiacion\_solar(t) * CTE \quad (5.4)$$

Luego de experimentar con distintas constantes que nos lleven a una curva similar a los datos reales proporcionados, llegamos al número 4,45.

## 5.5. El riesgo de trabajar con estimaciones

Las simulaciones nos permiten comparar nuestra estrategia contra cualquier otra en un período determinado. Sin embargo, más allá de saber si nuestro modelo es mejor para un día en particular, necesitamos poder saber si éste va a ser mejor que otras estrategias en general.

Hay dos factores que influyen mucho a la hora de saber qué tan bien le va a ir a un modelo: uno es la cantidad de energía que van a recibir los paneles solares a lo largo del día, y el otro son los precios del mercado de energía en ese día.

Parte de nuestro trabajo fue predecir estos dos factores. Luego de haber hecho esas predicciones, nosotros contamos con una estimación de lo que serán las condiciones climáticas y los precios del mercado de energía del día. En base a ese input, podemos construir nuestro modelo de programación lineal entera y llegar así a nuestra estrategia final.

Nuevamente, debemos recordar que estamos hablando de **estimaciones**, las cuales **no serán exactamente iguales a lo que se termine dando realmente**. Esto podría ser un problema, ya que nuestro modelo, que seguramente nos va a dar muy buenos resultados en caso de cumplirse nuestras estimaciones, podría darnos un **resultado no tan bueno en la realidad**.

Por ejemplo, si estimamos que a las 18 hs. de ese día el mercado del DAM tendrá un precio excesivamente alto, mucho mayor al resto de las horas, el modelo nos va a dar una estrategia que apunte a reservar la mayor cantidad de energía posible para las 18 hs. y venderla toda en ese momento. Luego, si llegado el día, el precio para las 18 hs. es menor al estimado, nuestro modelo no va a dar muchas ganancias. Esto es debido a que estaba basado en una condición muy específica de la estimación.

Esto se potencia en mercados como el RTM donde la varianza de los precios es más alta, la estimación podría decir que el precio de las 18:35 hs. es muy alto. Pero lo más probable es que luego esto no ocurra, por lo tanto si nuestro modelo le da mucha importancia a ese dato, es probable que termine no siendo tan bueno en la realidad.

## 5.6. Eligiendo el modelo más adaptable a cambios

Hablamos de qué podría pasar cuando nuestras estimaciones difieren mucho de lo que se da en la realidad. Por otro lado, sabemos que en caso de que nuestras estimaciones sean buenas, a pesar de diferir de la realidad, no van a ser muy diferentes. Por lo que vimos en la conclusión del capítulo 3, deberían ser bastante similares. Vimos que los precios se comportan siguiendo la *duck curve*, puede pasar que la estimación difiera un poco en la magnitud de los picos, pero **es probable que el comportamiento a grandes rasgos de los precios y la energía recibida en la realidad y en las estimaciones sean similares**.

Debido a esto pensamos que un aspecto a evaluar a la hora de medir si nuestro modelo es bueno o no, es **medir cómo se comporta al simular la estrategia generada usando condiciones levemente distintas** a las que se le pasaron como input.

Una estrategia que se comporta muy bien para las condiciones que se le pasaron como input, pero que baja mucho el rendimiento cuando estas condiciones cambian levemente, nos va a interesar menos que otro que se comporta de manera más consistente al modificar estas condiciones.

A su vez, no necesitamos que una estrategia se comporte bien cuando las condiciones cambian mucho, porque sabemos que nuestras estimaciones no van a ser muy distintas de la realidad.

## 5.7. Framework de simulaciones

Vimos que trabajar con estimaciones nos trae el riesgo de que nuestra estrategia no se comporte como esperamos al ejecutarla en la realidad. También vimos que para nuestro caso particular, nos interesa cómo el modelo se comporta ante cambios leves en las estimaciones, mientras que no nos importa tanto qué pasa con cambios más grandes.

Esto nos llevó a pensar en un mecanismo que nos posibilite elegir un modelo más consistente para nuestro caso particular. Este mecanismo consta de tres etapas.

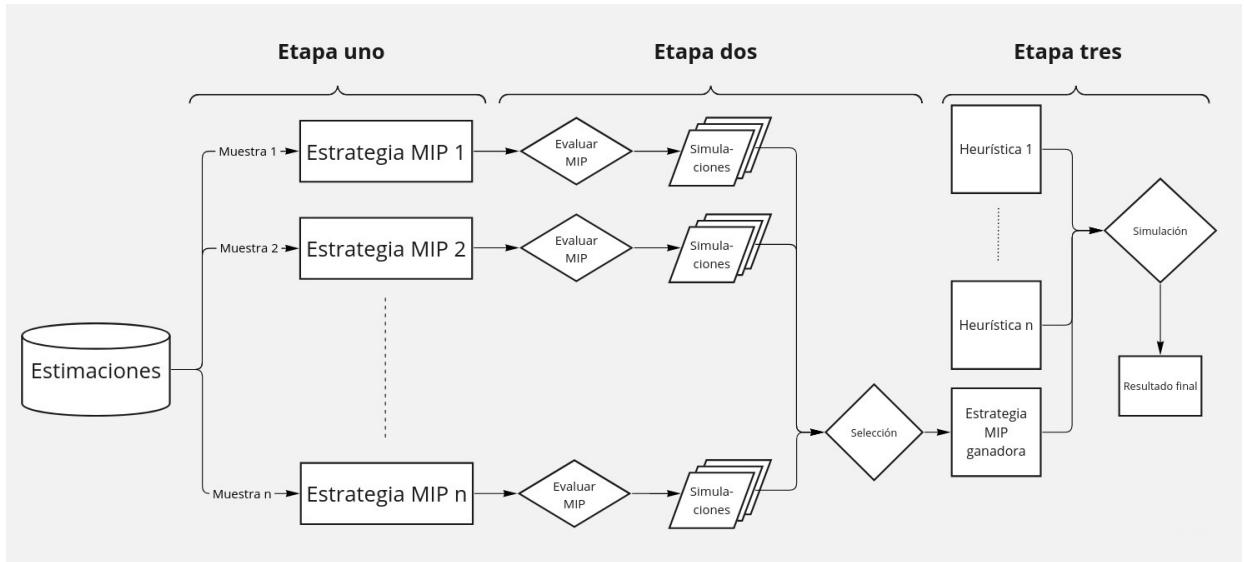


Fig. 5.1: ESQUEMA DEL FRAMEWORK

En la **etapa uno**, obtenemos  $N$  muestras de nuestras estimaciones. Generamos estas muestras usando una distribución normal, tomando nuestra estimación inicial como media y con un desvío bajo ya que como dijimos antes, nuestra estimación no debería diferir mucho de las condiciones que se den realmente. Luego **calculamos un MIP diferente para cada una de estas muestras**. De esta forma obtendremos  $N$  estrategias diferentes.

En la **segunda etapa**, una vez generados los MIPs, **realizamos muchas simulaciones del mismo utilizando más muestras** generadas a partir de las estimaciones que usamos como input para calcular el modelo. Obtenemos las muestras de la misma manera que en la etapa uno.

Luego, podremos **medir el desempeño de los distintos MIP** en base al resultado de las simulaciones. Así podremos quedarnos con el MIP que haya tenido mejores resultados en estas simulaciones.

Una vez elegido el MIP ganador, llegamos la **etapa tres**. En esta etapa realizamos la **comparación del MIP ganador con las otras heurísticas**. Para hacer esta comparación utilizamos los datos reales del día para el que estamos haciendo la prueba.

### 5.7.1. ¿Con qué criterio seleccionamos el mejor MIP?

Dijimos que en este *framework* generamos muchos MIP diferentes en la etapa uno, y los evaluamos en la etapa dos para luego poder elegir el mejor MIP.

La pregunta que nos deberíamos hacer es ¿cuál es el mejor MIP? La respuesta es que **depende de qué es lo que estemos buscando**.

En caso de que queramos tener un criterio más conservador, podríamos quedarnos con el MIP al cual le fue bien a lo largo de todas las evaluaciones. Por ejemplo, podríamos seleccionar el MIP que tenga una **mejor mediana o mejor promedio** para sus evaluaciones.

Si, en cambio, quisieramos adoptar un criterio más arriesgado, podríamos quedarnos con el MIP que haya tenido un **máximo mayor en sus evaluaciones**. De esta manera elegimos el MIP que tiene más potencial para que le vaya mejor, pero posiblemente le vaya mal en otras condiciones.

Consideramos que el hecho de poder fácilmente cambiar el criterio de selección para quedarnos con distintos MIP pudiendo adoptar un criterio más conservador o uno más arriesgado, es otra virtud de este *framework*, ya que le da una mayor flexibilidad en su utilización.

Yendo puntualmente al contexto de esta tesis, **elegimos ir por un camino más conservador y usar la mediana** como criterio de selección a la hora de elegir el mejor MIP. Tomamos esta decisión porque la mediana es una métrica fácil de calcular y además es robusta a *outliers*, a diferencia de otras métricas como por ejemplo la media.

### 5.7.2. ¿Cómo hacemos el muestreo?

Parte de nuestro *framework* de experimentación consiste en calcular muchos MIP con distintas muestras de las estimaciones. Y evaluarlos con aún más muestras.

Nuestras estimaciones consisten, por un lado, en los precios de los mercados, y por otro lado en la energía que recibirán los paneles.

Para los precios tomamos muestras con una distribución normal y un **desvío del 20 % para el mercado RTM y del 10 % para el mercado DAM**. Elegimos un desvío mayor para el RTM ya que es el mercado más variable. Cada muestra es independiente de las demás.

En el caso de la energía que recibirán los paneles, lo cual depende de la radiación solar y la nubosidad del momento, quisimos hacer un muestreo donde cada muestra depende de la muestra anterior. Tomamos esta decisión debido a que la radiación y la nubosidad no suelen comportarse de forma tan cambiante de una hora a otra. Si, por ejemplo, un día a las 14hs. tenemos pleno sol, es probable que a las 15hs. la situación no haya cambiado por completo y tengamos un día completamente nublado. Por este motivo hicimos que la muestra de una hora se tome a partir del promedio de la predicción que tenemos para esa hora y la muestra que obtenida en la hora anterior. Si se pensara a las muestras como una función matemática se la podría escribir de la siguiente forma:

$$\text{muestra}(\text{hora}) = \begin{cases} \text{normal}(\text{prediccion}(\text{hora})) & \text{si hora} = 0 \\ \frac{\text{normal}(\text{muestra}(\text{hora}-1)) + \text{prediccion}(\text{hora})}{2} & \text{si no} \end{cases} \quad (5.5)$$

### 5.7.3. ¿Cómo se componen las estimaciones?

Para poder calcular los distintos MIP que utilizaremos en el *framework* necesitamos estimaciones de los precios, de la radiación solar y de la nubosidad del día.

**Tanto para el DAM como para el RTM vamos a usar los precios del día anterior** como estimaciones. Ya vimos en el capítulo 3 que esto es la mejor alternativa para el DAM. En cuanto al RTM, un mercado que es muy variable y que no se puede predecir, creemos que usar el día anterior es la opción más razonable ya que es lo más cercano que tenemos a lo que pueda pasar.

**Para lo que es información del clima, vamos a utilizar la información de Visual Crossing** como estimaciones. *Visual Crossing* nos brinda un pronóstico del tiempo

para días futuros. Si estuviéramos usando el *framework* en la realidad y no en un experimento, utilizaríamos este pronóstico. Sin embargo, en este experimento tenemos que usar días pasados y no futuros, ya que queremos poder simular que hubiera pasado en la realidad luego al usar el *framework* o las heurísticas. Para este caso lo que vamos a hacer es directamente **usar la información real de *Visual Crossing*** para ese día. Al no usar directamente los datos de *Visual Crossing* y utilizar muestras que hacen pequeños cambios a esta información, estamos simulando lo que habría sido usar un pronóstico del tiempo para ese día.

#### 5.7.4. Validando el framework

Nuestro *framework* consta de tres etapas. En la etapa uno generamos distintas estrategias tomando distintas muestras como input. En la etapa dos evaluamos cada estrategia con más muestras y seleccionamos la mejor estrategia. En la tercera etapa simplemente comparamos nuestra estrategia seleccionada con las heurísticas.

