

Analisi dei consumi energetici dell'Università Milano-Bicocca

Autore 1: Miryam Bartoli, 867107, m.bartoli3@campus.unimib.it

Autore 2: Cesare Colzani, 821253, c.colzani3@campus.unimib.it

Autore 3: Mario Pedol, 830296, m.pedol1@campus.unimib.it

Autore 4: Nabil El Asri, 826040, n.elasri@campus.unimib.it

Indice

1	Sinossi	1
2	Parole chiave	1
3	Introduzione	1
4	Obiettivi	2
5	Aspetti metodologici	2
5.1	Modello ARIMA(p,d,q)	2
5.2	FBprophet	3
6	I dati	3
6.1	Data Quality	4
6.2	Osservazioni	4
7	Analisi e trattamento dei dati	5
7.1	Data Cleaning	5
7.2	Stima dell'andamento Teorico dei consumi energetici durante il periodo pandemico	6
7.2.1	Risultati	8
7.3	Dashboard di monitoraggio costi e consumi	9
7.3.1	Risultati	10
8	Conclusione e possibili sviluppi	10
9	Bibliografia	11

1 Sinossi

Qual è stato l'impatto del Covid-19 sui consumi energetici dell'Università degli studi di Milano-Bicocca? Attraverso un'accurata analisi relativa ai dati degli edifici U1 e U6 questo elaborato si pone l'obiettivo di fornire una panoramica di come le due ondate di pandemia abbiano avuto impatto sui consumi, al fine di suggerire, per quanto possibile, strategie che permettano di ottimizzare la gestione della potenza assorbita. Verrà inoltre costruito uno strumento interattivo che possa permettere il monitoraggio delle variazioni dei consumi, individuando delle apposite stime che permettano di identificare il corrispettivo

costo economico e l'impronta di CO_2 rilasciata dagli edifici.

2 Parole chiave

I temi principali oggetto di analisi possono essere individuati mediante tre parole chiave: Sostenibilità, Consumo Energetico, Efficienza Energetica.

In particolare, questi termini possono essere definiti come segue:

- **Sostenibilità:** si riferisce a quel tipo di produzione e utilizzo delle risorse energetiche che implica uno sviluppo sostenibile volto a preservare l'ambiente;
- **Consumo di energia elettrica:** domanda di energia prodotta dalla fornitura elettrica utilizzata;
- **Efficienza energetica:** insieme di operazioni che permettono di contenere i consumi energetici, ottimizzando il rapporto tra fabbisogno energetico e livello di emissioni di CO_2 .

Queste sono parole che ricorreranno spesso nei paragrafi che seguono.

3 Introduzione

"Nulla si crea e nulla si distrugge ma tutto si trasforma".

Per quanto intuitiva, tale affermazione, di cui nessuno metterebbe in dubbio la veridicità, non sempre rispecchia le azioni quotidiane e la gestione delle risorse. Spesso, la consapevolezza della necessità di seguire principi di sostenibilità non si traduce in comportamenti volti in tale direzione perché gli effetti non sono tangibili o immediatamente riscontrabili. Per questo, infatti, ricorriamo all'aiuto dei dati. La chiamata alla regolazione del consumo dell'energia è ormai impossibile da ignorare e le università, non solo in qualità di edifici ma soprattutto in quanto poli di formazione, non possono sottrarsi a tale compito. Il *National Bureau of Statistics of China* nel

2009 asseriva che "secondo alcune statistiche, il consumo energetico pro capite nelle strutture universitarie e il consumo energetico per unità di superficie sono superiori ai consumi di energia per unità di superficie degli edifici residenziali" [1].

Un gruppo di studiosi americani [2] ha individuato nove punti cardine verso cui le Università dovrebbero rivolgere la loro attenzione per guidare questa rivoluzione sostenibile che spaziano da azioni concrete in merito alla gestione delle proprie risorse e proprietà ad obiettivi a lungo termine che mirano all'educazione e la ricerca in tale campo.

Un ulteriore importante parametro di valutazione è il calcolo della Carbon Footprint, ovvero la quantità di gas serra emessi nell'atmosfera. In particolare l'energia elettrica, seppur indirettamente, concorre alla crescita di questo fenomeno per cui è importante regolarne il consumo.

L'Università Milano-Bicocca nell'anno 2020 si è posizionata 132^{ma} nella classifica mondiale Green Metric che valuta i poli universitari in base al loro impegno in tema di sostenibilità[3]. Infatti, sin dalla sua fondazione, l'Ateneo ha sempre avuto un occhio di riguardo sul tema dei consumi energetici e della sostenibilità concretizzato nella nascita dell'organo denominato BASE (Bicocca Ambiente Società Economia), composto da un comitato scientifico comprensivo di tutte le figure presenti nell'Università stessa. Gli obiettivi di BASE sono in linea con quelli proposti dallo studio sopracitato e sono già avviati e in continuo sviluppo[4].

