

Consumi energetici edificio U1 dell'Università degli Studi Milano-Bicocca

2021-2022

Davide Mobilia, d.mobilia@campus.unimib.it

Marta Salvi, m.salvi32@campus.unimib.it

Sinossi

Il report tratta l'analisi di una serie temporale relativa ai consumi energetici dell'edificio U1 dell'Università degli Studi di Milano-Bicocca. Lo studio osserva l'andamento della serie, tenta di effettuare delle previsioni su consumi futuri e si pone l'obiettivo di comprenderne dinamiche e sfaccettature con il fine di ottenere informazioni utili per migliorare la gestione energetica dell'edificio in esame, assumendo una prospettiva rivolta sia all'aspetto economico che a quello ecologico.

L'analisi è stata condotta attraverso modelli SARIMA a cui viene aggiunta l'informazione relativa a un regressore esterno, e Prophet. Quest'ultimo è un modello basato su una libreria open source sviluppata da Facebook e rilasciata nel 2017 in grado di gestire multi-stagionalità e *outlier* in maniera alternativa rispetto ai modelli classici.

Parole chiave

Serie storica, SARIMAX, Prophet, Milano-Bicocca, Energia, Temperatura.

1. Introduzione

Il tema dell'energia è sempre più centrale nel dibattito economico e politico. Il Nord Stream 2, sviluppi tecnologici delle energie rinnovabili, i crescenti prezzi dell'energia sono temi quotidianamente riscontrabili nelle testate giornalistiche. È indubbio che uno dei grandi problemi del futuro sia la gestione del fabbisogno energetico mondiale dalla sua generazione, possibilmente in modo sostenibile, agli utenti finali, i quali si presentano in numero inesorabilmente crescente. Non sprecare le risorse a disposizione è una componente fondamentale del cittadino responsabile e, per estensione, del bene pubblico ben gestito.

Il lavoro seguente pone come obiettivo l'analisi dei dati dei consumi energetici di un edificio (U1) dell'Università degli Studi di Milano-Bicocca. Esso rappresenta una tipologia di immobile pubblico dove è possibile intervenire per modificarne l'impronta energetica, contrariamente a stabilimenti più critici che hanno più sofisticate esigenze di climatizzazione. In particolare, si studiano i trend e i comportamenti stagionali del fabbisogno energetico, anche in relazione ad agenti esogeni, come possono essere le diverse condizioni meteo a cui l'edificio è esposto. È stato scelto l'edificio U1 in quanto esso offre un sistema di raffrescamento convenzionale (non teleraffrescamento), il cui utilizzo è quindi riscontrabile sui dati a disposizione.

2. Obiettivo/problema affrontato

L'analisi è basata sulla serie storica relativa ai consumi energetici dell'edificio U1. Sulla base di essi vengono identificati ed isolati eventuali trend e ciclicità che potrebbero essere di interesse per lo studio. Si vuole comprendere se andamenti e stagionalità possano essere giustificati da agenti esterni, in particolare da elementi atmosferici come la temperatura.

Queste relazioni potrebbero tornare utili per una pianificazione finanziaria efficace relativa ai consumi di energia elettrica all'interno del complesso universitario Milano-Bicocca che, per le sue dimensioni e complessità strutturali, può contenere delle inefficienze.

Inoltre, comprendere come variano i consumi in base all'andamento meteo può essere un'informazione utile anche per il fornitore di energia elettrica, che deve saper gestire e possibilmente prevedere il carico energetico della rete¹.

3. Aspetti metodologici

La serie in esame è caratterizzata da andamenti che si ripetono con la stessa magnitudine per diversi cicli temporali. Non si hanno quindi amplificazioni di trend proporzionali al tempo. Per questo motivo è stata applicata una scomposizione additiva che divide la serie in tre componenti (stagionalità, trend e residui) che, se sommate, riportano alla serie originale. È quindi possibile isolare la componente stagionale annuale dei consumi che, secondo le assunzioni di partenza, dovrebbe rispecchiare quella del meteo.

Per studiare la stazionarietà sono state utilizzate diverse funzioni e test. Inizialmente sono stati considerati i *plot* di ACF (*AutoCorrelation Function*) e PACF (*Partial AutoCorrelation Function*)² per la verifica di eventuali andamenti stagionali. Per quanto riguarda i test, si è optato per l'A.D.-Fuller³ la cui ipotesi nulla afferma l'esistenza di una radice unitaria all'interno della componente autoregressiva, per cui la serie si può definire non stazionaria. Per ogni radice esistente sarà necessaria una differenziazione per riportare i dati ad un livello di stazionarietà.

3.1 SARIMA e SARIMAX

Si è tentato di utilizzare un modello SARIMA(X)⁴, che permette di cogliere componenti *AutoRegressive* e di *Moving Average* (anche in combinazione stagionale) di processi integrati. La componente regressiva esterna X del modello è necessaria per verificare la dipendenza della variabile principale dai dati meteorologici.

La principale difficoltà riscontrata nell'utilizzo di SARIMA(X) è l'incapacità di gestire le multi-stagionalità. Per questo motivo, come spiegato nel *Paragrafo 5.2*, sono stati considerati dati aggregati settimanalmente così da eliminare stagionalità⁵ di livello inferiore (giornaliera ed oraria). I dati forniti sono i consumi espressi in potenza (kW) ogni 15 minuti. Si è quindi deciso di aggregarli usando la media dei livelli di assorbimento energetico offerti dalle osservazioni. In questo modo si sono ottenute potenze medie settimanali utilizzate poi per i modelli SARIMA e SARIMAX. Qualora sorgesse la necessità di convertire i livelli di potenza settimanale in consumo energetico settimanale (kWh), occorre moltiplicare la potenza per il numero di ore in una settimana.

Come metrica è stato considerato l'indice AIC, o *Akaike's Information Criterion*, che fornisce una misura della qualità di stima di un modello

statistico sulla base di una funzione di verosimiglianza, tenendo conto sia della bontà di adattamento sia della complessità del modello. Questo indice è stato scelto come unica metrica per confrontare i modelli. Tuttavia, come espansione del lavoro, sarebbe bene considerare una forma di giudizio più articolata, con più metriche in combinazione.

I residui⁶ del modello sono stati studiati attraverso quattro grafici:

- Il grafico dei residui standardizzati, che rappresenta i residui del modello divisi per una stima della deviazione standard del cosiddetto *white noise* ovvero la componente dell'errore che, essendo randomica, non può essere predetta;
- L'istogramma dei residui e la densità stimata, che fornisce informazioni riguardo la corretta costruzione del modello. Nella situazione ideale, i residui si distribuiranno come una Gaussiana con media 0 e varianza 1, $N(0,1)$;
- Il *Q-Q plot*, trasposizione grafica dei quantili della distribuzione, che deve essere idealmente lineare;
- Il correlogramma, equivalente all'ACF applicato ai residui del modello.

Per valutare la bontà complessiva dei modelli è stato poi calcolato l'errore medio assoluto percentuale MAPE (*Mean Absolute Percentage Error*), misurando quindi l'accuratezza di previsione, la cui formula è:

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|$$

Con A_t valore reale, F_t valore previsto e n il numero di punti stimati

3.2 Prophet

Prophet⁷ è un modello per la previsione di *time series* basato su metodi regressivi composti da una curva di crescita del trend e da componenti stagionali modellate con sviluppi in serie di Fourier. I modelli di Prophet riescono a gestire osservazioni orarie, giornaliere o settimanali purché il loro numero sia elevato. La capacità di lavorare in maniera automatica con dati mancanti e *outlier* è una peculiarità di questo modello. Esso permette di utilizzare i dati senza doverli pulire. Tali caratteristiche hanno permesso di utilizzare dati aggregati su base giornaliera e non più

settimanale, permettendo inoltre di valutare le multi-stagionalità della serie⁸.

4. I dati

Il dataset fornito contiene osservazioni rappresentanti i consumi energetici degli edifici denominati U1 ed U6 dell'Università degli Studi di Milano-Bicocca relativi agli anni 2018, 2019 e 2020, divisi mensilmente e campionati ogni 15 minuti.

Dato che l'edificio U6 è provvisto di teleraffrescamento, le analisi del progetto si sono rivolte esclusivamente all'edificio U1 che presenta invece un più tradizionale impianto di condizionamento elettrico.

I dati, come forniti, sono eterogenei nel loro formato (.csv, .xlsx, .xltx, .xls). Il primo passo in preparazione all'elaborazione è uniformarli in formato .xls per una più facile assimilazione automatica da parte dei software utilizzati per l'analisi, ad esempio la libreria *pandas*.

Per poter svolgere l'analisi, inoltre, sono stati integrati con un dataset che raccoglie alcune informazioni meteorologiche. Esso proviene da un database *PostgreSQL* che raccoglie registrazioni meteo di numerose città nel mondo. Tale raccolta è alimentata dall'API pubblica di *Openweathermap* (OWM)⁹. Le richieste all'*endpoint* di questa API vengono effettuate in modo tale che ogni città considerata presenti un rilevamento cadenzato tra i 12 e i 18 minuti, intervallo non costante per motivi di limitazione API. Dal database sono state prelevate le osservazioni relative alla città di Milano per gli anni presi in considerazione ed esportate in formato .csv. L'ipotesi è che le informazioni contenute in questo dataset possano ben rapportarsi con una profilazione energetica degli edifici pubblici.

Il dataset meteorologico offre, in dettaglio, le seguenti variabili esplicative:

- *t*: orario della rilevazione;
- *temperature*: temperatura istantanea rilevata dai sensori a cui *Openweathermap* fa riferimento;
- *humidity*: umidità istantanea in percentuale;
- *sunrise/sunset*: orario di alba e tramonto;
- *status*: fattore indicante la condizione meteo al momento della rilevazione (*Clear, Cloudy, Thunderstorm, Windy, ...*).

Viene quindi analizzato congiuntamente al dataset dei consumi, del quale si considerano le colonne *DATA*, *ORA* e *CONSUMO_ATTIVA_PRELEVATA*, quest'ultima esprimente il consumo in potenza dell'edificio U1.

5. Analisi/Processo di trattamento dei dati

Il codice per accorpare i dati è stato sviluppato in modo tale da rendere possibile la parallelizzazione delle operazioni *Input/Output* qualora ipoteticamente si volesse estendere il progetto per integrare dati precedenti al 2018. Per procedere all'analisi del dataset si è sviluppato un software scritto in *Python3*, talvolta processato in un *Jupyter notebook*. Quest'ultimo consente di integrare codice di programmazione e testo *Markdown/LaTeX*, aiutando il processo di prototipazione.

Per analizzare la serie temporale presa in esame con questo linguaggio di programmazione è necessario utilizzare librerie e funzioni esterne. Le due librerie principali sono state *statsmodels* e *Prophet*. La prima mette a disposizione classi e funzioni per la stima di vari modelli, tra cui *SARIMAX*, e, per ogni stimatore, offre una varietà di metriche. La seconda libreria principale che abbiamo scelto di usare, *Prophet*, permette di eseguire stime con modelli regressivi basati su serie di Fourier.