Supongamos que al modificar el *input* tomando distintas muestras el MIP fuera siempre el mismo o muy parecido para cada muestra. En este caso la etapa uno del *framework* no aportaría mucho valor a la hora de seleccionar la mejor estrategia. Los MIPs obtenidos al realizar las muestras tendrían los mismos resultados. Podríamos directamente calcular el MIP con las estimaciones y comparar contra las heurísticas.

Por otro lado, para validar que la etapa dos nos aporta información valiosa, también querríamos validar que al hacer muchas simulaciones, obtengamos resultados diferentes. El razonamiento es similar, si al evaluar las estrategias con distintas muestras, obtuviéramos siempre un resultado similar, entonces la etapa dos no nos estaría aportando demasiado, podríamos simplemente hacer una sola evaluación de cada estrategia y quedarnos con la mejor.

En las siguientes secciones **vamos a validar que las etapas del *framework* nos serán útiles** a la hora de quedarnos con la mejor estrategia.

#### 5.7.5. ¿Cómo afectan los sampleos a los MIP?

En esta sección vamos a observar cómo cambian los MIP al calcularlos utilizando como *input* distintas muestras de nuestras estimaciones de precios y de clima. Para poder ver esto, calculamos 100 MIPs diferentes utilizando como *input* distintas muestras generadas a partir de la estimación de precios que hizo nuestro modelo predictivo para el 17 de mayo del 2022, y evaluamos esos MIP simulándolos con datos reales del día siguiente.

Para visualizar la distribución graficamos los resultados de los diferentes MIP. Además mostramos en ese gráfico los resultados de evaluar las heurísticas usando los datos reales para poder compararlos.

Viendo el gráfico en la figura 5.2 podemos observar que cambiar el *input* de los MIP hace que el resultado de los mismos varíe considerablemente. La mediana fue de una ganancia de USD 1975.83, muy similar a la media, que fue de USD 1973.44. Mientras tanto, el desvío de los resultados fue de USD 146.99. Con un máximo de USD 2246.24 y un mínimo de USD 1369.13.

Mientras tanto, las heurísticas obtuvieron resultados de USD 1744.39 y USD 1657.87.

Podemos observar que según el input que se le de al MIP, podríamos terminar con una estrategia que supera por un buen margen a ambas heurísticas, o una que pierde contra ambas.

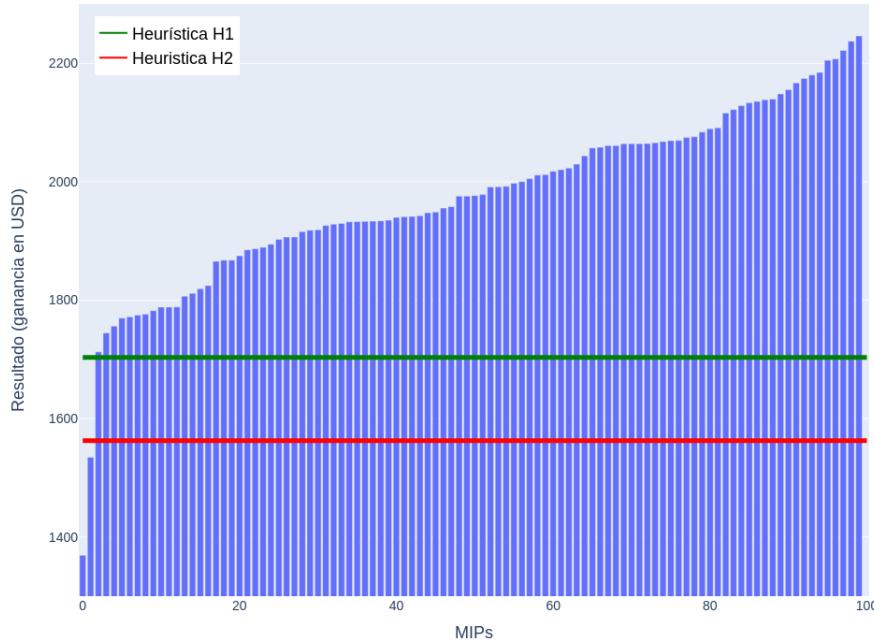


Fig. 5.2: Resultados de 100 MIPs con diferentes muestras de precios del 17 de mayo.

Esto muestra que **según las distintas muestras, el resultado final del MIP puede tener una diferencia digna de tener en cuenta**, por lo que la etapa uno del *framework* puede sernos útil para encontrar la mejor estrategia.

#### 5.7.6. ¿Cómo afectan los sampleos a las simulaciones?

Ya pudimos validar que la etapa uno del *framework* propuesto nos es útil. Ahora veamos si la segunda etapa también lo es.

Recordemos que para poder validar esto, **queríramos que al hacer muchas simulaciones con distintas muestras sobre un MIP ya calculado, obtengamos resultados diferentes**. El razonamiento es similar al de la sección anterior, si para cada muestra obtuviéramos un resultado similar, entonces evaluar cada estrategia muchas veces no nos estaría aportando demasiado, podríamos simplemente hacer una sola evaluación de cada estrategia y quedarnos con la mejor.

El experimento que realizamos para esto consistió en calcular el MIP con el *input* de las estimaciones de nuestro modelo predictivo del día 17 de mayo del 2022.

Una vez obtenido el MIP, tomamos 100 muestras de los precios reales del 17 de mayo, y evaluamos el MIP para cada muestra. Los resultados de las evaluaciones fueron los siguientes:

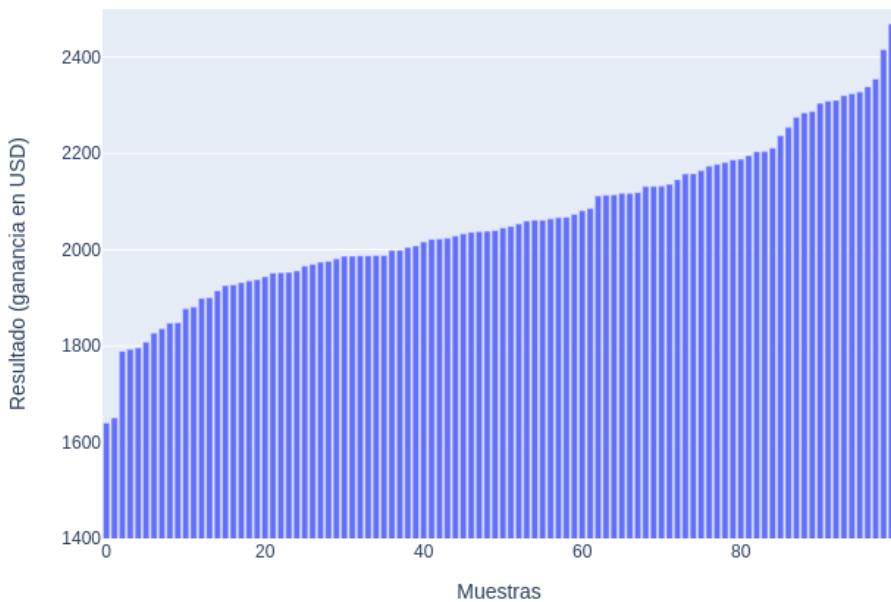


Fig. 5.3: Resultados de 100 simulaciones diferentes para el mismo MIP el 17 de mayo.

Podemos ver que nuevamente los resultados de las simulaciones varían bastante según la muestra. La simulación con menor ganancia fue de USD 1639.61, mientras que la que obtuvo mayores ganancias fue de USD 2469.12. Con una mediana de USD 1988.21 y una media de USD 2059.8714.

Con esto podemos validar que una **estrategia puede tener resultados muy diversos tomando distintas muestras**, por lo que el paso dos de nuestro *framework* nos aporta más información sobre qué MIP puede tener mejores resultados en la realidad.

#### 5.7.7. ¿Qué vamos a evaluar en la comparación final?

Al diseñar este *framework*, tuvimos que elegir qué métricas usar para poder comparar al MIP ganador con las heurísticas seleccionadas.

Sabemos que el **principal punto es la ganancia neta**, que se desprende de los ingresos totales y de las multas. Pero también hay otras métricas que nos resultaron interesantes para poder, además de ver cómo se diferencian las distintas estrategias en el resultado final, ver qué diferencia hay en el comportamiento de las métricas. ¿Qué estrategia usa más las baterías? ¿A qué mercado le da mayor preponderancia cada estrategia?

Por eso nos quedamos con las siguientes métricas a comparar:

- **Ingresos (USD)**: El total de ingresos que conseguimos vendiendo al DAM y al RTM
- **Multas (USD)**: Monto de las multas que tiene la estrategia por haber cumplido con lo subastado en el DAM. Es la cantidad de MWh faltante, al precio del RTM en ese momento.

- **Ganancia neta (USD):** Son los ingresos menos las multas.
- **Total energía cargada (MWh):** La energía de los paneles utilizada para cargar las baterías durante todo el día.
- **Total energía descargada (MWh):** La cantidad de energía de las baterías que fue descargada para vender a los mercados.
- **Energía BAT vendida a DAM (MWh)**
- **Energía BAT vendida a RTM (MWh)**
- **Energía panel vendida a RTM (MWh)**
- **#Intervalos de carga BAT:** Cantidad de intervalos de 5 min que fueron utilizados para cargar las baterías.
- **#Intervalos de descarga BAT:** Cantidad de intervalos de 5 min en los que se descargaron las baterías.
- **Ingresos totales del mercado DAM (USD):** De los ingresos, cuántos fueron por vender al DAM.
- **Ingresos totales del mercado RTM:** De los ingresos, cuantos fueron originados por vender al RTM.
- **Energía de panel sin uso (MWh):** Energía recibida por los paneles que no fue utilizada para vender al RTM ni para cargar las baterías.
- **Energía del panel recibida:** Es la suma de la cantidad de energía recibida por los paneles utilizada para cargar baterías, para vender al RTM y la que no fue utilizada.

#### 5.7.8. Ejemplo del framework

Para poder ilustrar como funciona el framework. Hicimos una prueba del mismo, en el que tomamos las estimaciones de precios y clima del 17 de mayo del 2022. Tomamos 4 muestras de estas estimaciones para generar **4 MIPs diferentes**. Luego hicimos **1000 simulaciones de estos MIPs** utilizando otras muestras de las estimaciones. Las evaluaciones dieron los siguientes resultados:

MIP	Media	Varianza	Mediana
1	2,145.19	28,319.85	2,146.87
2	2,071.78	12,502.22	2,077.08
3	2,196.92	21,782.03	2,203.05
4	1,998.45	15,320.86	2,002.71

Al usar la mediana como criterio de selección, el MIP elegido es el número 3, con una mediana de USD 2,223.3.

Luego de esto probamos esta estrategia contra las heurísticas anteriormente mencionadas.

Podemos ver que en este ejemplo el MIP obtuvo ganancias similares a las heurísticas. En cuanto a los ingresos, el MIP supera a las heurísticas en alrededor de un 10%, incluso

Categoría	MIP	Heurística H1	Heurística H2
Ingresos	2,201.41	2,028.85	2,008.29
Multas	105.67	0.0	0.0
Ganancia neta	2,095.74	2,028.85	2,008.29
Total energía cargada	18.74	14.99	18.74
Total energía descargada	18.74	14.99	18.74
Energía BAT vendida al DAM	18.01	14.99	18.74
Energía BAT vendida al RTM	0.72	0	0
Energía panel vendida al RTM	6.18	9.59	6.39
#Intervalos de carga BAT	82	51	71
#Intervalos de descarga BAT	48	24	48
Ingresos totales del mercado DAM	1,854.23	1,601.23	1,669.1
Ingresos totales del mercado RTM	347.18	427.62	339.19
Energía panel sin uso	3.93	4.26	3.71
Energía panel recibida	28,84	28,84	28,84

Tab. 5.2: Ejemplo del *framework* con 4 MIPs y 1000 tests por MIP.

cuando la heurística **H2** vende una mayor cantidad de energía. Esto muestra que el MIP hizo un uso más eficiente de la energía obtenida.

La ganancia de las dos heurísticas es prácticamente igual. La heurística **H2** vende una mayor cantidad de energía al DAM, pero la **H1** vende más energía al RTM.

Tambien podemos ver que el MIP tiene multas, a diferencia de las heurísticas que no han tenido ninguna multa en este caso.

## 5.8. Primer experimento

Habiendo ya implementado el *framework* y logrado utilizar el mismo en un pequeño ejemplo, el siguiente paso fue usarlo en un contexto más amplio.

Hicimos el experimento de **ejecutar el *framework* varios días seguidos, particularmente del 17 de mayo del 2022 al 27 del mismo mes**.