Le altre librerie utilizzate sono:

- *matplotlib* per rappresentare in forma grafica la serie;
- *pandas* per la gestione dei dati;
- *numpy* mette a disposizione funzioni matematiche di alto livello per poter operare efficientemente su strutture tipo array;
- *tqdm* è servita per rappresentare una barra di progressione raffigurante l'avanzamento dei cicli iterativi in fase di processing dati;
- *itertools*, in particolare *product*, è stata usata per calcolare il prodotto cartesiano di un iterabile;
- *Prophet.diagnostics* utile per eseguire le *cross-validation* e per ottenere metriche di performance;
- *fbprophet*, in particolare *plot*, utilizzato per le visualizzazioni.

5.1 Preprocessing

Le colonne *DATA* e *ORA* presenti nei file relativi ai consumi energetici sono state raggruppate in

un'unica colonna di tipo *datetime* chiamata *T* e, usando quest'ultima come riferimento temporale, è stato effettuato un campionamento medio orario con *forward-fill* in presenza di valori non validi. Inoltre, i dati meteo relativi alle precipitazioni e all'orario di alba e tramonto sono stati rielaborati per una maggiore fruibilità a livello di analisi. Sono state generate le colonne *is_raining*, che esprime la percentuale di rilevazioni marcate come piovose nella variabile status sul totale delle misurazioni accorpate in un'ora, e *ore_luce*, un valore intero che rappresenta le ore di luce effettive in una giornata.

Per quanto riguarda la gestione degli *outlier* si è tentato inizialmente lo studio dei singoli episodi considerati anomali. Per identificare tali osservazioni anomale, come primo passo si è provato a generare un *box plot* costruito su una finestra *rolling* di 7 giorni. Questo approccio, tuttavia, non si è rivelato efficace poiché i dati anomali talvolta cadevano molto vicini e, di conseguenza, non venivano identificati come tali. Alcuni valori anomali potevano essere dovuti a festività che però in differenti anni portavano solo a cambiamenti di trend, senza risultare *outlier* veri e propri.

Si è deciso dunque di ricorrere all'algoritmo di previsioni per serie storiche, Prophet¹⁰. Questo è considerato robusto agli *outlier* e in grado di identificare eventuali valori anomali, modellando la serie al netto di essi. La sua caratteristica è che esegue una stima anche per i dati utilizzati per il training. Su queste stime viene costruito un intervallo di confidenza all'80% e lo si utilizza per identificare e gestire gli *outlier*. Nel caso in cui un'osservazione cada fuori dall'intervallo di confidenza viene marcata e sostituita con il valore del limite dell'IC ad esso associato.

In questo modo si mantiene l'andamento della serie evitando quindi la rimozione di eventi caratterizzanti come ad esempio le ricorrenze festive.

5.2 Processo di SARIMA e SARIMAX

L'obiettivo dell'analisi è la modellizzazione della serie storica attraverso modelli cosiddetti classici, come SARIMA e SARIMAX.

Il primo problema incontrato è rappresentato, come accennato in precedenza, dagli *outlier*: questi modelli sono molto sensibili e non permettono una corretta modellizzazione in presenza di valori anomali. Un ulteriore elemento

di disturbo è la multi-stagionalità che i modelli classici non sono in grado di gestire. La serie, infatti, presenta una stagionalità giornaliera (di-notte), una settimanale (giorni feriali-weekend) e una annuale (estate-inverno). Questo problema è stato risolto eseguendo un campionamento medio a sette giorni tale da includere sempre un fine settimana, giorni in cui i consumi generalmente si abbassano. Così facendo, si è eliminata la stagionalità giornaliera e quella settimanale.

La libreria *statsmodels* offre il sotto modulo *seasonal_decompose*¹¹ che svolge una scomposizione della serie. Ci si è concentrati su stagionalità annuali, per le quali si ha una più diretta correlazione con il meteo.

Dopo aver indagato la stazionarietà della serie con il test di Dickey-Fuller (A.D.-Fuller), si sono cercati dei segnali di modelli AR (*AutoRegressive*) o MA (*Moving Average*) attraverso l'analisi della funzione di autocorrelazione (ACF) e di autocorrelazione parziale (PACF)³ al fine di formulare un'ipotesi riguardo possibili trend o stagionalità presenti nei nostri dati.

Tenendo in considerazione i risultati ottenuti dalle precedenti analisi, sono state testate, attraverso un algoritmo iterativo, tutte le possibili combinazioni parametriche $(p, d, q)(P, D, Q)s$ di un modello SARIMA in modo da selezionare quella che minimizza l'AIC.

La costruzione del modello è stata eseguita e testata attraverso un *holdout* delle osservazioni: questa tecnica prevede di dividere il dataset completo in due sottogruppi, training set e test set che, nel caso in esame, sono costituiti rispettivamente da 120 e 37 osservazioni (77%-23% circa).

Sono stati sviluppati i modelli SARIMA a 157 settimane e a 120 e, successivamente, quelli SARIMAX per entrambi gli orizzonti temporali per ciascuna variabile esogena presa in considerazione (temperatura, umidità, pioggia e ore di luce). Grazie ai modelli SARIMAX si può affermare se e quali regressori siano significativi per la modellizzazione e quindi utili per la previsione dei consumi di energia elettrica.

Dopo la costruzione di ciascun modello ne sono stati valutati i residui attraverso le visualizzazioni spiegate in precedenza, nel *Paragrafo 3.1*.

Infine, è stata valutata l'accuratezza previsionale dei modelli utilizzando il MAPE. In particolare, questa misura è stata applicata alla parte di

training e di test così da poter valutare al meglio il comportamento della modellizzazione e identificare eventuali casi di *under* o *overfitting*.

Una volta identificato il modello che meglio rappresenta i dati, si è proseguito con una proiezione dell'andamento della serie storica su diverse scale temporali: 3, 6 e 9 mesi.

5.3 Analisi Prophet

Prophet, a differenza di SARIMA, non ha difficoltà nel gestire più stagionalità; si è scelto quindi di utilizzare i dati aggregati giornalmente. È stato creato un modello univariato semplice e successivamente con regressore per includere le variabili esogene riguardanti il meteo. In particolare, si è scelto di considerare la temperatura essendo l'unica risultata significativa nei precedenti modelli SARIMAX.

Prophet necessita l'impostazione di alcuni parametri che definiscono la curva del trend, le stagionalità e i *change-points*⁸. Nella modellazione del trend, in particolare, occorre definire il parametro *growth* il quale ne indica la forma, lineare o logistica; nel caso di ricerca è stato impostato come lineare. Ci si aspetta, infatti, che la potenza media erogata non abbia un andamento tendente alla saturazione.

Questo algoritmo stima le stagionalità usando una somma di Fourier parziale, il numero dei termini della somma parziale corrisponde al numero di funzioni utilizzate come regressori per stimare la stagionalità. Dato che aumentando il numero di termini della somma parziale, aumenta il rischio di *overfitting*, i parametri relativi alle stagionalità sono stati impostati effettuando delle prove e selezionando quelli che mostravano le performance migliori in termini di MAPE ottenuto con la funzione *cross-validation* di Prophet. A differenza di quanto fatto per i modelli SARIMAX, questo passaggio non è stato implementato utilizzando una funzione di *parameter tuning*. La capacità computazionale a disposizione, infatti, si è rivelata insufficiente. Ne è risultato che i parametri di stagionalità *weekly_seasonality* e *yearly_seasonality* sono stati impostati rispettivamente a 3 e a 10, il loro valore di default.

La funzione di *cross-validation* è stata configurata selezionando: un periodo di training (*initial*) a 730 giorni, per comprendere due anni di stagionalità; un orizzonte da prevedere (*horizon*) che è stato impostato prima a 60, poi a 90, 180 e 270 giorni (2, 3, 6 e 9 mesi); il parametro *cut-off*

point è invece stato fissato a 30. Quest'ultimo implica che ad ogni iterazione della *cross-validation*, sia il training set che il test set si sposteranno avanti di 30 osservazioni. Impostati questi parametri la funzione restituirà delle misure di performance corrispondenti alla media di quelle ottenute ad ogni iterazione. Queste sono: errore quadratico medio (MSE), radice del MSE o varianza interna (RMSE), errore medio assoluto (MAE), errore medio assoluto percentuale (MAPE) ed errore assoluto percentuale mediano (MDAPE). Tra queste si è scelto di utilizzare il MAPE come misura di riferimento.

Inoltre, Prophet è stato utilizzato per individuare dei trend *change-points*, vale a dire i punti in cui l'andamento del trend varia in maniera rilevante. La funzione di *change-point detection* seleziona inizialmente 25 potenziali punti nei quali calcola la variazione dell'inclinazione del trend, identificando come effettivi *change-points* solo quelli in cui il valore assoluto della variazione (*delta*) è maggiore di 0.01.

6. Risultati

6.1 SARIMA e SARIMAX

6.1.1 Scomposizione serie

Come analisi della serie storica in esame è stata effettuata una scomposizione additiva, in questo modo è stato possibile isolare le caratteristiche che la compongono.

In particolare, nella *Figura 1* grazie alla funzione *seasonal_decompose* sono stati isolati il trend, la componente stagionale e la distribuzione dei residui dati.

Ci si è quindi concentrati sulla distribuzione dei residui. Come spiegato in precedenza, al fine di pulire i dati dagli *outlier* è stato utilizzato Prophet. A supporto di questa decisione, il *range* della distribuzione dei residui utilizzando il metodo dei *box plot* è molto più ampio, si è infatti passati da un intervallo intorno a $[-20, +20]$ a uno intorno a $[-10, +10]$ con Prophet (*Figura 2*).

Per il commento di trend e stagionalità si rimanda all'analisi effettuata con Prophet essendo questo in grado di produrre risultati più precisi e puntuali attraverso l'utilizzo di dati giornalieri.

6.1.2 Verifica stazionarietà, ACF e PACF

Successivamente è stato necessario capire se la serie fosse stazionaria o meno per poi operare differenziazioni tali da eliminare eventuali autocorrelazioni.

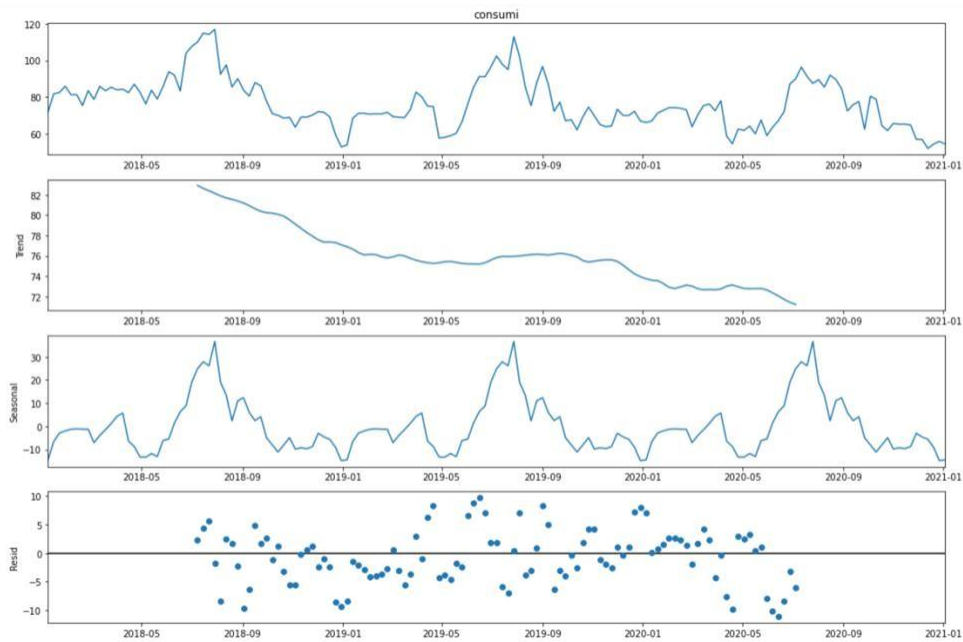


Figura 1. Analisi delle componenti della serie, gestione *outlier* con Prophet

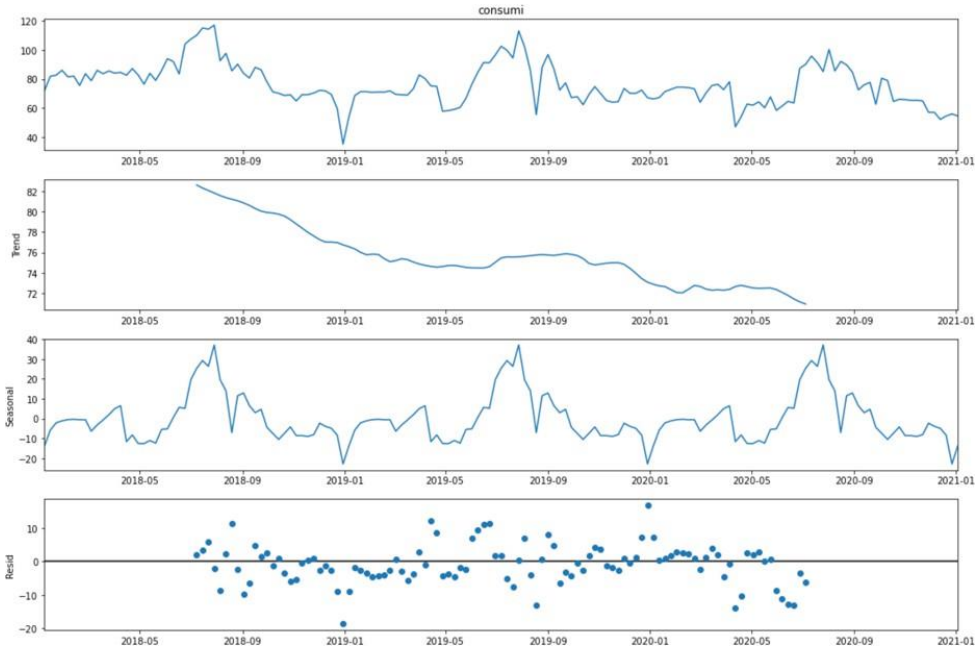


Figura 2. Analisi delle componenti della serie, gestione *outlier* con metodo dei box plot

Rappresentando i *plot* ACF e PACF si può notare una forte correlazione di primo ordine che va poi diminuendo all'aumentare dei lag. Interessante è l'andamento armonico del *plot* con un picco positivo intorno al lag 52, questo valore coincide con la ciclicità annuale data l'aggregazione settimanale dei dati. Tuttavia, questa stagionalità non risulta essere significativa, la causa è probabilmente l'ampiezza del ciclo. Sostanzialmente il valore della serie è influenzato dalle osservazioni immediatamente precedenti piuttosto che da una stagionalità a ciclo così elevato.

Nel *plot* di autocorrelazione parziale (PACF) si può notare similmente una forte correlazione di primo ordine senza però osservare ulteriori lag significativi. La decrescita quasi esponenziale osservata nel *plot* dell'ACF e la natura del PACF appena visto porta subito a pensare ad una componente autoregressiva di ordine 1.

Con un primo test A.D.-Fuller non si può confermare con un alto livello di significatività che la serie sia stazionaria, si esegue quindi una differenziazione di primo ordine ($lag=1$) e si analizzano nuovamente le funzioni di

autocorrelazione ACF e PACF. Da questi nuovi grafici non è possibile identificare chiaramente parametri *AR* e *MA* della nostra serie. Le autocorrelazioni significative sono diverse e non sembrano seguire uno schema definibile con una semplice analisi grafica dei *plot* ACF e PACF.

È stata verificata dunque la stazionarietà con un secondo test A.D.-Fuller sulla serie differenziata. In questo caso, il *p-value* molto piccolo (e^{-25}) permette di rifiutare l'ipotesi nulla e, quindi, di affermare la stazionarietà della serie trasformata.

Poiché dai *plot* precedenti non è stato possibile ricavare la totalità dei parametri del modello SARIMA è stata creata una funzione di *parameter tuning*. Nel dettaglio, si è testato un *range* di valori (estremi compresi) per i singoli parametri così formato:

- $p = \text{range}(0, 4)$;
- $d = 1$;
- $q = \text{range}(0, 3)$;
- $P = \text{range}(0, 2)$;
- $D = 1$;
- $Q = \text{range}(0, 2)$;
- $s = 52$.

Si è deciso di fissare i due parametri di differenziazione d e D per le analisi precedentemente fatte attraverso le funzioni ACF e PACF. È risultata evidente la componente *AR*(1) la cui differenziazione di primo ordine ha reso la nostra serie stazionaria. Per quanto riguarda il parametro D , nonostante la funzione ACF mostri un andamento non significativo, si è deciso di inserire una stagionalità annuale poiché è una caratteristica evidente della nostra serie.

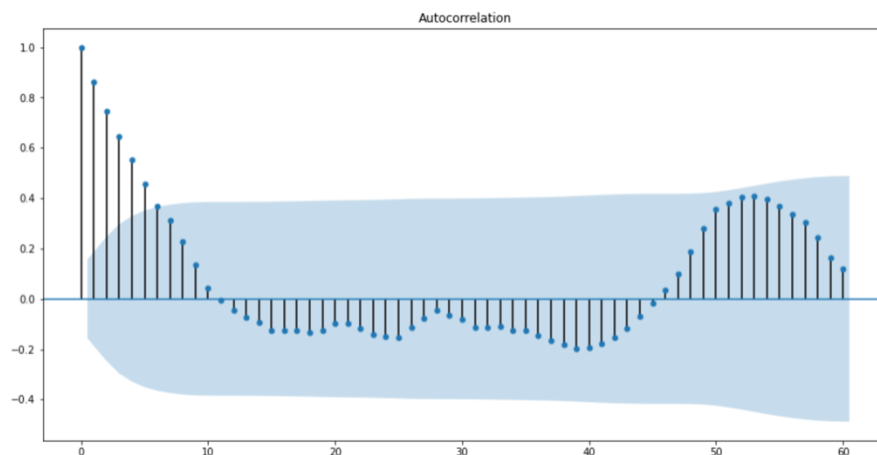


Figura 3. *Plot* autocorrelazione (ACF)

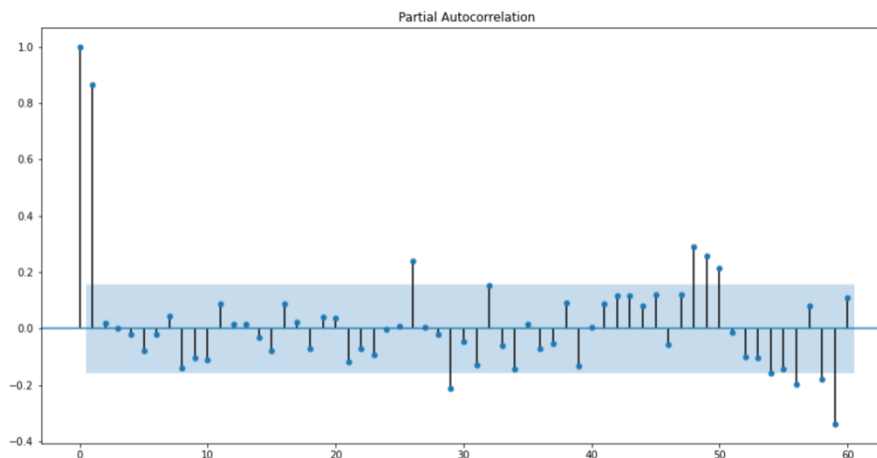


Figura 4. *Plot* autocorrelazione parziale (PACF)

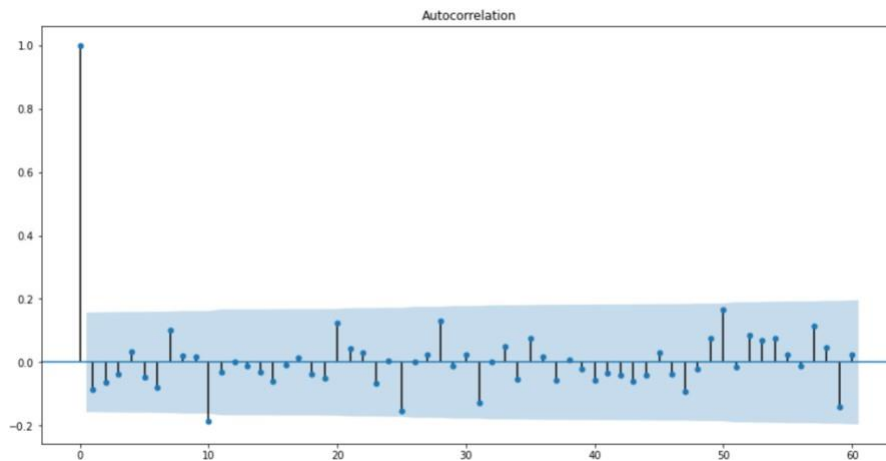


Figura 5. Plot autocorrelazione (ACF) con $lag=1$

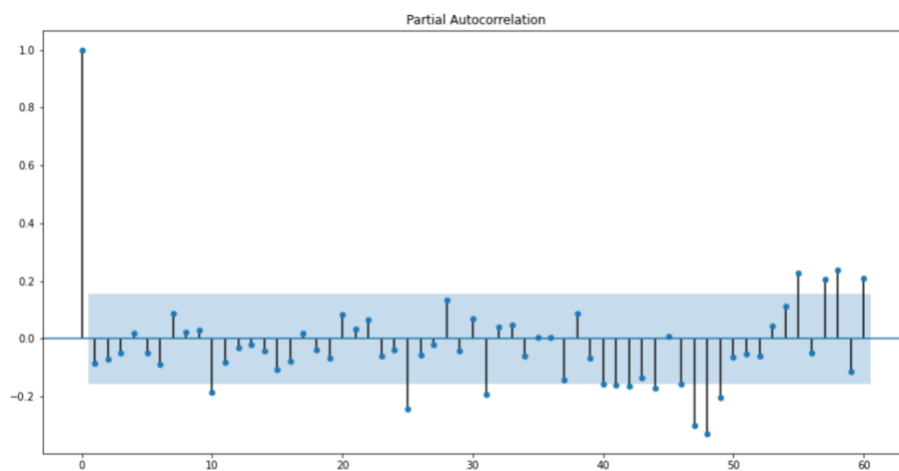


Figura 6. Plot autocorrelazione parziale (PACF) con $lag=1$

6.1.3 SARIMA

Sono stati costruiti i modelli SARIMA a 157 settimane e a 120 sia per la casistica $D = 1$ (stagionalità annuale) sia per $D = 0$. Questo è stato fatto per dimostrare che il modello con stagionalità annuale è in grado di adattarsi meglio ai dati pur considerando che dal *plot* dell'ACF questa stagionalità non risulta essere significativa.