Para cada día la ejecución fue de calcular 10 MIPs con muestras distintas y evaluar cada MIP 1000 veces.

Podemos ver los resultados en la siguiente tabla.

	Fecha	17/05/2022	18/05/2022	19/05/2022	20/05/2022	21/05/2022	22/05/2022	23/05/2022	24/05/2022	25/05/2022	26/05/2022
Ingresos	MIP	2,265.27	2,464.18	2,712.77	1,869.11	1,932.64	1,887.69	2,478.31	3,214.76	3,251.87	3,263.51
	H1	2,028.85	2,315.74	2,056.83	1,786.88	1,869.41	1,619.05	2,291.10	2,924.14	2,918.75	3,050.9
	H2	2,008.29	2,226.71	2,051.00	1,770.86	1,755.14	1,617.15	2,172.11	2,754.63	2,549.15	2,904.44
Multas	MIP	44.40	0.00	77.47	0.00	0.00	0.00	11.15	0.00	0.00	32.94
	H1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Ganancia neta	MIP	2,220.87	2,464.18	2,635.30	1,869.11	1,932.64	1,887.69	2,467.15	3,214.76	3,251.87	3,230.57
	H1	2,028.85	2,315.74	2,056.83	1,786.88	1,869.41	1,619.05	2,291.10	2,924.14	2,918.75	3,050.9
	H2	2,008.29	2,226.71	2,051.00	1,770.86	1,744.09	1,617.15	2,172.11	2,754.63	2,549.15	2,904.44
Total energía cargada	MIP	17.78	18.86	17.90	19.15	17.75	17.58	16.49	17.75	16.93	18.18
	H1	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99
Total energía descargada	MIP	17.78	18.86	17.90	19.15	17.75	17.58	16.49	17.75	16.93	18.18
	H1	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99
Energía BAT vendida a DAM	MIP	17.74	18.74	18.74	18.74	18.74	18.63	18.74	18.74	18.74	18.74
	H1	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99
Energía BAT vendida a RTM	MIP	17.74	18.74	18.74	18.74	18.74	18.63	18.74	18.74	18.74	18.74
	H1	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99
Energía panel vendida a RTM	MIP	17.24	4.36	6.89	2.75	10.50	14.96	10.16	0.00	0.00	9.01
	H1	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99	14.99
#Intervalos de carga BAT	MIP	0.54	14.50	11.01	16.40	7.25	2.63	6.33	17.75	16.93	9.17
	H1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Energía panel vendida a RTM	MIP	6.49	9.21	8.46	8.05	10.10	5.32	11.98	10.72	10.34	13.46
	H1	9.59	13.08	12.99	12.01	12.86	7.92	13.48	13.48	12.28	14.99
#Intervalos de descarga BAT	MIP	83.00	90.00	79.00	86.00	90.00	80.00	86.00	90.00	90.00	90.00
	H1	51.00	58.00	60.00	55.00	64.00	56.00	62.00	61.00	59.00	60.00
Ingresos totales del mercado	MIP	71.00	90.00	85.00	76.00	96.00	72.00	94.00	94.00	94.00	94.00
	H1	61.00	51.00	32.00	52.00	37.00	41.00	37.00	31.00	28.00	39.00
Ingresos totales del mercado	MIP	24.00	24.00	24.00	24.00	24.00	24.00	24.00	24.00	24.00	24.00
	H2	48.00	48.00	48.00	48.00	48.00	48.00	48.00	48.00	48.00	48.00
Ingresos totales del mercado	MIP	1,936.77	173.82	858.06	63.25	94.96	1,689.62	1,306.13	0.00	0.00	1,257.78
	H1	1,601.23	1,801.23	1,656.35	1,466.24	1,483.50	1,400.90	1,716.26	1,944.57	2,174.73	1,899.68
Ingresos totales del mercado	MIP	1,669.10	1,915.88	1,707.21	1,518.52	1,514.90	1,399.74	1,812.72	2,079.65	2,170.98	2,098.04
	H1	427.62	2,290.35	1,854.71	1,805.86	983.67	198.07	1,172.18	3,214.76	3,251.87	2,005.73
Energía panel sin uso	MIP	339.19	310.83	349.87	320.64	385.91	218.14	574.84	979.57	744.02	1,150.41
	H1	28.84	28.07	27.98	28.23	27.85	240.24	217.41	359.40	674.98	806.40
Energía total recibida	MIP	28.84	28.07	27.98	28.23	27.85	27.86	28.47	28.47	27.27	28.45
	H2	28.84	28.07	27.98	28.23	27.85	27.86	28.47	28.47	27.27	28.45

**Los resultados del primer experimento fueron muy positivos.** En los 10 días el modelo generado mediante un MIP a partir de nuestras estimaciones superó a ambas heurísticas por márgenes considerables. En promedio, la estrategia **MIP tuvo ganancias netas que superaron a las heurísticas en un 10%**, obteniendo una diferencia máxima del 28% el 19 de mayo.

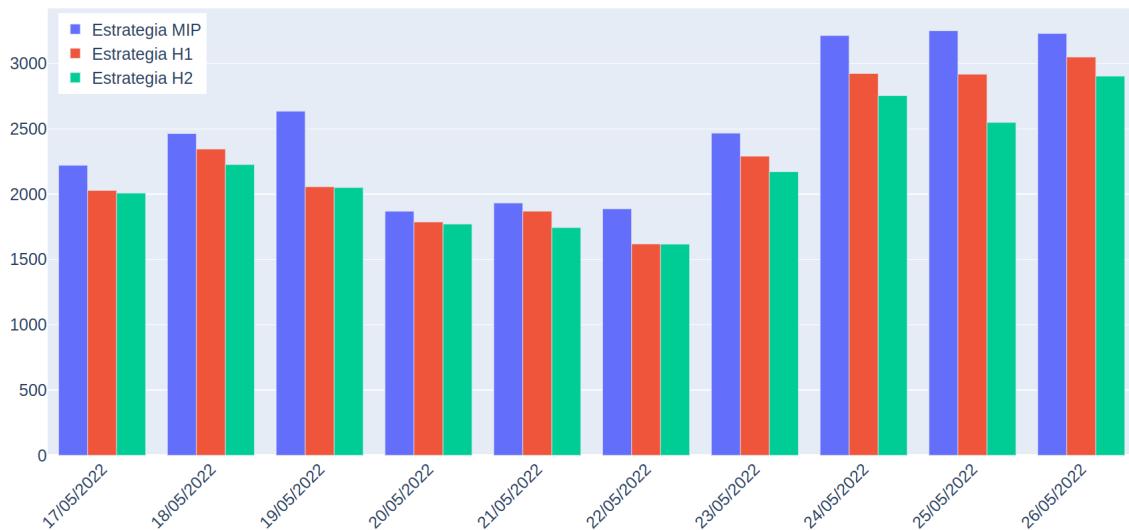


Fig. 5.4: Ganancia neta de las distintas estrategias.

Podemos observar que la heurística **H1 nunca tiene multas**, esto es debido a que se trata de estrategia más conservadora, vendiendo la energía de sus baterías sólo una vez, habiéndolas cargado durante 8 horas. Mientras tanto, podemos ver que la estrategia brindada por el **MIP y por la heurística H2**, tienen algunas multas debido a seguir estrategias algo más arriesgadas.

Podemos corroborar que todas las estrategias utilizan toda la energía que carga en las baterías. Toda la energía cargada es luego descargada.

No ocurre lo mismo con toda la energía recibida por los paneles. **Podemos ver que la energía recibida por los paneles es la misma en todas las estrategias, pero la cantidad de energía utilizada varía según la estrategia.** Esto ocurre porque hay algunos días donde las estrategias deciden no utilizar una parte de la energía de los paneles. Esto puede verse en la columna “Energía panel sin uso”, y ocurre porque hay momentos donde el precio del RTM es cero y las baterías están cargadas por completo. Por lo que no hay nada que podamos hacer con la energía recibida.

Por ejemplo, en la figura 5.5 podemos ver un gráfico de los precios del RTM el día 17 de mayo, donde hubo una mayor cantidad de energía sin utilizar. Desde las 10 a las 15 los precios del RTM fueron cero. Por lo que no se pudo vender energía al RTM en ese lapso.

De la misma manera se puede observar que hay días como el 19 y 22 de mayo donde la energía descargada por las baterías es menor a la cargada. Esto ocurre por el mismo motivo, cuando recibimos más energía de la prevista por la estrategia y los precios en el RTM son cero, la energía es usada para cargar las baterías. Al no haber sido contemplada por las estrategias, esta energía extra podría no ser utilizada, llevando a que se cargue energía que luego no es descargada.

Algo llamativo es que **el MIP varía bastante la estrategia en los distintos**

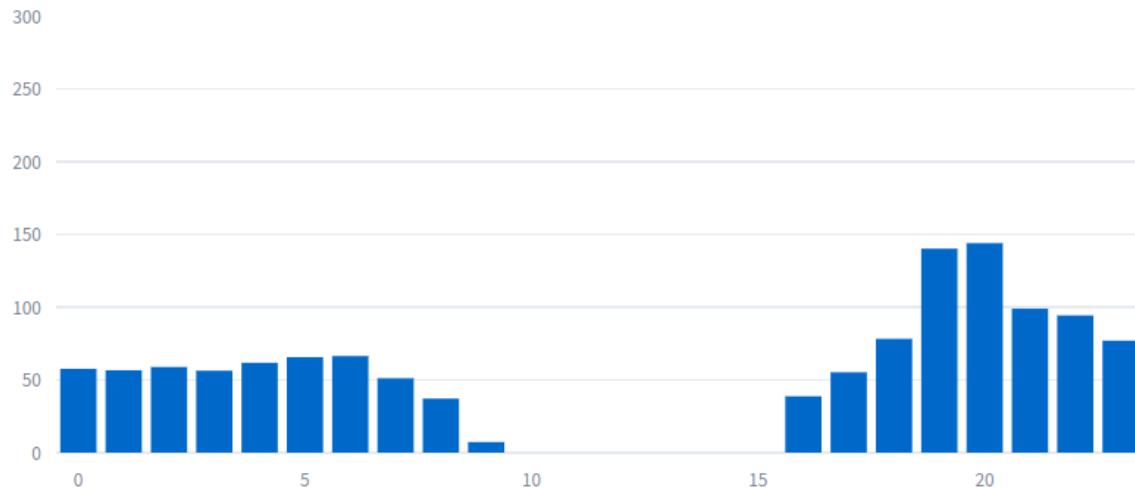


Fig. 5.5: Precios del RTM el 17 de mayo del 2022.

**días.** Podemos ver esto si observamos como distribuye la energía de las baterías en cada día. Algunos días, como el 17 o el 22 de mayo, vende casi toda su energía al DAM, comportándose de forma más similar a las otras heurísticas. Sin embargo, muchas veces no usa la energía de la batería para vender al DAM, sino que la usa en gran medida para el RTM, llegando a extremos como el 24 y 25 de mayo, donde vende toda su energía al RTM. Esto tiene sentido si vemos el gráfico de la figura 5.6 donde vemos que los precios del RTM son mayores a los del DAM en esa fecha.

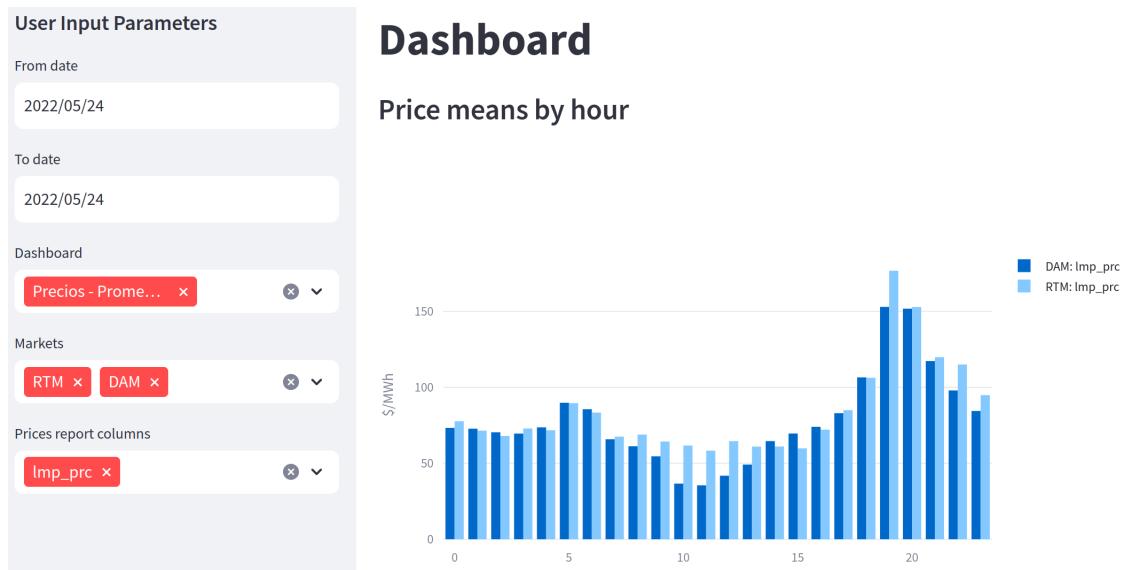


Fig. 5.6: Precios del RTM vs DAM 24 de mayo del 2022.

## 5.9. Experimentacion final

Para poder hacer una comparación final entre nuestra estrategia MIP y las heurísticas. Quisimos hacer un último experimento un poco más amplio que el anterior.

En este experimento comparamos a las distintas estrategias a lo largo de 60 días del año 2022. Utilizamos las siguientes fechas para comparar:

- Del 1 al 20 de enero del 2022.
- Del 1 al 20 de julio del 2022.
- del 15 de septiembre al 4 de octubre del 2022.

Estos días fueron elegidos con el objetivo de tener un panorama más amplio al comparar. Los primeros 20 días son durante el pleno invierno de California, los siguientes días fueron en pleno verano y los últimos días son durante el otoño. De esta manera tenemos días con mucho sol y temperaturas elevadas, días con poco sol y temperaturas bajas, y días con una temperatura y una radiación media.

En este experimento configuramos el *framework* de experimentación para que calcule 40 MIP con distintas muestras y que luego evaluen estos MIPs con 1000 muestras diferentes.

### 5.9.1. Ejecución de los experimentos

Estos experimentos fueron ejecutados en una instancia de *Google Cloud Platform*. Utilizamos una instancia de tipo *e2-standard-4*,<sup>1</sup> con 4 núcleos y 16 GB de memoria.

Se utilizó Python 3.9 y CPLEX 22.1 para resolver los problemas de programación lineal entera.

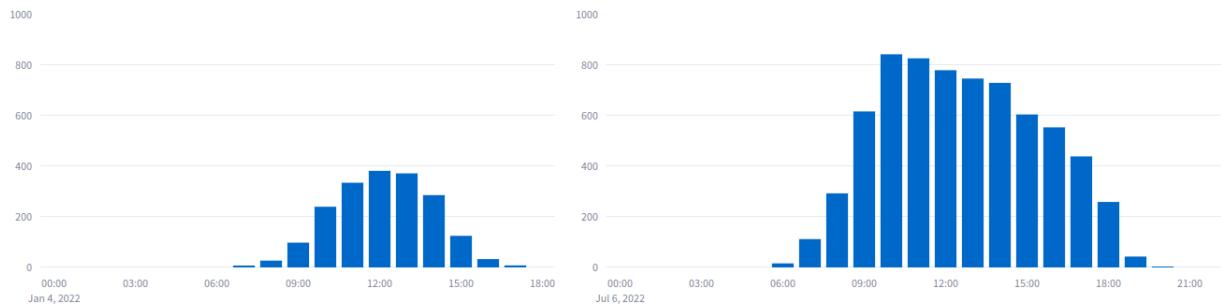
### 5.9.2. Adaptando las heurísticas

Las condiciones climáticas son muy distintas a lo largo de estos días, por lo que las heurísticas fueron levemente modificadas para adaptarse mejor a los distintos escenarios.

Durante los días de verano y otoño, las heurísticas se mantuvieron igual que en la sección anterior. La heurística **H1** carga las baterías en su totalidad y las vende a las 18hs. y 19hs, vendiendo la energía sobrante al RTM. La **H2** hace esto mismo y además carga un cuarto de la batería durante la mañana para vender a las 10hs.

Durante el invierno, la energía recibida por los paneles es mucho menor. Debido a esto, los paneles no logran cargar las baterías por completo durante el día. Por lo que debemos adaptar las heurísticas para vender menos energía y de esta manera no tener tantas multas por incumplir con el mercado DAM.

<sup>1</sup> Familia de máquinas de uso general para Compute Engine | Documentación de Compute Engine. Google Cloud. URL: <https://cloud.google.com/compute/docs/general-purpose-machines?hl=es-419>.



(a) Radiación solar el 4 de enero del 2022.

(b) Radiación solar el 9 de julio del 2022.

La heurística **H1** solo venderá la mitad de su energía, o sea 7.5 MWh. Esta energía será vendida a las 18hs.

La heurística **H2** solo venderá un 75 % de la batería. Un 50 % será vendido a las 18hs, y el 25 % restante a las 10hs.

### 5.9.3. Resultados crudos

A continuación podremos observar en distintas tablas, los resultados crudos de los experimentos. Estos serán luego analizados en las siguientes secciones del capítulo.

Fecha	MIP	Ingresos		Multas		Ganancia neta		Total energía cargada		Total energía descargada	
		Valor fecha	H1	H2	MIP	H1	H2	MIP	H1	H2	MIP
01/01/2022	794,51	682,23	627,30	0,00	17,38	88,67	645,51	664,85	7,29	7,26	7,30
02/01/2022	645,68	44,563,00	645,70	539,14	0,00	49,31	176,57	645,68	7,05	6,87	5,93
03/01/2022	781,14	44,564,00	852,84	767,45	0,00	32,98	205,93	781,14	819,87	7,36	7,01
04/01/2022	634,28	44,565,00	623,04	511,38	47,19	35,07	127,29	587,09	6,58	6,90	6,04
05/01/2022	744,98	44,566,00	663,60	656,51	0,00	26,14	744,98	663,60	8,03	7,60	8,15
06/01/2022	634,48	44,567,00	614,70	545,37	51,40	14,15	149,53	553,08	6,83	6,27	6,84
07/01/2022	562,41	44,568,00	535,57	478,31	15,58	9,91	124,39	546,83	525,65	7,19	7,28
08/01/2022	561,81	44,569,00	556,71	490,15	17,53	9,75	547,28	539,46	7,12	6,28	7,20
09/01/2022	578,57	44,570,00	555,80	449,41	0,00	0,00	69,51	578,57	555,80	7,73	8,11
10/01/2022	614,13	44,571,00	603,16	532,83	0,00	72,53	197,24	614,13	530,63	6,80	6,25
11/01/2022	677,63	44,572,00	594,61	515,65	26,96	0,00	95,99	650,67	594,61	7,74	6,66
12/01/2022	561,29	44,573,00	536,74	494,94	38,20	57,40	165,82	573,34	629,26	6,27	6,40
13/01/2022	113,19	44,574,00	89,27	79,27	13,92	34,44	425,22	99,26	-255,24	1,21	0,62
14/01/2022	618,51	44,575,00	530,24	577,44	0,00	6,18	618,51	530,24	8,61	8,22	0,63
15/01/2022	147,18	44,576,00	138,68	25,48	315,83	121,70	-121,70	-121,70	-177,16	1,21	0,93
16/01/2022	477,04	44,577,00	492,20	395,45	27,74	63,52	178,61	449,30	428,68	6,13	6,21
17/01/2022	13,00	44,578,00	17,61	15,40	0,36	367,37	446,32	12,64	-349,75	0,22	0,18
18/01/2022	666,57	44,579,00	600,36	593,33	18,04	0,00	47,12	648,53	600,36	8,10	7,64
19/01/2022	732,02	44,580,00	670,73	630,03	0,00	104,42	705,62	670,73	8,12	7,97	7,00
20/01/2022	702,90	44,581,00	588,53	596,20	0,00	0,00	52,66	702,90	588,53	8,51	8,73
01/07/2022	1,931,85	44,743,00	2,133,81	2,124,07	0,00	0,00	0,00	1,931,85	2,133,81	19,40	14,99
02/07/2022	1,652,11	44,744,00	1,540,81	1,455,05	4,36	0,00	0,00	1,647,74	1,540,81	18,47	14,99
03/07/2022	1,277,19	44,745,00	1,086,23	1,084,88	0,00	0,00	0,00	1,277,19	1,086,23	20,86	14,99
04/07/2022	1,435,03	44,746,00	1,232,30	1,235,17	65,46	0,00	0,00	1,369,57	1,232,30	20,04	14,99
05/07/2022	1,964,02	44,747,00	1,755,07	1,776,45	102,30	0,00	0,00	1,861,72	1,755,07	18,98	14,99
06/07/2022	1,926,45	44,748,00	1,758,61	1,743,68	14,94	0,00	0,00	1,911,51	1,758,61	19,12	14,99
07/07/2022	2,119,74	44,749,00	2,045,82	2,043,84	49,14	0,00	0,00	2,045,82	2,045,82	20,93	14,99
08/07/2022	2,207,01	44,750,00	2,123,84	2,111,59	84,83	0,00	0,00	2,122,18	2,123,84	18,31	14,99
09/07/2022	2,182,55	44,751,00	2,087,76	2,009,08	18,48	0,00	0,00	2,164,07	2,087,76	18,71	14,99
10/07/2022	2,477,37	44,752,00	2,316,37	2,270,35	23,71	0,00	0,00	2,453,67	2,316,37	18,76	14,99
11/07/2022	3,522,31	44,753,00	3,477,83	3,447,83	0,00	0,00	0,00	3,522,31	3,477,94	18,92	14,99
12/07/2022	2,958,64	44,754,00	2,899,99	2,865,36	22,82	0,00	0,00	2,958,62	2,899,99	19,70	14,99
13/07/2022	2,940,25	44,755,00	2,891,99	2,854,39	97,69	0,00	0,00	2,842,56	2,854,39	18,80	14,99
14/07/2022	3,170,12	44,756,00	3,046,12	2,784,10	44,88	0,00	0,00	123,53	3,125,23	3,046,12	16,53
15/07/2022	3,346,42	44,757,00	3,408,81	3,326,76	80,23	0,00	0,00	3,266,19	3,408,81	18,61	14,99
16/07/2022	3,436,10	44,758,00	3,428,35	3,228,57	39,87	0,00	0,00	123,40	3,462,35	3,296,23	20,64
17/07/2022	3,674,17	44,759,00	3,299,19	3,215,19	0,00	0,00	0,00	3,674,17	3,299,19	18,64	14,99
18/07/2022	3,407,76	44,760,00	3,596,91	3,483,08	0,00	0,00	0,00	168,04	3,407,76	3,596,91	18,00
19/07/2022	4,405,83	44,761,00	4,129,06	4,059,81	4,05	0,00	0,00	17,74	4,129,06	19,74	14,99
20/07/2022	3,068,70	44,762,00	3,507,38	3,409,39	0,00	0,00	0,00	3,068,70	3,507,38	19,70	14,99
15/09/2022	1,940,68	44,819,00	1,893,43	1,835,64	45,87	0,00	1,47	1,894,81	1,893,43	17,69	14,99
16/09/2022	1,806,21	44,820,00	1,733,48	1,731,93	33,07	0,00	34,71	1,773,14	1,733,48	17,33	14,99
17/09/2022	1,388,90	44,821,00	1,392,44	1,427,91	0,00	0,00	0,00	1,891,77	1,392,44	16,38	14,99
18/09/2022	1,605,83	44,822,00	1,643,42	1,404,88	37,41	0,00	0,00	106,99	1,568,42	1,643,42	16,73
19/09/2022	2,034,43	44,823,00	1,980,75	1,908,50	0,00	0,00	0,00	99,89	2,034,43	1,980,75	15,89
20/09/2022	2,001,07	44,824,00	1,795,43	1,705,84	0,00	0,00	0,00	125,65	2,001,07	1,795,43	15,20
21/09/2022	1,539,93	44,825,00	1,866,45	1,826,04	0,00	0,00	0,00	146,74	1,866,45	15,76	14,99
22/09/2022	1,922,69	44,826,00	1,982,59	1,885,77	30,91	0,00	0,00	32,99	1,891,77	1,982,59	15,99
23/09/2022	2,058,82	44,827,00	2,019,08	1,865,71	1,427,91	0,00	0,00	143,53	2,058,82	2,019,08	17,04
24/09/2022	1,605,83	44,828,00	2,072,85	1,937,80	17,46	0,00	0,00	128,85	2,072,85	16,12	14,99
25/09/2022	2,174,98	44,829,00	2,131,16	2,006,33	0,00	0,00	0,00	107,40	2,174,98	2,131,16	16,13
26/09/2022	4,202,52	44,830,00	3,173,20	2,849,98	18,19	0,00	0,00	308,23	4,184,33	3,179,20	15,19
27/09/2022	4,163,82	44,831,00	3,494,00	3,149,58	63,22	0,00	0,00	281,52	4,163,82	15,01	14,99
28/09/2022	4,420,59	44,832,00	3,261,14	2,980,53	13,15	0,00	0,00	235,72	4,407,44	3,261,14	15,29
29/09/2022	1,764,40	44,833,00	1,854,02	1,761,53	0,00	0,00	0,00	164,01	1,764,40	1,854,02	15,00
30/09/2022	1,474,39	44,834,00	1,507,52	1,422,04	35,47	0,00	0,00	202,44	1,438,93	1,507,52	15,31
01/10/2022	1,342,91	44,835,00	1,361,56	1,227,33	33,20	0,00	0,00	194,13	1,340,71	1,361,56	15,01
02/10/2022	1,459,50	44,836,00	1,369,97	1,204,44	0,00	0,00	0,00	227,87	1,459,50	1,369,97	15,22
03/10/2022	1,598,25	44,837,00	1,707,54	1,581,14	0,00	0,00	0,00	245,17	1,598,25	1,707,54	15,55
04/10/2022	1,916,75	44,838,00	2,003,34	1,940,77	0,00	0,00	0,00	82,29	1,916,75	2,003,34	15,70