6.1.3.1 SARIMA $D = 1$

Per quanto riguarda il modello con $D = 1$, i valori con cui è stato settato il modello hanno prodotto un AIC pari a 465.298 e sono i seguenti: $endog = consumi$, $p = 0$, $d = 1$, $q = 2$, $P = 1$, $D = 1$, $Q = 0$, $s = 52$. Questi sono stati ricavati usando le prime 120 osservazioni e poi sono stati utilizzati per i due orizzonti temporali in analisi 120 e 157 settimane, tecnica usata anche in tutte le altre modellizzazioni svolte, senza variabili esogene. I parametri del modello a 120 settimane (Figura 7) risultano essere significativi con buoni valori di p -value. Si può notare in Figura 9 che il *plot* degli errori standardizzati non ha particolari

andamenti o variazioni nel tempo e che il *Q-Q plot* rispetta una distribuzione normale dei quantili dove le code sono leggermente pesanti. Non ideale è invece la distribuzione di densità: l'istogramma rileva, infatti, una leggera asimmetria destra rispetto alla distribuzione $N(0,1)$. Inoltre, non risulta evidenza di autocorrelazioni significative. Si può quindi affermare che in generale c'è una buona distribuzione dei residui nel nostro modello.

Gli stessi parametri individuati per il modello precedente risultano anche più significativi quando utilizzati per la modellizzazione a 157 settimane (Figura 8). Come osservabile in Figura 10, il *Q-Q plot* e la distribuzione di densità di questo secondo modello risultano molto simili a quello precedente. È necessario sottolineare invece un *range* più ampio degli errori standardizzati e la presenza di autocorrelazione inspiegabile alla settimana 7.

Successivamente è stata testata la capacità previsionale del modello con un training set di 120

settimane e test sulle restanti 37. In *Figura 11* viene messa a confronto la nostra serie con il rispettivo modello SARIMA e si può notare come quest'ultimo sembri avere un adattamento sufficientemente buono ai nostri dati. Il passaggio da training a test set è evidenziato dalla barra nera verticale e in *Figura 12* è rappresentata solo la parte previsionale relativa al training set.

	coef	std err	z	P> z	[0.025	0.975]
ma.L1	-0.3890	0.115	-3.378	0.001	-0.615	-0.163
ma.L2	-0.3768	0.120	-3.153	0.002	-0.611	-0.143
ar.S.L52	-0.5340	0.225	-2.373	0.018	-0.975	-0.093
sigma2	41.0539	12.025	3.414	0.001	17.485	64.623

Figura 7. Parametri del modello SARIMA $D = 1$ a 120 settimane

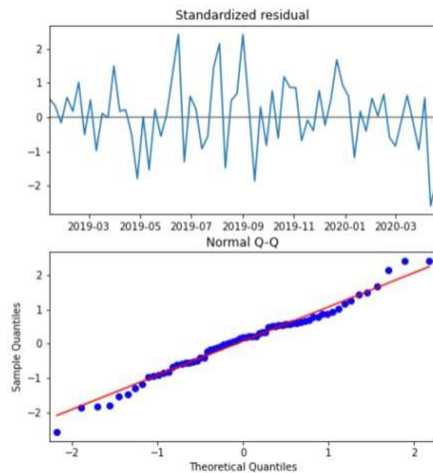


Figura 9. Plot errori del modello SARIMA $D = 1$ a 120 settimane

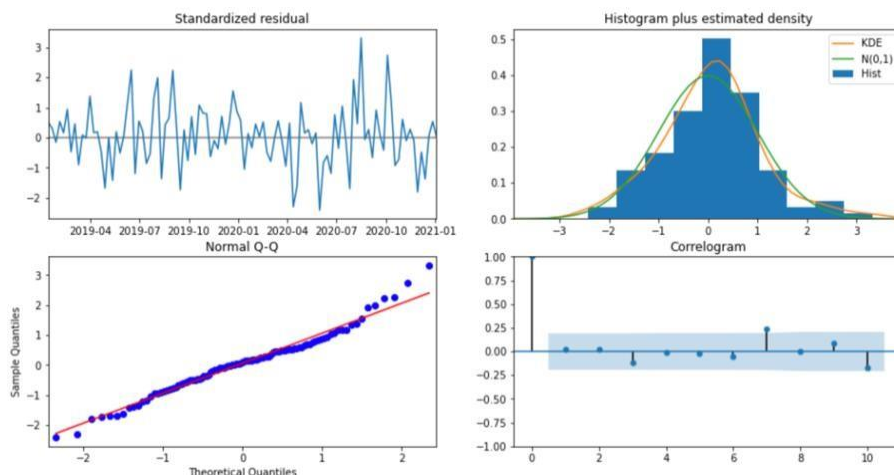
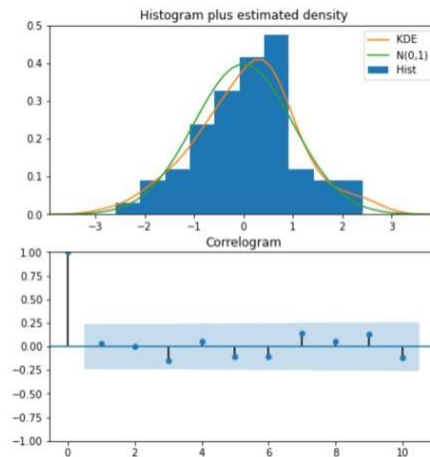


Figura 10. Plot errori del modello SARIMA $D = 1$ a 157 settimane

Per concludere l'analisi si verifica la bontà di adattamento del modello attraverso l'indice MAPE. Il valore calcolato complessivamente restituisce un MAPE di 8.24%, in particolare per la parte di training il MAPE risulta essere 8.27%, mentre per il test set 10.87%. Si può quindi affermare di avere un ottimo adattamento sia per la parte di training che per quella di test.

	coef	std err	z	P> z	[0.025	0.975]
ma.L1	-0.3971	0.108	-3.678	0.000	-0.609	-0.185
ma.L2	-0.3692	0.117	-3.164	0.002	-0.598	-0.140
ar.S.L52	-0.4673	0.116	-4.017	0.000	-0.695	-0.239
sigma2	52.2751	6.104	8.564	0.000	40.311	64.239

Figura 8. Parametri del modello SARIMA $D = 1$ a 157 settimane



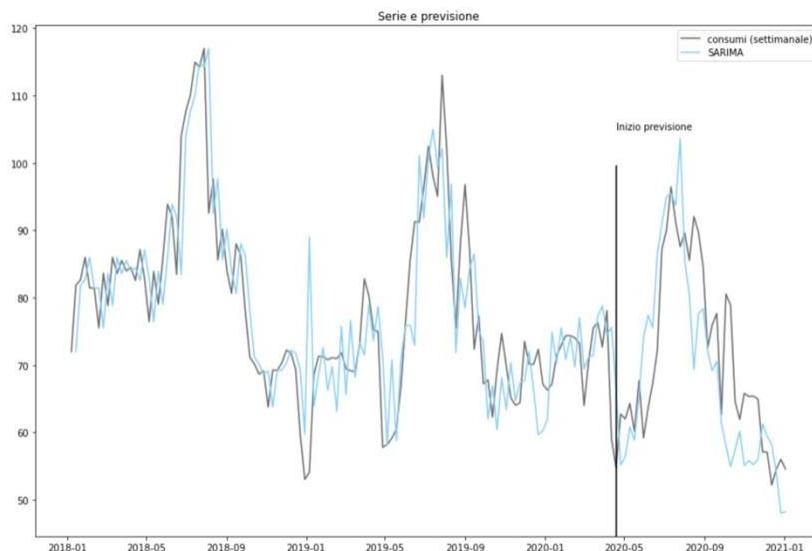


Figura 11. Serie storica a confronto con training e test set ottenuti con SARIMA $D = 1$

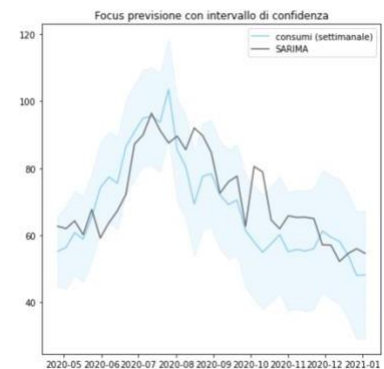


Figura 12. Particolare di Figura 11, confronto serie storica e test set ottenuti con SARIMA $D = 1$

6.1.3.2 SARIMA $D = 0$

Provando a rimuovere il parametro di differenziazione stagionale $D = 0$, in una prima analisi, il modello sembra avere un buon adattamento alla serie. Tutte le variabili risultano infatti essere significative in entrambi gli orizzonti temporali presi in considerazione (Figura 13 e 14). Tuttavia, i modelli risultanti dalla funzione di *parameter tuning* producono valori di AIC molto elevati, rispettivamente 796.3 e 1044.1 (in Appendice la distribuzione degli errori).

Valutando la capacità previsionale di questo modello, si nota come le performance siano peggiori rispetto a quello con stagionalità annuale. Le rappresentazioni grafiche della serie storica e della previsione sono infatti riconducibili a un chiaro caso di *overfitting* (Figura 15). Nella prima parte (training set) le due serie sono quasi sovrapposte, mentre nella seconda la serie in azzurro

con il modello SARIMA con $D = 0$) ha un andamento completamente differente. In Figura 16 si può vedere chiaramente come l'andamento estivo, solitamente caratterizzato da una forte crescita dei consumi venga appiattito, risultando quindi in una predizione priva di quella componente stagionale che contraddistingue la nostra serie.