Fecha	Energía BAT vendida a DAM				Energía BAT vendida a RTM				Energía panel vendida a DAM				#Intervalos de carga BAT				#Intervalos de descarga BAT
	MIP	H1	H2	MIP	H1	H2	MIP	H1	H2	MIP	H1	H2	MIP	H1	H2	MIP	
01/01/2022	5,10	7,27	7,51	2,20	0,00	0,00	0,58	0,62	0,38	93,00	98,00	116,00	16,00	12,00	12,00	22,00	
02/01/2022	0,00	6,88	5,94	7,96	0,00	0,00	1,35	1,53	2,47	90,00	94,00	96,00	12,00	12,00	12,00	13,00	
03/01/2022	3,11	7,02	6,15	4,25	0,00	0,00	1,32	1,67	2,54	91,00	86,00	97,00	12,00	12,00	12,00	13,00	
04/01/2022	5,96	6,91	6,05	0,63	0,00	0,00	1,91	1,59	2,45	90,00	90,00	90,00	114,00	114,00	114,00	13,00	
05/01/2022	7,50	7,50	8,16	0,54	0,00	0,00	0,67	1,10	0,55	90,00	84,00	96,00	13,00	13,00	13,00	21,00	
06/01/2022	5,59	7,25	6,25	1,25	0,00	0,00	2,00	1,60	2,57	90,00	84,00	96,00	13,00	13,00	13,00	14,00	
07/01/2022	7,20	7,29	6,29	0,90	0,00	0,00	1,74	1,64	2,64	90,00	84,00	96,00	24,00	24,00	24,00	14,00	
08/01/2022	7,13	7,15	6,13	0,00	0,00	0,00	1,59	1,56	2,58	91,00	86,00	96,00	12,00	12,00	12,00	13,00	
09/01/2022	7,70	7,50	7,15	0,93	0,00	0,00	1,18	0,79	1,77	91,00	93,00	105,00	25,00	25,00	25,00	21,00	
10/01/2022	4,29	6,26	5,63	2,52	0,00	0,00	1,32	1,88	2,51	90,00	84,00	96,00	17,00	11,00	11,00	13,00	
11/01/2022	7,75	7,50	6,67	0,00	0,00	0,00	1,63	1,71	2,70	90,00	84,00	96,00	16,00	12,00	12,00	14,00	
12/01/2022	5,65	6,41	5,82	0,63	0,00	0,00	1,49	1,36	1,95	90,00	84,00	96,00	12,00	11,00	11,00	13,00	
13/01/2022	1,22	6,63	0,84	0,00	0,00	0,00	0,32	0,91	0,70	90,00	84,00	96,00	10,00	2,00	2,00	5,00	
14/01/2022	7,78	7,50	9,03	0,63	0,00	0,00	0,55	0,43	0,14	94,00	101,00	119,00	25,00	25,00	25,00	24,00	
15/01/2022	1,66	1,22	0,94	0,63	0,00	0,00	1,00	1,28	1,28	91,00	84,00	96,00	10,00	2,00	2,00	5,00	
16/01/2022	5,51	6,22	5,04	0,22	0,00	0,00	1,83	1,79	2,97	91,00	85,00	96,00	13,00	10,00	10,00	13,00	
17/01/2022	0,00	0,19	0,17	0,22	0,00	0,00	0,05	0,09	0,11	90,00	84,00	96,00	7,00	1,00	1,00	2,00	
18/01/2022	8,35	7,50	7,65	0,00	0,00	0,00	1,58	1,82	2,28	90,00	84,00	100,00	20,00	12,00	12,00	17,00	
19/01/2022	8,13	7,50	7,01	0,90	0,00	0,00	1,72	1,87	2,84	90,00	84,00	96,00	19,00	12,00	12,00	16,00	
20/01/2022	7,84	7,50	7,58	0,69	0,00	0,00	1,80	1,72	2,74	90,00	84,00	96,00	26,00	26,00	26,00	17,00	
01/07/2022	0,00	15,00	18,75	19,41	0,00	0,00	11,32	15,73	11,98	90,00	54,00	82,00	32,00	24,00	24,00	24,00	
02/07/2022	14,78	15,00	18,75	3,71	0,00	0,00	12,26	15,75	12,00	88,00	54,00	82,00	32,00	24,00	24,00	24,00	
03/07/2022	17,16	15,00	18,75	3,71	0,00	0,00	1,82	3,28	3,58	101,00	68,00	82,00	51,00	24,00	24,00	48,00	
04/07/2022	20,05	15,00	18,75	0,00	0,00	0,00	5,47	10,51	10,40	105,00	67,00	83,00	83,00	24,00	24,00	48,00	
05/07/2022	17,45	15,00	18,75	1,54	0,00	0,00	11,41	15,40	11,65	90,00	55,00	83,00	62,00	24,00	24,00	48,00	
06/07/2022	18,50	15,00	18,75	0,63	0,00	0,00	11,41	15,54	11,79	90,00	53,00	82,00	50,00	24,00	24,00	48,00	
07/07/2022	20,50	15,00	18,75	0,44	0,00	0,00	9,85	15,79	12,04	90,00	53,00	81,00	49,00	24,00	24,00	48,00	
08/07/2022	11,85	15,00	18,75	6,47	0,00	0,00	11,98	15,30	11,55	90,00	53,00	82,00	34,00	24,00	24,00	48,00	
09/07/2022	14,49	15,00	18,75	4,24	0,00	0,00	12,15	15,87	12,12	90,00	54,00	82,00	42,00	24,00	24,00	48,00	
10/07/2022	14,44	15,00	18,75	4,33	0,00	0,00	12,03	15,80	12,05	90,00	53,00	81,00	33,00	24,00	24,00	48,00	
11/07/2022	7,50	15,00	18,75	11,43	0,00	0,00	11,98	15,91	12,16	90,00	53,00	81,00	32,00	24,00	24,00	48,00	
12/07/2022	14,04	15,00	18,75	5,68	0,00	0,00	11,35	16,07	12,32	90,00	53,00	79,00	44,00	24,00	24,00	48,00	
13/07/2022	17,56	15,00	18,75	1,25	0,00	0,00	11,99	15,81	12,06	90,00	53,00	81,00	50,00	24,00	24,00	48,00	
14/07/2022	14,33	15,00	17,27	2,20	0,00	0,00	10,06	11,59	9,32	90,00	69,00	96,00	28,00	24,00	24,00	46,00	
15/07/2022	14,09	15,00	18,75	4,53	0,00	0,00	10,32	13,94	10,19	90,00	57,00	89,00	32,00	24,00	24,00	48,00	
16/07/2022	19,00	15,00	16,57	1,25	0,00	0,00	9,25	14,50	12,93	90,00	50,00	77,00	24,00	24,00	24,00	35,00	
17/07/2022	11,75	15,00	18,75	6,90	0,00	0,00	11,66	15,31	11,56	90,00	57,00	75,00	35,00	24,00	24,00	48,00	
18/07/2022	8,27	15,00	16,51	9,74	0,00	0,00	8,17	11,18	9,67	90,00	51,00	78,00	42,00	24,00	24,00	34,00	
19/07/2022	7,50	15,00	18,75	12,25	0,00	0,00	12,64	17,39	13,64	90,00	50,00	75,00	34,00	24,00	24,00	47,00	
20/07/2022	0,00	15,00	12,21	0,00	0,00	0,00	10,47	15,18	11,69	90,00	56,00	83,00	33,00	24,00	24,00	47,00	
21/09/2022	17,53	15,00	17,83	0,17	0,00	0,00	6,37	9,07	6,24	90,00	59,00	96,00	47,00	24,00	24,00	43,00	
22/09/2022	17,34	15,00	17,51	0,90	0,00	0,00	6,51	8,85	6,34	90,00	59,00	96,00	48,00	24,00	24,00	43,00	
23/09/2022	15,73	15,00	18,75	0,66	0,00	0,00	7,02	8,42	4,83	82,00	61,00	94,00	38,00	24,00	24,00	48,00	
18/09/2022	16,16	15,00	16,49	0,63	0,00	0,00	5,38	7,16	5,67	94,00	68,00	98,00	37,00	24,00	24,00	40,00	
19/09/2022	11,39	15,00	17,26	4,51	0,00	0,00	7,50	8,40	6,14	90,00	61,00	96,00	32,00	24,00	24,00	41,00	
20/09/2022	0,00	15,00	16,93	15,21	0,00	0,00	4,98	5,48	5,69	90,00	79,00	96,00	25,00	24,00	24,00	35,00	
21/09/2022	15,00	16,18	14,99	15,77	0,00	0,00	6,31	7,07	5,90	90,00	64,00	96,00	27,00	24,00	24,00	36,00	
22/09/2022	14,53	15,00	17,93	1,48	0,00	0,00	6,54	7,55	4,62	89,00	64,00	101,00	28,00	24,00	24,00	46,00	
23/09/2022	16,98	15,00	15,93	0,97	0,00	0,00	4,72	6,77	5,84	90,00	63,00	96,00	37,00	24,00	24,00	37,00	
24/09/2022	14,55	15,00	16,39	1,00	0,00	0,00	6,63	7,18	5,79	90,00	64,00	96,00	27,00	24,00	24,00	38,00	
25/09/2022	13,27	15,00	16,57	2,87	0,00	0,00	6,24	7,38	5,81	90,00	64,00	96,00	30,00	24,00	24,00	38,00	
26/09/2022	12,42	15,00	14,42	2,78	0,00	0,00	4,98	5,18	5,76	90,00	79,00	96,00	30,00	24,00	24,00	35,00	
27/09/2022	10,35	15,00	14,99	4,68	0,00	0,00	5,57	5,60	5,60	90,00	65,00	96,00	30,00	24,00	24,00	34,00	
28/09/2022	4,49	15,00	15,30	10,81	0,00	0,00	5,53	5,83	5,52	90,00	64,00	96,00	29,00	24,00	24,00	35,00	
29/09/2022	7,50	15,00	15,81	8,25	0,00	0,00	5,69	6,44	5,63	90,00	67,00	96,00	28,00	24,00	24,00	37,00	
30/09/2022	11,82	15,00	14,85	3,50	0,00	0,00	5,28	5,61	5,76	90,00	69,00	96,00	29,00	24,00	24,00	35,00	
01/10/2022	12,87	15,00	14,85	2,15	0,00	0,00	5,46	5,47	5,63	90,00	71,00	96,00	27,00	24,00	24,00	34,00	
02/10/2022	7,50	15,00	14,31	7,72	0,00	0,00	4,60	4,82	5,51	90,00	70,00	96,00	27,00	24,00	24,00	33,00	
03/10/2022	0,00	15,00	14,84	15,59	0,00	0,00	4,63	5,28	5,44								

Fecha	Ingresos totales del mercado DAM			Ingresos totales del mercado RTM			Energía panel sin uso			Energía total recibida		
	MIP	H1	H2	MIP	H1	H2	MIP	H1	H2	MIP	H1	H2
01/01/2022	470,67	674,40	626,08	323,84	7,83	1,21	0,00	0,00	0,00	7,88	7,88	7,88
02/01/2022	0,00	616,99	503,04	645,68	28,71	36,10	0,00	0,00	0,00	8,41	8,41	8,41
03/01/2022	376,58	763,69	632,64	402,56	89,15	134,80	0,00	0,00	0,00	8,68	8,68	8,68
04/01/2022	560,24	607,44	504,87	74,05	15,36	6,51	0,00	0,00	0,00	8,49	8,49	8,49
05/01/2022	693,73	644,89	632,85	51,25	18,71	3,66	0,00	0,00	0,00	8,70	8,70	8,70
06/01/2022	474,49	548,61	455,04	159,99	66,09	90,34	0,00	0,00	0,00	8,83	8,83	8,83
07/01/2022	507,07	481,67	401,04	55,34	53,89	77,26	0,00	0,00	0,00	8,92	8,92	8,92
08/01/2022	534,94	543,96	445,29	29,87	12,75	44,86	0,00	0,00	0,00	8,70	8,70	8,70
09/01/2022	562,42	549,46	448,17	16,15	6,35	1,23	0,00	0,00	0,00	8,90	8,90	8,90
10/01/2022	390,72	521,56	444,16	223,41	81,60	88,68	0,00	0,00	0,00	8,13	8,13	8,13
11/01/2022	645,31	561,01	477,93	32,33	33,60	37,73	0,00	0,00	0,00	9,37	9,37	9,37
12/01/2022	461,45	484,17	418,95	99,84	52,57	76,99	0,00	0,00	0,00	7,76	7,76	7,76
13/01/2022	97,69	46,16	47,71	15,49	43,11	31,56	0,00	0,00	0,00	1,53	1,53	1,53
14/01/2022	570,69	515,55	576,38	47,82	14,69	1,07	0,00	0,00	0,00	9,16	9,16	9,16
15/01/2022	119,96	88,67	54,11	27,22	50,00	60,00	0,00	0,00	0,00	2,21	2,21	2,21
16/01/2022	400,43	457,95	335,43	76,62	34,25	60,01	0,00	0,00	0,00	8,01	8,01	8,01
17/01/2022	0,36	13,97	11,29	12,64	3,64	4,11	0,00	0,00	0,00	0,27	0,27	0,27
18/01/2022	624,65	552,44	542,03	41,92	47,92	51,30	0,00	0,00	0,00	9,92	9,92	9,92
19/01/2022	669,67	605,60	538,57	62,34	65,13	91,46	0,00	0,00	0,00	9,84	9,84	9,84
20/01/2022	610,78	551,65	528,31	92,12	36,88	67,90	0,00	0,00	0,00	10,32	10,32	10,32
01/07/2022	0,00	1,321,76	1,321,76	1,931,85	812,06	647,67	0,00	0,00	0,00	30,72	30,72	30,72
02/07/2022	1,312,72	1,322,36	1,203,49	339,33	408,45	251,55	0,00	0,00	0,00	30,74	30,74	30,74
03/07/2022	1,000,07	1,021,45	1,020,08	277,12	64,79	7,95	12,37	8,31	30,63	30,63	30,63	30,63
04/07/2022	1,357,98	1,024,97	1,028,26	77,05	207,33	206,91	5,13	5,13	1,49	30,63	30,63	30,63
05/07/2022	1,526,79	1,223,35	1,348,37	437,23	531,72	427,88	0,00	0,00	0,00	30,39	30,39	30,39
06/07/2022	1,492,07	1,169,72	1,223,34	1,300,37	434,38	589,27	473,31	0,00	0,00	30,53	30,53	30,53
07/07/2022	1,654,32	1,262,19	1,410,69	465,43	783,63	633,15	0,00	0,00	0,00	30,78	30,78	30,78
08/07/2022	1,188,19	1,385,07	1,561,42	1,018,91	550,17	550,17	0,00	0,00	0,00	30,29	30,29	30,29
09/07/2022	1,418,19	1,351,31	1,477,94	764,37	736,45	531,14	0,00	0,00	0,00	30,86	30,86	30,86
10/07/2022	1,556,41	1,522,58	1,642,13	920,97	793,79	628,21	0,00	0,00	0,00	30,79	30,79	30,79
11/07/2022	1,484,52	2,442,55	2,656,01	2,037,79	1,035,39	791,82	0,00	0,00	0,00	30,90	30,90	30,90
12/07/2022	1,852,40	1,914,01	2,120,25	1,102,65	985,98	745,11	0,00	0,00	0,00	31,06	31,06	31,06
13/07/2022	2,038,67	1,815,39	2,015,52	901,59	1,076,60	838,87	0,00	0,00	0,00	30,80	30,80	30,80
14/07/2022	2,205,54	2,226,99	2,189,87	964,57	819,13	594,23	0,00	0,00	0,00	26,58	26,58	26,58
15/07/2022	2,322,41	2,466,14	2,691,82	1,024,01	942,68	634,94	0,00	0,00	0,00	28,93	28,93	28,93
16/07/2022	2,569,42	2,202,12	2,284,73	866,64	1,226,23	943,84	0,00	0,00	0,00	29,49	29,49	29,49
17/07/2022	1,726,44	2,107,58	2,308,99	1,947,73	1,191,60	906,20	0,00	0,00	0,00	30,30	30,30	30,30
18/07/2022	1,480,71	2,532,28	2,449,05	1,927,05	1,244,62	1,035,93	0,00	0,00	0,00	26,17	26,17	26,17
19/07/2022	1,515,62	2,640,71	2,696,05	2,890,21	1,489,03	1,163,76	0,00	0,00	0,00	32,38	32,38	32,38
20/07/2022	1,019,27	2,321,95	2,551,05	2,049,43	1,185,43	857,44	0,00	0,00	0,00	30,17	30,17	30,17
21/09/2022	1,704,54	1,552,48	1,718,05	236,14	340,95	116,99	0,00	0,00	0,00	24,06	24,06	24,06
22/09/2022	1,625,30	1,676,39	1,771,03	231,47	276,99	150,12	0,00	0,00	0,00	22,54	22,54	22,54
23/09/2022	1,875,15	1,764,44	1,456,49	1,581,82	1,354,80	85,00	118,24	73,11	0,00	21,76	21,76	21,76
24/09/2022	1,370,45	1,398,77	1,342,98	235,38	244,65	61,89	0,00	0,00	0,00	22,15	22,15	22,15
25/09/2022	1,111,33	1,489,84	1,568,11	923,10	480,91	340,39	0,00	0,00	0,00	23,39	23,39	23,39
26/09/2022	0,00	1,449,50	1,487,18	2,001,07	345,92	218,66	0,00	0,00	0,00	20,68	20,68	20,68
27/09/2022	2,375,11	3,093,63	1,527,12	1,539,93	359,63	298,92	0,00	0,00	0,00	22,06	22,06	22,06
28/09/2022	609,60	2,811,26	2,693,05	1,771,03	297,39	306,20	114,04	0,00	0,00	22,54	22,54	22,54
29/09/2022	1,303,89	1,274,20	1,354,80	183,67	254,64	160,93	0,00	0,00	0,00	21,76	21,76	21,76
30/09/2022	994,84	1,242,50	1,149,97	479,55	362,14	265,02	272,07	0,00	0,00	20,60	20,60	20,60
01/10/2022	989,17	1,147,46	1,044,53	2,518,42	2,603,86	1,684,31	258,72	246,12	0,00	20,46	20,46	20,46
02/10/2022	573,28	1,192,44	1,047,39	886,22	1,787,11	400,37	257,68	0,00	0,00	20,59	20,59	20,59
03/10/2022	0,00	1,048,09	1,308,49	1,598,25	276,05	0,00	0,00	0,00	0,00	20,27	20,27	20,27
04/10/2022	1,316,93	1,664,46	1,744,23	599,82	338,45	606,54	0,00	0,00	0,00	21,37	21,37	21,37

Los resultados del experimento fueron positivos. La estrategia **MIP da una mayor ganancia neta en las tres épocas del año.**

Si observamos la ganancia neta, los resultados obtenidos por el MIP superan a los de la heurística **H1** en un promedio del 2.51% durante el periodo del 1 al 20 de julio. En la franja del 15 de septiembre al 4 de octubre, la diferencia es de un promedio de 3.13% en favor del MIP.

Durante la franja del 1 al 20 de enero, que transcurre en el invierno de California, el MIP cuenta con un margen aún mayor. No es tan sencillo de expresar en términos porcentuales debido a que **hay días donde las heurísticas directamente tienen perdidas**, por lo que el porcentaje se distorsiona. Precisamente, esto ocurre el 13, el 15 y el 17 de enero. Pero incluso haciendo un análisis dejando afuera estos días y observamos los 17 días restantes, la diferencia es de 7.66% en promedio.

En las siguientes secciones haremos un análisis de los resultados de cada franja de fechas por separado.

#### 5.9.4. Análisis general

Al analizar los resultados notamos que hay algunas observaciones que se repiten en las tres franjas de fechas.

Un aspecto a mencionar es la forma en la que las heurísticas utilizan la energía de las baterías y contrastarla con el MIP. Las heurísticas venden, por definición, toda la energía de sus baterías al DAM. El MIP, se diferencia en este sentido y la utilización de la energía varía a lo largo de los días. Hay fechas donde efectivamente vende todo al DAM, como el 4 de julio o el 16 de septiembre, pero también tiene muchos días donde la energía vendida al DAM es mucho menor, llegando a extremos como el 1 de julio o el 2 de enero donde no vende energía al DAM. Esto ocurre porque hay fechas donde los valores del mercado RTM son mayores a los del DAM y el MIP elige volcar su energía allí. Por ejemplo, para el 1 de julio, el MIP utiliza como estimaciones los precios del 30 de junio, si comparamos los precios del DAM y el RTM en esa fecha, observaremos que los precios del RTM son mayores.

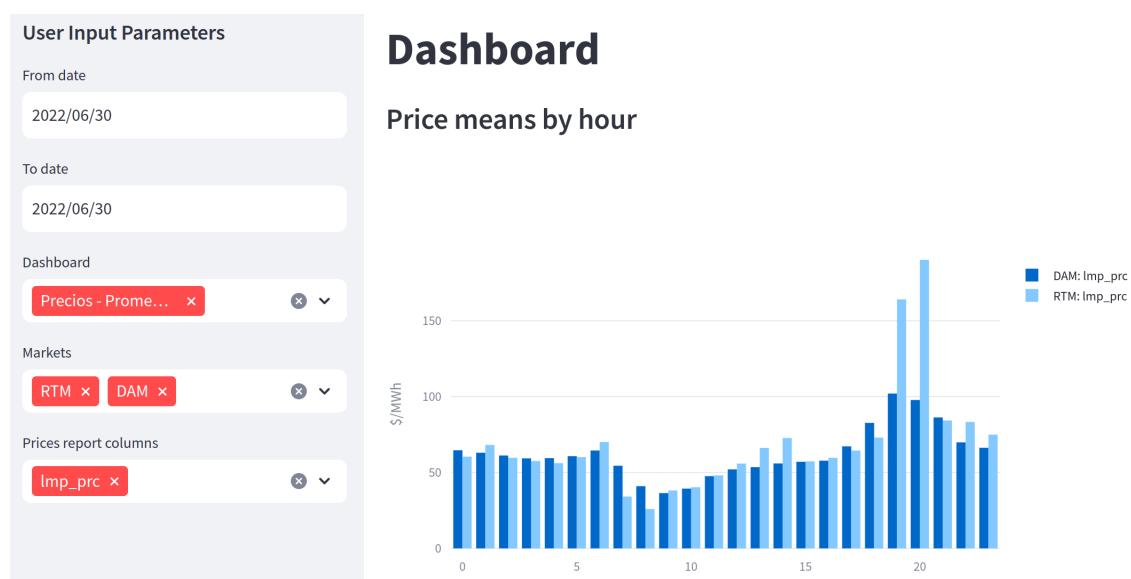


Fig. 5.8: Precios del RTM vs DAM 30 de junio del 2022.