Queste considerazioni, già evidenti a livello grafico, sono rinforzate dai risultati della bontà di previsione studiata attraverso l'indicatore MAPE. Nonostante il valore generale sia pari a 7.27%, quindi inferiore rispetto al modello con stagionalità annuale, andando ad analizzare separatamente training e test set la situazione cambia. Il MAPE del training set è pressoché lo stesso, 7.26%, ma il test set ha un valore molto più elevato arrivando ad un valore di 13.48% confermando l'*overfitting* di questo modello.

	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.8458	0.064	13.268	0.000	0.721	0.971
ma.L1	-0.9821	0.040	-24.787	0.000	-1.060	-0.904
ar.S.L52	0.3178	0.129	2.460	0.014	0.065	0.571
sigma2	41.8330	5.199	8.046	0.000	31.643	52.024

Figura 13. Parametri del modello SARIMA $D = 0$ a 120 settimane

	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.8494	0.057	14.843	0.000	0.737	0.962
ma.L1	-0.9843	0.036	-27.016	0.000	-1.056	-0.913
ar.S.L52	0.1750	0.100	1.745	0.081	-0.022	0.372
sigma2	44.1403	4.523	9.759	0.000	35.275	53.006

Figura 14. Parametri del modello SARIMA $D = 0$ a 157 settimane

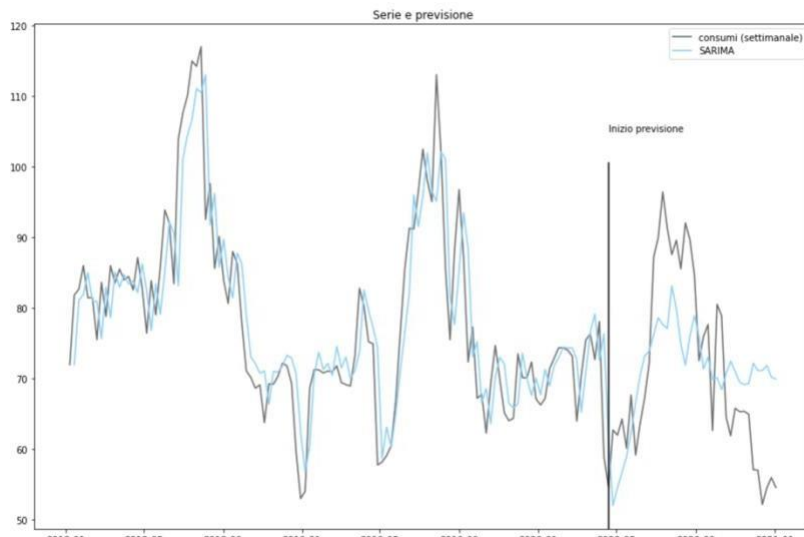


Figura 15. Serie storica a confronto con training e test set ottenuti con SARIMA $D = 0$

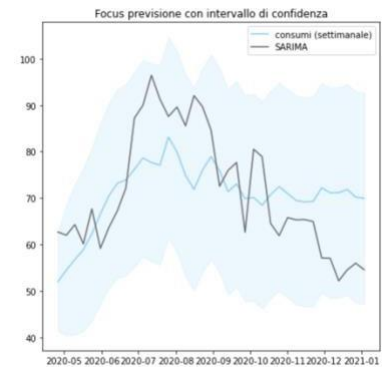


Figura 16. Particolare di Figura 15, confronto serie storica e test set ottenuti con SARIMA $D = 0$

6.1.4 SARIMAX

Si cerca ora di capire se delle variabili esogene alla serie storica dei consumi possano risultare significative. Sfruttiamo, in questo senso, i dati meteo aggiungendoli come regressori ai nostri modelli e valutiamo quindi i diversi modelli SARIMAX. Le variabili considerate nel nostro modello saranno: temperatura, umidità, pioggia e ore di luce. L'algoritmo restituirà oltre ai parametri caratteristici anche il regressore con il suo livello di significatività, in questo modo possiamo capire quali variabili siano effettivamente influenti per l'andamento dei consumi.

Per eseguire una prima indagine sulla possibile correlazione tra la serie e il meteo, sono state rappresentate su un piano cartesiano la componente stagionale e l'andamento di ciascuna variabile esogena negli anni presi in esame.

6.1.4.1 SARIMAX - Temperatura

La prima variabile presa in considerazione è *temperature*. Come si può osservare in Figura 17, l'andamento tra estate e inverno segue

abbastanza fedelmente l'andamento dei consumi energetici.

Per ciascuno dei due orizzonti temporali presi in considerazione, sono stati trovati parametri diversi che meglio rappresentassero la serie. Quelli individuati per l'orizzonte temporale di 120 settimane, producono un AIC di 459.123, invece quelli a 157 settimane di 720.669 (in *Appendice* i dettagli di questo modello). L'altra differenza tra i due modelli è la significatività della variabile esogena: la significatività aumenta all'aumentare delle settimane prese in considerazione.

Analizzando la distribuzione dei residui si può affermare che siano distribuiti in maniera ottimale per entrambi i modelli.

Inoltre, al fine di testare la capacità previsionale del modello è stata effettuata una previsione a partire dal modello a 120 settimane sulle ultime 37. Il modello a 157 settimane non è stato indagato ulteriormente avendo parametri diversi dall'altro e non avendo possibilità di testarlo.

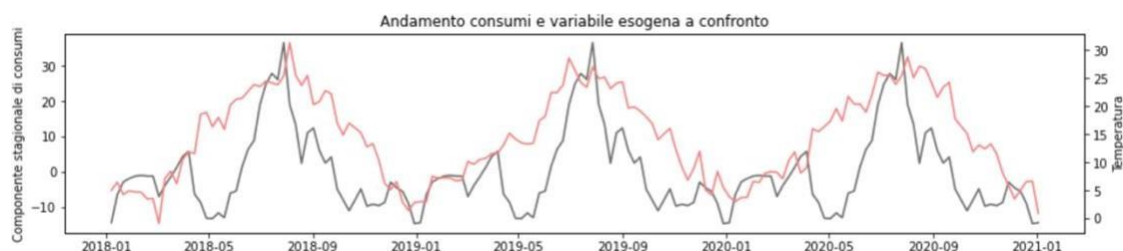


Figura 17. Serie storica a confronto con la temperatura media a settimana

Come si può osservare in *Figura 20*, la serie prevista in rosso segue solo in parte l'andamento di quella originale, in grigio. Si è quindi andati a verificare l'accuratezza della previsione con l'indice MAPE. Questo si posiziona a 7.46% per il

test set e per il training a 18.29%. La differenza è di oltre dieci punti percentuali e porta subito a pensare che siamo in presenza di un caso di *overfitting*, data anche la differenza di andamento osservabile nel dettaglio in *Figura 21*.

	coef	std err	z	P> z	[0.025	0.975]
temperature	0.6579	0.225	2.918	0.004	0.216	1.100
ar.L1	1.2836	0.100	12.811	0.000	1.087	1.480
ar.L2	-0.6724	0.082	-8.249	0.000	-0.832	-0.513
ma.L1	-1.8212	1.946	-0.936	0.349	-5.635	1.992
ma.L2	0.9987	2.129	0.469	0.639	-3.175	5.172
ar.S.L52	-0.5528	0.221	-2.498	0.013	-0.987	-0.119
sigma2	31.6169	67.940	0.465	0.642	-101.543	164.777

Figura 18. Parametri del modello SARIMAX (temperatura) a 120 settimane

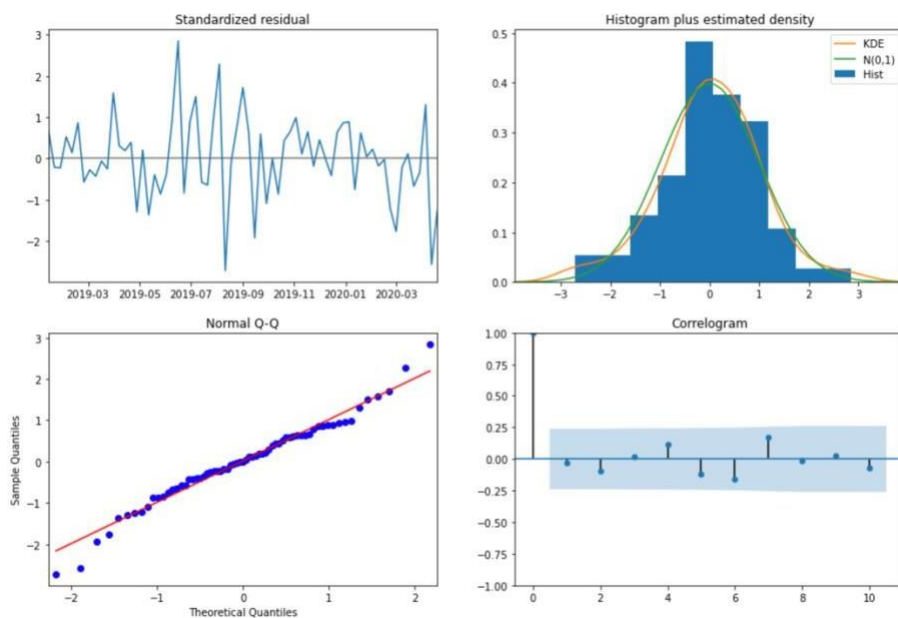


Figura 19. Plot errori del modello SARIMAX (temperatura) a 157 settimane

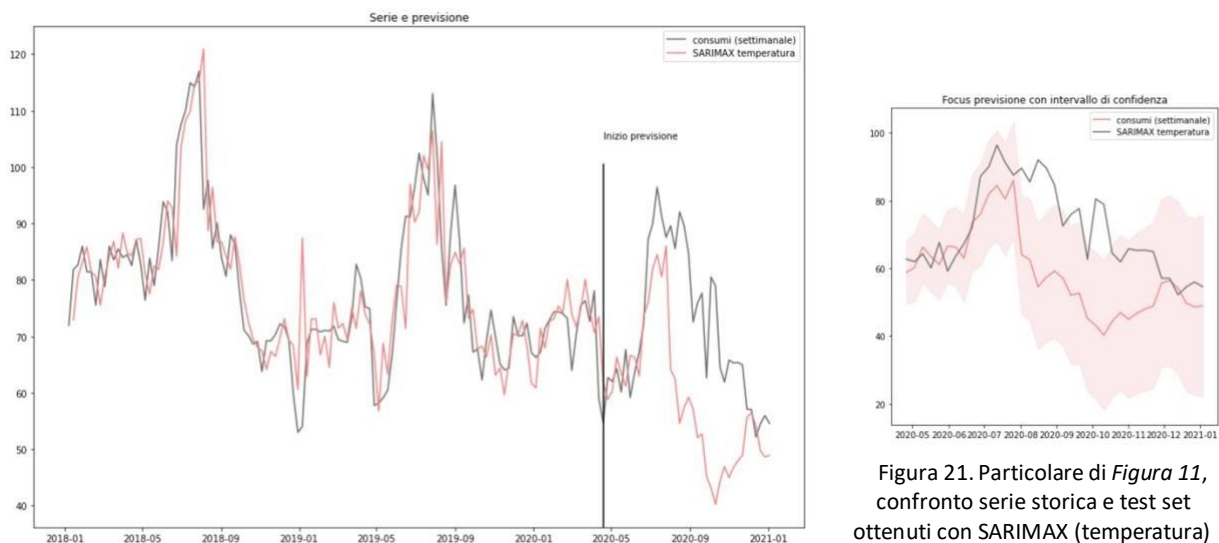


Figura 20. Serie storica a confronto con training e test set ottenuti con SARIMAX (temperatura)

Figura 21. Particolare di *Figura 11*, confronto serie storica e test set ottenuti con SARIMAX (temperatura)

6.1.4.2 Altre variabili esogene

Le altre variabili esogene del meteo prese in considerazione sono *ore_luce*, *humidity* e *is_raining*.