Hay otra observación a destacar que también ocurrió en las 3 franjas de fechas y está muy relacionado al anterior, y es que **la cantidad de energía total vendida al RTM por parte del MIP es mayor a la que es vendida al utilizar las**

**heurísticas.** Durante el invierno el MIP vende un promedio de 2.33 MWh, mientras que las heurísticas tienen un promedio de 1.33 MWh y 1.91 MWh respectivamente. En el verano, el promedio del MIP es de 15.97 MWh, mientras que las heurísticas venden 14.29 MWh y 11.23 MWh. Y en otoño la diferencia es de 10.92 MWh contra 6.71 y 5.52 MWh. Esto nos indica que el MIP hace un mayor uso del mercado RTM.

### 5.9.5. Verano

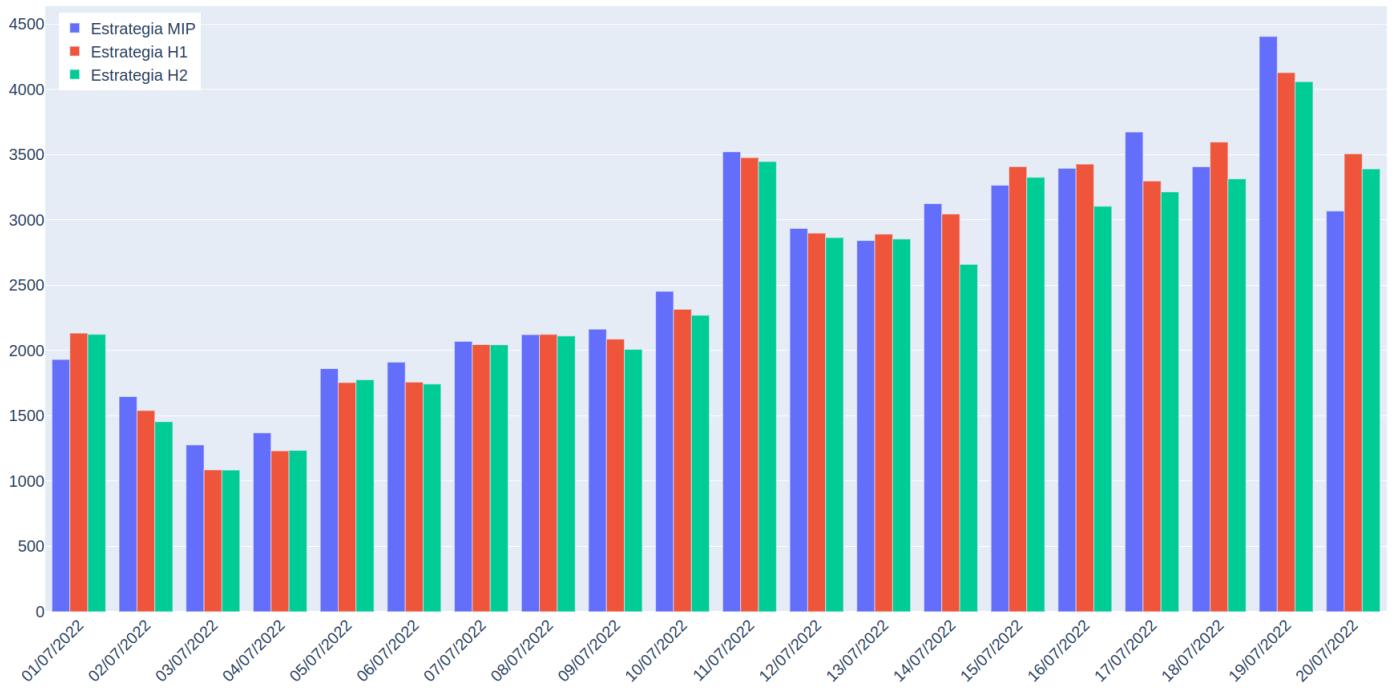


Fig. 5.9: Ganancia neta de las distintas estrategias durante el verano.

Como ya dijimos, la diferencia en promedio entre las ganancias netas durante estos días es de un 2.51% a favor del MIP si la comparamos con la heurística **H1**, y de un 5.5% si se la compara con la heurística **H2**. Haciendo un análisis más de día por día, vemos que de los 20 días en los cuales realizamos el experimento, el MIP tiene una mayor ganancia neta que la heurística **H1** en 14, obteniendo una diferencia máxima del 17.58% el 3 de julio. Comparando contra la heurística **H2**, el MIP tiene más ganancias en 16 de los 20 días, con un máximo de 17.72%, también el 3 de julio.

Las ganancias netas totales durante los 20 días es de USD 52,454.9 para el MIP, de USD 51,766.36 para la heurística **H1** y de USD 50,095.92 para la heurística **H2**. Lo cual nos da una diferencia de un 1.32% y de un 4.7% a favor del MIP respectivamente.

El promedio de energía recibida es de 30.16 MWh. Esta energía es mayormente utilizada por el MIP y las heurísticas. Hay solo dos días donde tenemos energía sin uso, el 3 y 4 de julio.

Al recibir tanta energía por parte de los paneles, tanto las heurísticas como el MIP pueden cumplir con el mercado DAM, lo cual se ve reflejado en el bajo nivel de multas que tienen. La heurística **H1** no tiene multas, la **H2** tiene solo 4 días de multas que no pasan los 200 USD. El MIP tiene 13 días con multas, pero de un monto bajo, promediando los USD 49.9.

El total de energía cargada en la batería es algo mayor en el MIP, con una media de 19.12 MWh, superando a las heurísticas **H1** y **H2**, que cargaron 14.99 MWh y 18.42 MWh. Esto se ve reflejado en la cantidad de intervalos donde cada estrategia cargó estas baterías. El MIP cargó sus baterías en promedio durante 91 intervalos, mientras que las

heurísticas lo hicieron durante 55 y 81 intervalos. Esto nos da un tiempo de carga de 7:35hs para el MIP, 4:35hs para **H1** y 6:45hs para **H2**.

### 5.9.6. Invierno

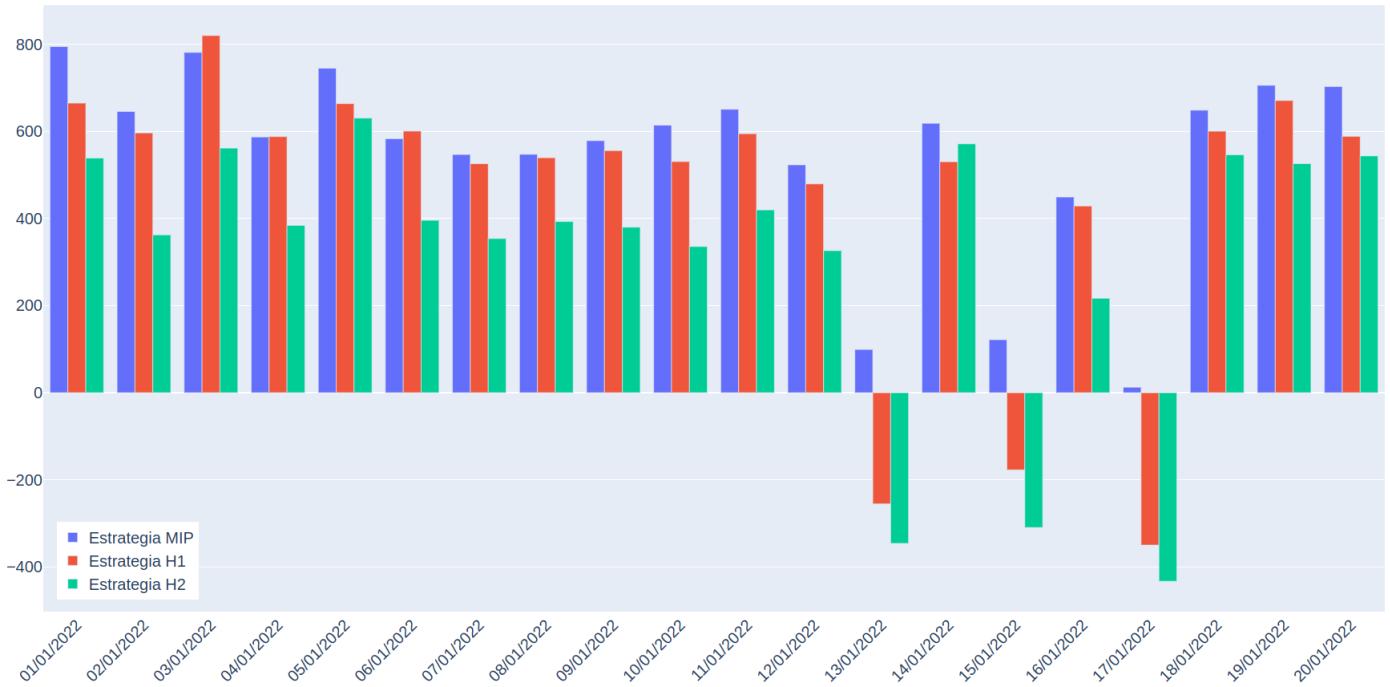


Fig. 5.10: Ganancia neta de las distintas estrategias durante el invierno.

Al tratarse de pleno invierno donde los días son más cortos, la energía recibida por los paneles es mucho menor. En promedio los paneles reciben 7.7 MWh en todo el día, apenas un 25 % de la energía recibida en los días de verano.

Esto hace que las heurísticas, a pesar de las adaptaciones mencionadas en la sección 5.9.2, tengan muchas multas por no poder cumplir con el DAM. Las multas tienen un promedio diario de 70 USD y 160 USD para las heurísticas **H1** y **H2** respectivamente. Sin embargo, debido a tener una estrategia más flexible, el MIP tiene una media diaria de 15 USD de multa.

En cuanto a las ganancias netas, el MIP es superior a la heurística **H1** en 17 de los 20 días y es superior a la **H2** en todos los días del experimento.

Se observa que en las 3 estrategias los ingresos son mucho menores comparado a los del verano. Esto se debe a que la energía recibida es mucho menor y también a que los precios de energía son más bajos en esta época del año. Sumado esto a las altas multas que presentan las heurísticas, lleva a las mismas al extremo de perder dinero en algunos días. Como se puede observar en el gráfico, esto ocurre para ambas heurísticas en las fechas 13, 15 y 17 de enero. El MIP, en cambio, presenta ganancias a lo largo de los 20 días.

Quitando estos días y quedándonos solo con las fechas donde las heurísticas obtienen un resultado positivo, el MIP también tiene una diferencia notoria, con un promedio de 7.66 % contra la heurística **H1** y del 48 % para la heurística **H2**.

La ganancia total del MIP es de USD 10,955, mientras que la de **H1** es de USD 9,195 y la de **H2** es de USD 6,396.

En este caso no hay días donde haya quedado energía sin uso. Esto es lógico, pues al no haber ningún día donde los paneles reciban 15 MWh, las baterías nunca llegan a estar completas.

En cuanto al uso de la batería, el MIP es quien más la carga, con un promedio de 6.45 MWh por día, mientras que **H1** tiene un promedio de 6.16 MWh y **H2** uno de 5.8 MWh. Llamativamente, esto no se refleja en la cantidad de intervalos de carga, donde el MIP tiene un promedio de 91, mientras que la heurística **H2** tiene un promedio de 100.

El MIP es quien tiene más intervalos de descarga, lo cual tiene sentido si se observa que cargó (y descargó) más energía a la batería que las heurísticas.