Sono stati costruiti i modelli SARIMAX per ciascuno di questi regressori; tuttavia, nessuno di questi è risultato significativo. I parametri derivati

dalla funzione di *parameter tuning* sono identici a quelli di SARIMA con $D = 1$, a sottolineare come una variabile esogene non significativa non aggiunga nessuna informazione rilevante al modello e, di conseguenza, venga ignorata da questo.

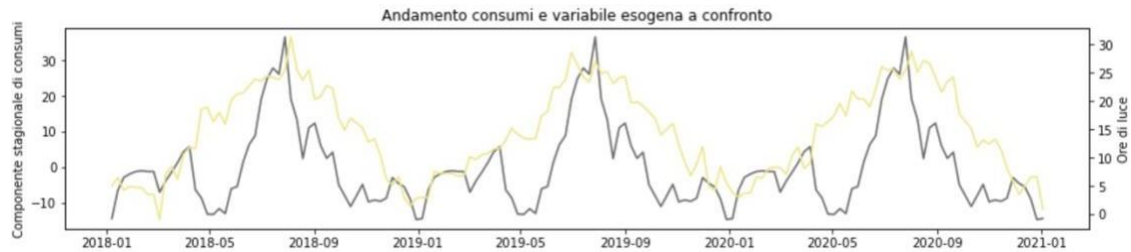


Figura 22. Serie storica a confronto con le ore di luce medie a settimana

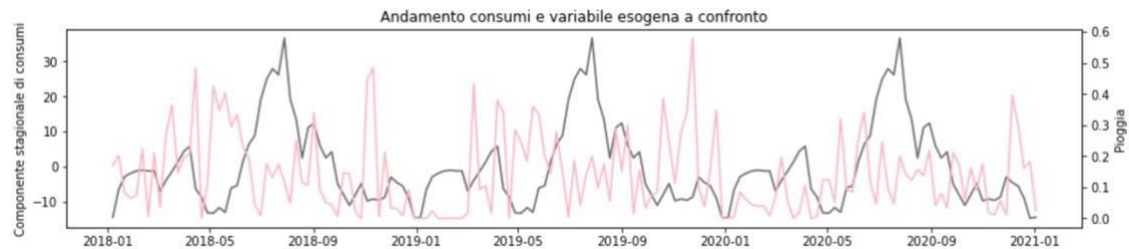


Figura 23 Serie storica a confronto con le precipitazioni medie a settimana

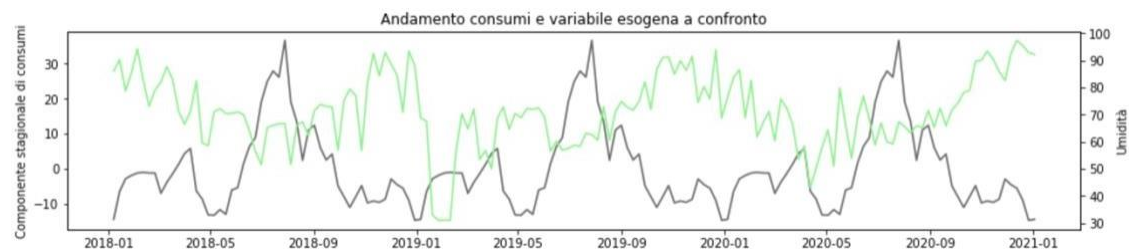


Figura 24. Serie storica a confronto con l'umidità media a settimana

6.1.5 Modello migliore

Tra tutti i modelli esposti precedentemente quello con il miglior comportamento è SARIMA con $D = 1$. Questo modello, oltre ad essere un buon compromesso in termini di AIC, è quello che adatta al meglio i risultati predittivi in fase di training e test. I modelli SARIMAX con temperatura e SARIMA privo di stagionalità annuale ($D = 0$), al contrario, presentano *overfitting*.

Usando quest'ultimo modello si è deciso di provare a prevedere l'andamento dei consumi nei mesi successivi a dicembre 2020, data di fine del

dataset di partenza. Questa previsione è stata fatta utilizzando come training set tutte le 157 settimane fornite in partenza per questo motivo si ha la possibilità di valutare l'accuratezza della previsione.

In Figura 25, 26 e 27 sono rappresentate le previsioni a rispettivamente 14, 27 e 40 settimane (3, 6 e 9 mesi). Osservando l'andamento ed essendo consapevoli di come dovrebbe essere l'andamento stagionale a parità di condizioni, è possibile affermare che, soprattutto nell'orizzonte temporale maggiore, l'andamento sembra rispecchiare una possibile realtà.

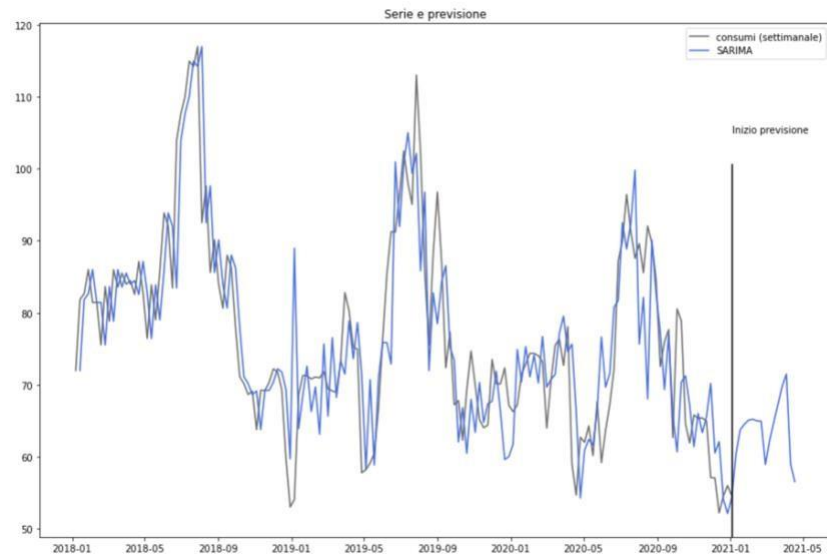


Figura 25. Previsione con SARIMA $D = 1$ di 3 mesi (14 settimane)

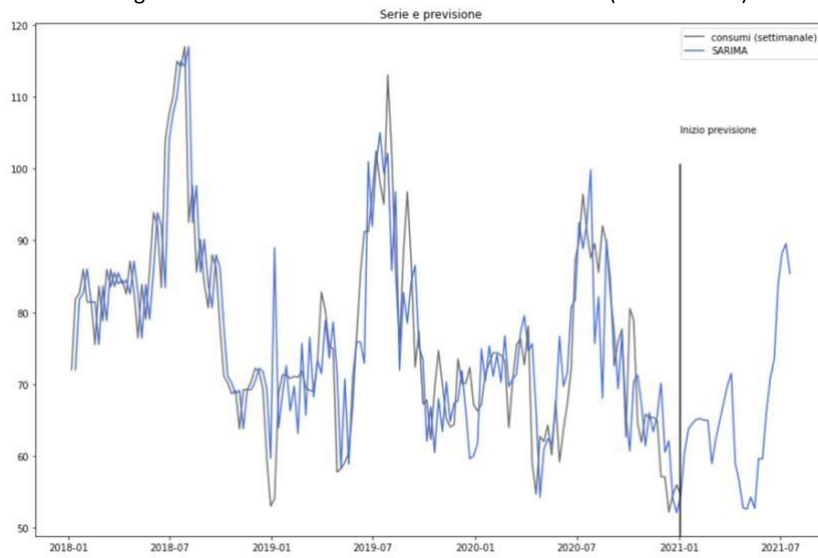


Figura 26. Previsione con SARIMA $D = 1$ di 6 mesi (27 settimane)

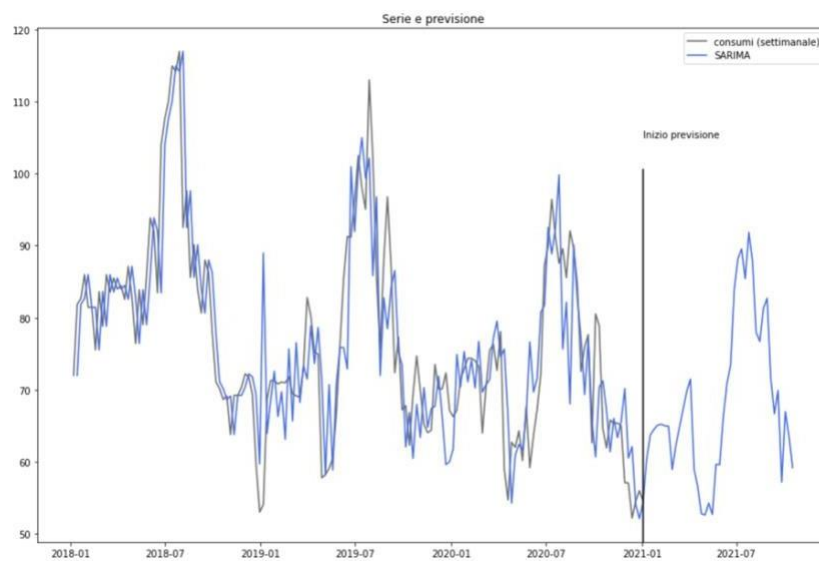


Figura 27. Previsione con SARIMA $D = 1$ di 9 mesi (40 settimane)

Una delle ragioni per cui il modello che considera anche il regressore temperatura non risulta essere il migliore è relativa al numero contenuto delle osservazioni oggetto del nostro studio: avendo reso il modello più complesso con l'aggiunta di un regressore più dati e più cicli a disposizione potrebbero dare risultati migliori rispetto al modello senza variabili esogene. Un'altra motivazione è data dalla natura della ciclicità della nostra serie: come osservato inizialmente in *Figura 17*, i consumi presentano una forte correlazione con le temperature nei periodi estivi ma nelle restanti stagioni la correlazione non è così evidente.

6.2 Prophet

Prophet è stato scelto per la capacità di gestire molteplici stagionalità. I suoi modelli, infatti, si adattano senza difficoltà alle stagionalità

settimanali e annuali, purché ci siano almeno due cicli di serie storica. Un aspetto interessante è la funzione di individuazione dei trend *change-points*, cioè quei punti in cui la funzione che descrive il trend cambia inclinazione o segno.

Il modello di regressione che utilizza la temperatura, risulta essere leggermente migliore rispetto al modello senza regressori e per questo lo si propone per l'analisi dei risultati.

6.2.1 Analisi Trend

Il trend risulta modellato più precisamente attraverso Prophet, in quanto è stato utilizzato per l'analisi il dataset nella sua forma giornaliera.

La prima analisi effettuata riguarda l'andamento del trend nei tre anni in esame e i trend *change-points* stimati.

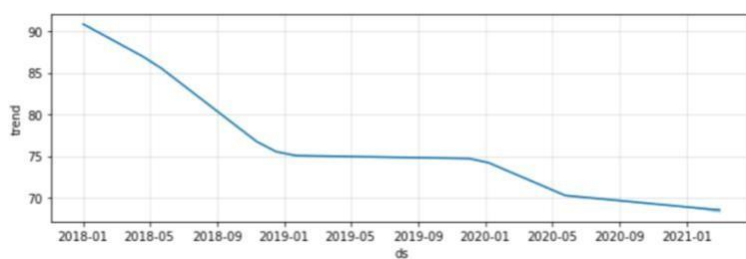


Figura 28. Trend con regressore

	ds	temperature	delta
105	2018-04-16	16.390875	0.050665
140	2018-05-21	17.744146	0.081700
315	2018-11-12	13.243528	0.137032
350	2018-12-17	1.893313	0.180196
385	2019-01-21	7.200000	0.110511
700	2019-12-02	6.613667	0.111985
735	2020-01-06	2.562667	0.125376
875	2020-05-25	21.858667	0.192508

Figura 29. Date dei trend *change-points*

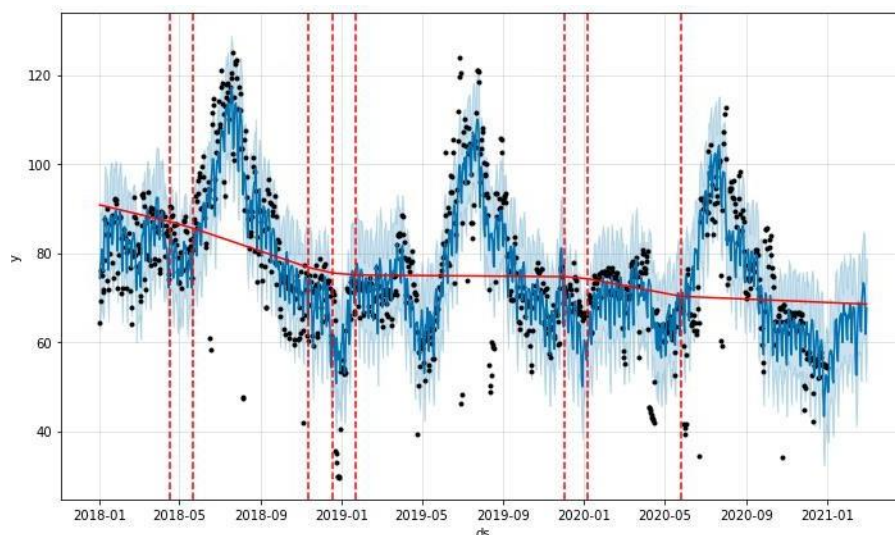


Figura 30. Trend *change-points*

Come si osserva in *Figura 28*, il trend risulta decrescente su tutto il periodo della serie storica.

Tra il 2018 e il 2019 si ha il calo maggiore, probabilmente per via dei lavori di ammodernamento degli impianti di illuminazione e raffrescamento in Università. Si ritiene, infatti, possibile che i *change-points* di dicembre 2018 segnalino l'assestamento, comunque decrescente, del consumo di fondo sul nuovo livello inferiore dopo le ristrutturazioni svolte durante l'anno. Si nota poi come i consumi tornino a calare in maniera più rilevante nel 2020, probabilmente anche a causa delle chiusure dovute al COVID-19.

Questo andamento sembra essere coerente con la politica *green* adottata dall'Ateneo negli ultimi anni. Nel dicembre 2018, infatti, l'Università ha aderito al RUS (Rete delle Università per lo Sviluppo sostenibile): un progetto che prevede la realizzazione di un piano per l'efficientamento energetico dell'Ateneo con l'obiettivo di ridurre al minimo l'impatto ambientale e, quindi, anche i consumi di energia elettrica. In seguito all'adesione al RUS, l'Università ha effettuato degli interventi strutturali atti a ridurre il proprio impatto energetico. In particolare, ha reso interamente a LED l'illuminazione degli spazi comuni e ha sostituito l'impianto di condizionamento con uno più efficiente. L'Ateneo, inoltre, ha dato vita al BASE¹ (Bicocca Ambiente Sostenibilità Economia) attraverso il quale ha affiancato agli interventi

strutturali una serie di attività di sensibilizzazione verso studenti e personale.

Analizzando i *change-points*, si è notato che alcuni di essi risultano coincidere temporalmente a eventi organizzati dal BASE e ad inizio e fine di sessioni d'esame (*Figura 29*: 16/04/2018, 21/01/2019, 25/02/2019, 25/05/2020) altri, invece, sono in prossimità di inizio e fine di periodi di vacanza (*Figura 29*: 17/12/2018, 06/01/2020). Un punto critico è che le date delle sessioni d'esame non coincidono per tutti i dipartimenti, quelli di Economia e Sociologia, per esempio, offrono una sessione aggiuntiva ad aprile.

Alcuni cambiamenti di trend potrebbero essere legati anche alla pandemia da Covid-19 che ha influenzato profondamente l'attività in presenza dell'Università.

Infine, la causa che più verosimilmente ha influenzato i trend *change-points* potrebbe essere l'implementazione di innovazioni strutturali o di modifiche nella gestione e nell'uso degli spazi universitari orientate ad un minore consumo energetico. Tuttavia, non avendo accesso a queste informazioni non si è in grado di determinare precisamente né le cause dei trend *change-points* né di effettuare analisi più approfondite.

6.2.2 Analisi stagionalità

Si analizzano ora le stagionalità. Nel modello di Prophet ne sono state riscontrate di due tipi: annuale e settimanale.

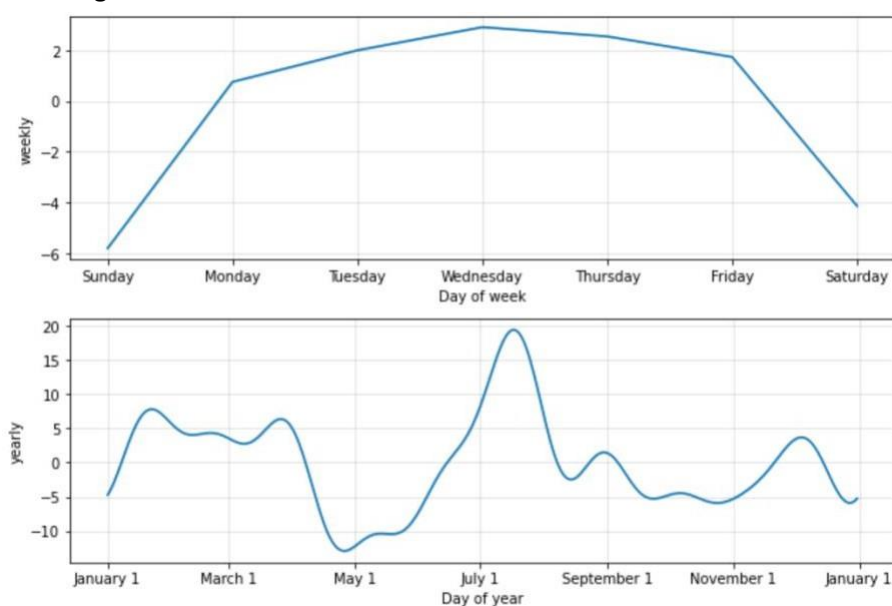


Figura 31. Stagionalità settimanale e annuale.

All'interno della stagionalità annuale si osserva come i consumi crescano sensibilmente nel periodo primaverile ed estivo, sintomo che il raffrescamento è verosimilmente la voce più influente sul consumo energetico. La presenza degli studenti sembrerebbe avere un'influenza solo parziale sui consumi, il picco infatti si ha nel mese di luglio, quando, per via della sessione d'esami, l'affluenza dovrebbe essere abbastanza contenuta. La presenza degli studenti non è quindi l'unico fattore rilevante, ciò è avvalorato dal fatto che ad agosto il consumo è comunque alto nonostante l'assenza di lezioni ed esami. Tuttavia, il fatto che da luglio ad agosto si registri un calo dei consumi fa comunque pensare che non sia un elemento completamente trascurabile.

Il consumo energetico registra un ulteriore calo nel mese di settembre, attestandosi su un livello quasi costante nel mese di ottobre (temperatura moderata e ancora abbastanza ore di luce). Si registra poi un aumento dei consumi nel mese di novembre (meno ore di luce), seguito da un calo a dicembre, che potrebbe essere causato dalle vacanze di Natale e dal fatto che molte lezioni sono già terminate e la sessione d'esami inizia solo a metà gennaio. A gennaio, infatti, il consumo risale, si ha una leggera oscillazione tra febbraio e marzo. Successivamente si verifica un calo considerevole tra aprile e maggio che porta a un consumo più basso nell'anno tra fine aprile e inizio maggio. Questo è dovuto al raffrescamento spento unito a più ore di luce naturale e alle vacanze di Pasqua, che non hanno data fissa, ma nei 3 anni a disposizione sono rispettivamente 1, 21 e 12 aprile. A maggio inizia ad aumentare il consumo che impenna tra giugno e luglio, probabilmente a causa dell'accensione del raffrescamento (15 giugno).

La stagionalità settimanale invece mostra consumi minimi nei fine settimana che, essendo giorni festivi, vedono, verosimilmente, un minor numero di studenti e personale presenti in università, la domenica l'università è chiusa. Nel complesso risulta un consumo in crescita nei primi tre giorni della settimana che porta al consumo massimo il mercoledì. Il lunedì risulta sempre il giorno feriale con il consumo minore.