### 5.9.7. Otoño

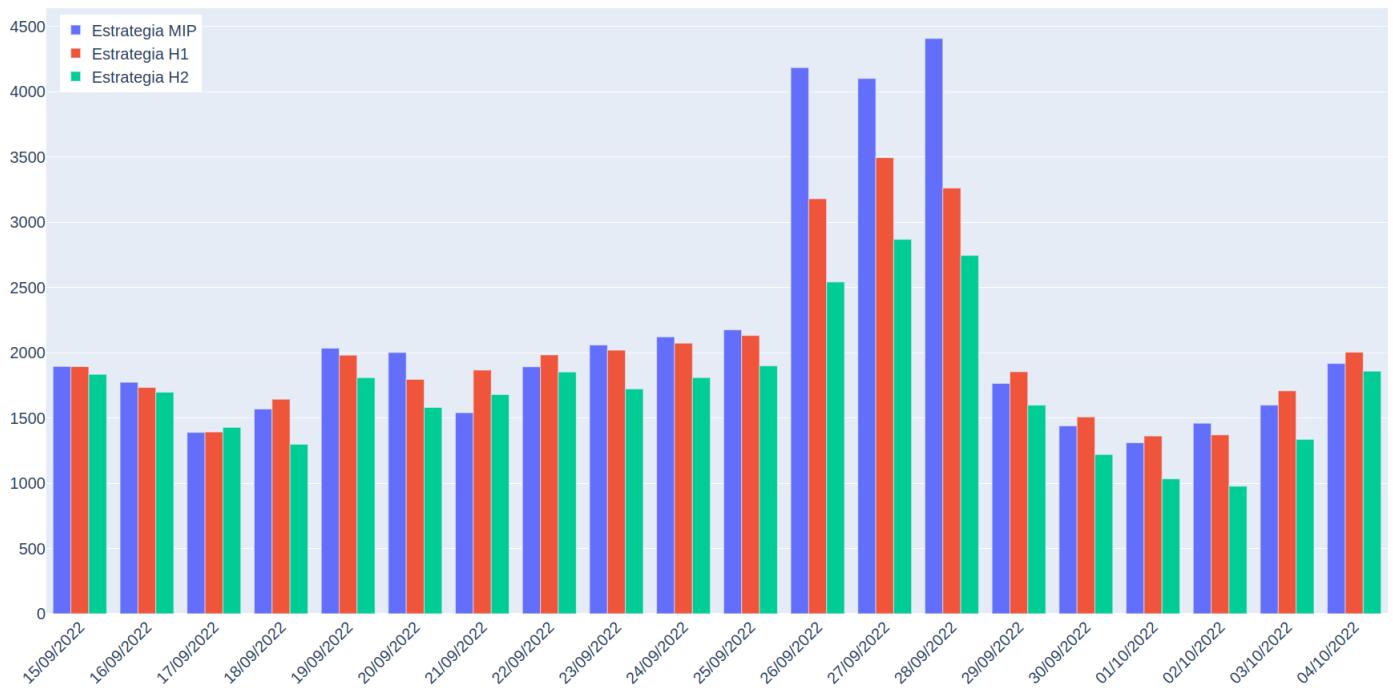


Fig. 5.11: Ganancia neta de las distintas estrategias durante el invierno.

Podemos observar que la energía recibida en estos días es de 21.71 MWh, en promedio. Esto nos sitúa por debajo de la energía recibida en verano pero por encima de la recibida en invierno.

En cuanto a la ganancia neta de las estrategias, el MIP fue superior a **H1** en 11 de los 20 días, y fue superior a **H2** en 18. En promedio el MIP fue superior a **H1** en un 3.14 %, y a **H2** en un 20.29 %. El día donde el MIP saca una mayor diferencia es el 26 de septiembre, donde tiene una diferencia del 31 % con **H1** y del 64 % con **H2**. El día donde el MIP tiene una menor ganancia, es el 21 de septiembre, donde pierde contra **H1** por un 17 % y contra **H2** por un 8 %.

El total de las ganancias netas del MIP a lo largo de los 20 días fue de USD 42,627, mientras que la heurística **H1** obtuvo una ganancia de USD 40,249 y **H2** una ganancia de USD 34,783. Esto significa una diferencia de casi el 6 % a favor del MIP contra **H1** y de un 22 % contra **H2**.

Al igual que en el verano la heurística **H1** no tiene multas, ya que siempre llega a cargar las baterías en su totalidad. El MIP y la heurística **H2** tienen multas pero son relativamente bajas. Estamos hablando de una media de USD 16.4 para el MIP y una de USD 143 para **H2**.

La energía total cargada en la batería es mayor en **H2**, que tiene una media de 16.18 MWh, que en el MIP, con una media de 15.89 MWh. Esto se ve reflejado en la cantidad de intervalos de carga de la batería, donde el MIP tiene un promedio de 89.75 y la

heurística **H1** de 96.25. La heurística **H1**, hace un menor uso de la batería, cargando siempre 14.99 MWh en un promedio de 65.25 intervalos.

### 5.9.8. Cierre del capítulo

El objetivo de esta última etapa del trabajo, plasmada en este capítulo, fue **validar la efectividad de usar un sistema para determinar una estrategia de venta de energía** que pueda ser utilizada en una planta real que participa del mercado de California y que haga el mejor uso posible de sus paneles y baterías maximizando las ganancias.

**Generamos un framework que utiliza las estimaciones de precios y clima y nuestro modelo de programación lineal entera** para generar una estrategia. Para lograr una estrategia más robusta, este sistema genera 40 MIP diferentes y los prueba a lo largo de 1000 muestras tomadas sobre el *input* para quedarse con el que tenga un mejor desempeño general.

Luego probamos nuestro *framework* ejecutándolo junto a otras heurísticas y comparando sus resultados. Para hacer esta comparación fue necesario poder realizar simulaciones donde se ejecuten las distintas estrategias y se tomen distintas métricas a comparar.

Pudimos hacer estas comparaciones a lo largo de distintas épocas del año, tomando métricas como la ganancia neta de cada estrategia, y el uso de la energía de las mismas.

**La experimentación tuvo resultados positivos**, donde nuestro sistema generó estrategias que tuvieron ganancias mayores a las heurísticas en todas las épocas del año en las que fueron comparadas.

## 6. CONCLUSIÓN

A lo largo del trabajo, adquirimos conocimiento sobre el mercado de energía de California y a partir de datos recopilados logramos generar un sistema que brinde estrategias para la utilización de paneles solares y baterías que maximicen la ganancia. Sometimos nuestro sistema de decisiones a diversos experimentos, comparándolo con dos heurísticas establecidas, una de las cuales es la utilizada actualmente por nuestra empresa de referencia. En todas las situaciones evaluadas, nuestro modelo superó consistentemente a estas heurísticas, en algunos casos con márgenes de diferencia significativos.

Más allá de la superioridad en términos de margen de ganancia, la utilización de un modelo matemático como el nuestro presenta ventajas adicionales, particularmente en términos de flexibilidad. Este modelo puede adaptarse con agilidad a cambios en diversos factores que influyen en el problema, tales como fluctuaciones climáticas o variaciones en los precios. Durante los experimentos, que abordaron situaciones variadas, pudimos confirmar la robustez de nuestro modelo. Fue notoria su capacidad para adaptarse eficientemente a diferentes escenarios. En contraste, las heurísticas, con su naturaleza más rígida, mostraron limitaciones, especialmente en condiciones cambiantes, como las que replicamos en experimentos situados en fechas invernales. En tales contextos, las heurísticas no lograron adaptarse con la misma efectividad que nuestro modelo, resultando en desempeños inferiores.

En resumen, este trabajo no solo confirma la eficacia de nuestro sistema de decisiones en comparación con heurísticas tradicionales, sino que también destaca la relevancia y necesidad de contar con herramientas flexibles y adaptables en un mercado de energía tan dinámico y en constante cambio como el de California. Es imperativo seguir innovando y adaptándose para enfrentar los desafíos futuros y maximizar las oportunidades en el ámbito de las energías renovables.

## 7. TRABAJO FUTURO

A lo largo de esta tesis se optó por simplificar ciertas cuestiones para no complejizar tanto el problema. A continuación mencionamos posibles aspectos donde se podría profundizar.

**Mejorar el predictor de precios:** Tener un modelo predictor de precios que tenga un excelente desempeño no fue un foco principal de este trabajo, ya que el objetivo fue poder cubrir el flujo completo desde la obtención y visualización de datos hasta el diseño de la estrategia final. Sin embargo, tener un predictor de precios con un alto grado de precisión podría impactar fuertemente en los resultados de nuestro sistema de decisiones. Existen varios enfoques en los que se podría trabajar. Por un lado, se podría agregar información al predictor existente, aprovechando la flexibilidad de *Prophet* ya mencionada en este trabajo. También se podría probar otra alternativa aprovechando el hecho de que los precios del día anterior sean siempre parecidos a los precios a predecir, e intentar predecir solo la diferencia entre estos dos días, en vez de querer predecir los precios absolutos del día. Otro enfoque sería no utilizar modelos de series de tiempo para esto y utilizar otras técnicas.

**Contemplar casos en que ofertas de venta de energía no son aceptadas:** En esta tesis, se trabajó con un caso de estudio que realizaba ofertas de tipo *price taker* y no se contempló la posibilidad de realizar ofertas a un precio mayor corriendo el riesgo de que dichas ofertas no sean aceptadas.

**Utilización de un modelo de programación lineal en forma operativa:** Nuestro sistema de decisiones se basa en predicciones para diseñar las estrategias. En los casos en los que las predicciones distan mucho de lo real, podría pasar que la estrategia no sea la mejor. Algo que podríamos hacer es utilizar otro modelo de programación lineal cuando se disponen de los datos reales para ver si la estrategia se puede modificar de forma de obtener más ganancias. Por ejemplo, podría ejecutarse un nuevo modelo una vez que se disponen los precios del mercado DAM y se tienen datos concretos sobre la radiación solar y nubosidad del día. Este modelo podría recalcular la estrategia con estos datos para contar con una mayor ganancia. Para esto habría que hacer modificaciones al modelo agregando en su input las ofertas ya aceptadas por el DAM y agregando la posibilidad de pagar multas en caso de incumplir con dicho mercado.

## BIBLIOGRAFÍA

- California ISO.* URL: <http://www.caiso.com/Pages/default.aspx>.
- California ISO - Markets training.* URL: <https://www.caiso.com/participate/Pages/LearningCenter/Markets.aspx>.
- CPLEX for Python users.* 25 de nov. de 2021. URL: <https://www.ibm.com/docs/en/icos/20.1.0?topic=apis-cplex-python-users>.
- CPLEX User's Manual.* 25 de nov. de 2021. URL: <https://www.ibm.com/docs/en/icos/20.1.0?topic=cplex-users-manual>.
- Familia de máquinas de uso general para Compute Engine | Documentación de Compute Engine.* Google Cloud. URL: <https://cloud.google.com/compute/docs/general-purpose-machines?hl=es-419>.
- Hong, Y.-Y. y C.-Y. Hsiao. «Locational marginal price forecasting in deregulated electricity markets using artificial intelligence». En: *IEE Proceedings - Generation, Transmission and Distribution* 149.5 (1 de sep. de 2002), págs. 621-626. ISSN: 1359-7051. DOI: [10.1049/ip-gtd:20020371](https://doi.org/10.1049/ip-gtd:20020371). URL: [https://digital-library.theiet.org/content/journals/10.1049/ip-gtd\\_20020371](https://digital-library.theiet.org/content/journals/10.1049/ip-gtd_20020371).
- OASIS Prod.* URL: <http://oasis.caiso.com/mrioasis/logon.do>.
- OASIS User-Interface Application Version 7.0.0: Deployment Scheduled on 2/19/19.* URL: <https://www.caiso.com/Documents/OASISUser-InterfaceApplicationVersion700-DeploymentScheduled-021919.html>.
- Prophet.* Prophet. URL: <http://facebook.github.io/prophet/>.
- Sheffrin, Anjali. «Empirical Evidence of Strategic Bidding in the California ISO Real-time Market». En: *Electricity Pricing in Transition*. Ed. por Ahmad Faruqui y B. Kelly Eakin. Topics in Regulatory Economics and Policy Series. Boston, MA: Springer US, 2002, págs. 267-281. ISBN: 9781461508335. DOI: [10.1007/978-1-4615-0833-5\\_18](https://doi.org/10.1007/978-1-4615-0833-5_18). URL: [https://doi.org/10.1007/978-1-4615-0833-5\\_18](https://doi.org/10.1007/978-1-4615-0833-5_18).
- Streamlit Docs.* URL: <https://docs.streamlit.io/>.
- Taylor, Sean J. y Benjamin Letham. *Forecasting at scale.* e3190v2. PeerJ Preprints, 27 de sep. de 2017. DOI: [10.7287/peerj.preprints.3190v2](https://doi.org/10.7287/peerj.preprints.3190v2). URL: <https://peerj.com/preprints/3190>.
- Tong, S., Handa Yangy y W. Torre. En: *Energy Storage System Dispatching Optimization in Stacked Applications for Utility Grid*. 2018. URL: <https://www.semanticscholar.org/paper/Energy-Storage-System-Dispatching-Optimization-in-Tong-Yangy/433a90fedbc0a84ad9cb97240d4a08ab8c08b2c1>.
- Weather Data & Weather API | Visual Crossing.* URL: <https://www.visualcrossing.com/>.
- Wolsey, Laurence. *Integer Programming.* 1.<sup>a</sup> ed. Wiley, 20 de oct. de 2020. DOI: [10.1002/9781119606475](https://doi.org/10.1002/9781119606475).