7. Conclusione e possibili sviluppi

In questo report sono stati esplorati i trend e le stagionalità relative ai consumi dell'edificio U1 dell'Università degli Studi di Milano-Bicocca

nonché l'influenza che agenti atmosferici esterni potrebbero avere su questi. In particolare, sono state prese in considerazione la temperatura, le ore di luce giornaliere, l'umidità e le precipitazioni.

Diversamente da quanto ipotizzato sulla base del confronto visivo tra l'andamento delle variabili e quello dei consumi, l'unica variabile che si è dimostrata significativa ed influente per i modelli illustrati è la temperatura. Dalla rappresentazione grafica in *Figura 17* si può notare come temperatura e consumi risultino correlati soprattutto d'estate. È lecito quindi pensare che il raffrescamento in questo periodo dell'anno incida maggiormente sui consumi. D'inverno, invece, l'edificio è riscaldato tramite teleriscaldamento che, per caratteristiche tecniche, non inficia sul consumo di energia elettrica. L'analisi dei consumi, considerati assieme alla temperatura, in definitiva porterebbe a una pianificazione efficace del consumo di energia elettrica nel periodo estivo.

Relativamente alla presenza di studenti nell'edificio tramite indagine dei trend *change-points*, risulta complicato trarre conclusioni definitive. Infatti, l'U1 è frequentato da studenti di molti corsi di laurea con calendari accademici discordare e diversa organizzazione per dipartimento di appartenenza. Basarsi solamente sui calendari accademici porta quindi a risultati inconcludenti.

Prendendo spunto dal tentativo di analisi dei trend *change-points*, si ritiene che un possibile sviluppo futuro sia l'integrazione dei dati sul consumo con informazioni puntuali relative alla presenza di studenti, personale e visitatori nell'edificio. Raffrontando questi flussi con i consumi energetici si potrebbe intervenire, ad esempio, attraverso una migliore gestione degli spazi o degli eventi, magari evitando di illuminare o raffrescare zone inutilizzate.

Un ulteriore spunto riguarda i rinnovamenti e gli efficientamenti strutturali dell'edificio. Sarebbe opportuno poter distinguere accuratamente quando i cambiamenti di trend sono relativi ad un miglioramento impiantistico e quando, invece, sono provocati da altri fattori come, ad esempio, una variazione d'uso degli spazi.

Riferimenti bibliografici

[1] Università degli studi di Milano-Bicocca, BASE (2020). Report di sostenibilità. <https://www.unimib.it/sites/default/files/Base/Report di Sostenibilita 2020 web.pdf> [Accessed 08/02/2022]

[2] Brownlee, J. (2017). A Gentle Introduction to Autocorrelation and Partial Autocorrelation. <https://machinelearningmastery.com/gentle-introduction-autocorrelation-partial-autocorrelation/> [Accessed 08/02/2022]

[3] Prabhakaran, S. (2019). Augmented Dickey Fuller Test (ADF Test) – Must Read Guide. <https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/> [Accessed 08/02/2022]

[4] Brownlee, J. (2018). A Gentle Introduction to SARIMA for Time Series Forecasting in Python. <https://machinelearningmastery.com/sarima-for-time-series-forecasting-in-python/> [Accessed 08/02/2022]

[5] Brownlee, J. (2017). How to Remove Trends and Seasonality with a Difference Transform in Python. <https://machinelearningmastery.com/remove-trends-seasonality-difference-transform-python/> [Accessed 08/02/2022]

[6] Brownlee, J. (2017). How to Visualize Time Series Residual Forecast Errors with Python. <https://machinelearningmastery.com/visualize-time-series-residual-forecast-errors-with-python/> [Accessed 08/02/2022]

[7] Prophet (2022). Quick Start. https://facebook.github.io/prophet/docs/quick_start.html#python-api [Accessed 08/02/2022]

[8] Prophet (2022). Trend Changepoints. https://facebook.github.io/prophet/docs/trend_changepoints.html [Accessed 08/02/2022]

[9] OpenWeatherMap.org (2012). Current weather and forecast – OpenWeatherMap <https://openweathermap.org/> [Accessed 08/02/2022]

[10] Prophet. Outliers. <https://facebook.github.io/prophet/docs/outliers.html> [Accessed 08/02/2022]

[11] 6.3 Classical decomposition | Forecasting: Principles and Practice (2nd ed). otexts.com. <https://otexts.com/fpp2/classical-decomposition.html> [Accessed 08/02/2022]

[12] PennState. 4.1 Seasonal ARIMA models | Applied Time Series Analysis. <https://online.stat.psu.edu/stat510/lesson/4/4.1> [Accessed 08/02/2022]

[13] neptune.ai (2021). Anomaly Detection in Time Series. <https://neptune.ai/blog/anomaly-detection-in-time-series> [Accessed 08/02/2022]

[14] Brockwell, P.J. and Davis, R.A. (2010). Introduction to Time Series and Forecasting. New York: Springer.

Appendice

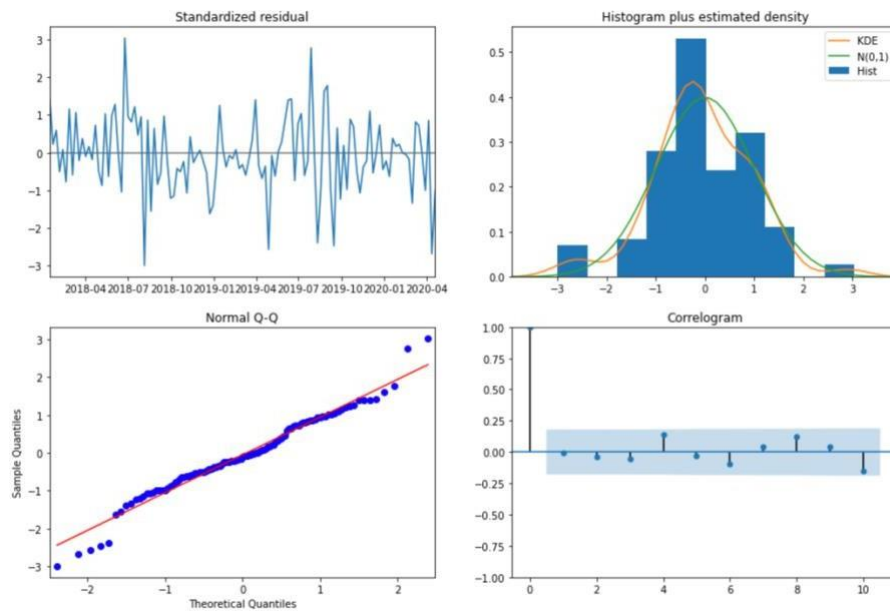


Figura A. *Plot* errori del modello SARIMA $D = 0$ a 120 settimane

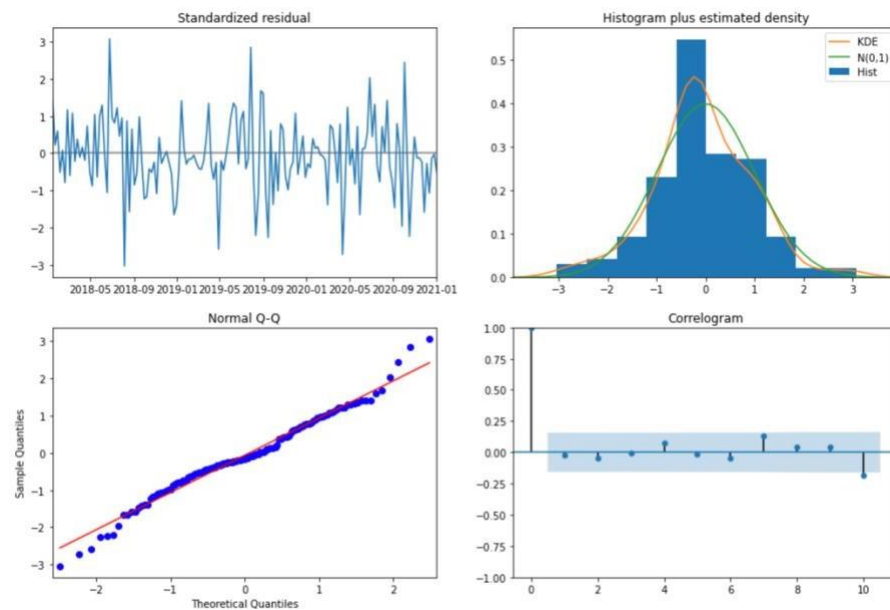


Figura B. *Plot* errori del modello SARIMA $D = 0$ a 157 settimane

	coef	std err	z	P> z	[0.025	0.975]
temperature	0.8640	0.230	3.755	0.000	0.413	1.315
ma.L1	-0.3970	0.096	-4.135	0.000	-0.585	-0.209
ma.L2	-0.3775	0.124	-3.040	0.002	-0.621	-0.134
ar.S.L52	-0.4092	0.119	-3.448	0.001	-0.642	-0.177
sigma2	49.2015	5.546	8.872	0.000	38.332	60.071

Figura C. Parametri del modello SARIMAX (temperatura) a 157 settimane

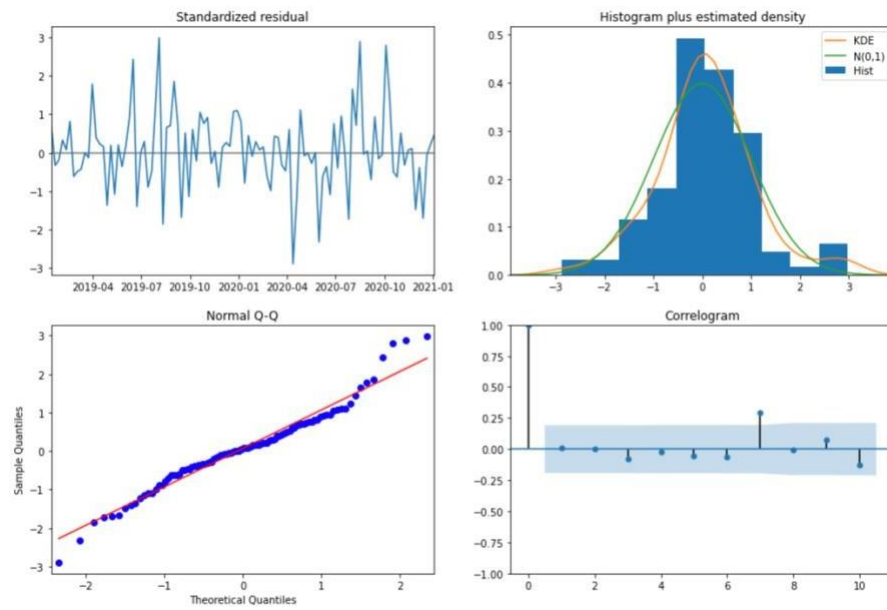


Figura D. *Plot* errori del modello SARIMAX (temperatura) a 157 settimane