L'intento di questo elaborato è quindi quello di partire da un'analisi dei consumi effettivi rilevati nell'Università dall'anno 2017 ad oggi al fine di avere un'idea più precisa della distribuzione degli stessi in base alle destinazioni dei vari edifici e l'impatto che possono avere sull'ambiente. Anche in termini di spesa è utile comprendere in quali ambiti è preventivabile un risparmio e dove si posizionano i costi più ingenti sia temporalmente che logisticamente. Il fine ultimo è quello di incentivare una condotta guidata dalla consapevolezza e dal desiderio di preservare l'ambiente non solo a livello di gestione ma anche e soprattutto da parte degli studenti e di coloro che vivono l'Università quotidianamente.

4 Obiettivi

Questo elaborato propone due obiettivi chiave che hanno lo scopo di aiutare l'Ateneo a prendere e suggerire efficaci ed efficienti decisioni. In particolar modo, l'Area Infrastrutture e Approvvigionamenti, che si occupa della promozione di azioni strutturali-

gestionali ed educativo-comportamentali finalizzate al risparmio e all'efficienza energetica.[5] Nello specifico questi sono:

1. Lo studio dell'impatto dello smart working e della DAD, dovuti al Covid-19 sui consumi mediante una comparazione dell'andamento teorico stimato con i dovuti modelli previsionali, contro quello effettivo. Più nel dettaglio: per l'edificio U6 è stato stimato l'effetto della seconda ondata pandemica sui consumi mentre per quanto riguarda l'edificio U1 sono state determinate le stime dei consumi durante il periodo di pandemia;
2. Costruzione di una dashboard che permette il monitoraggio dei consumi energetici, dei prezzi e della Carbon Footprint per individuare le anomalie al fine di suggerire eventuali miglioramenti che possono essere effettuati nel tempo.

Nei paragrafi che seguono verranno analizzati nello specifico entrambi gli obiettivi appena descritti, facendo riferimento al periodo che va dal 2018 al 2020, per gli edifici U1 e U6.

5 Aspetti metodologici

Prima di procedere con le analisi, è opportuno esplicitare in maniera formale i metodi adottati in questo elaborato.

Per rispondere al primo obiettivo, data la natura temporale dei dati, sono stati utilizzati dei modelli legati alle serie storiche, in particolare si è fatto affidamento alla classe dei modelli ARIMA e al modello "prophet", esplicitato in seguito.

Per il secondo obiettivo invece, si è reso necessario l'utilizzo di un software ad hoc per la creazione di una dashboard interattiva e informativa.

5.1 Modello ARIMA(p,d,q)

I modelli ARIMA, *AutoRegressive Integrated Moving Average*, sono una classe di modelli statistici utilizzati per analizzare e ottenere delle previsioni partendo da una serie temporale.

L'acronimo ARIMA racchiude tutti gli aspetti chiave del modello:

- Parte autoregressiva (**AR**): utilizza la relazione di dipendenza tra un'osservazione e un certo numero di osservazione ritardate;
- Parte integrata (**I**): consente di rendere stazionarie serie temporali attraverso la differenziazione della serie;

- Parte a media mobile (**MA**): utilizza la dipendenza che sussiste tra un'osservazione e un errore residuo attraverso un modello a media mobile applicato alle osservazioni ritardate.

Formalmente possiamo scrivere il modello come segue:

$$\varphi_b(B)\Delta^d Z_t = \vartheta_0 + \vartheta q(B)a_t : ARIMA_{(p,d,q)} \quad (1)$$

Dove:

- $\varphi_b(B)\Delta^d Z_t$ è la componente AR di ordine p ;
- $\Delta^d Z_t$ è la parte integrata del modello con d numero di differenziazioni;
- ϑ_0 è la costante;
- $\vartheta q(B)a_t$ è la componente a media mobile di ordine q .

5.2 FBprophet

FBprophet è uno strumento di previsione intuitivo e personalizzabile che non richiede competenze tecniche specifiche sul modello utilizzato. Similmente ai modelli additivi generalizzati (GAM), la previsione si ottiene sommando gli eventuali trend e le componenti stagionali e regredendo, poi, sul tempo. La formulazione generale è la seguente:

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \quad (2)$$

Dove:

- $g(t)$: rappresenta i cambiamenti non periodici (trend);
- $s(t)$: rappresenta le componenti periodiche (stagionalità settimanali, mensili e/o annuali);
- $h(t)$: rappresenta le feste ricorrenti ma che non ricadono sempre nello stesso giorno o periodo (es. Pasqua);
- ε_t : rappresenta l'errore che il modello non riesce a spiegare.

I trend $g(t)$ possono essere non lineari e costanti (come l'aumento della popolazione) o lineari ma non costanti. Per definire questa funzione si identificano dei punti specifici (detti punti di cambiamento) dove, appunto, il tasso di crescita κ subisce una modifica a cui vengono aggiunti degli aggiustamenti (analogamente si tratta il parametro compensativo m). Considerando la portata $C'(t)$ anch'essa mutabile nel tempo si giunge alle seguenti formulazioni:

- trend non lineari ma costanti:

$$g(t) = \frac{C'(t)}{1 + e^{-(\kappa + a(t)^T \delta)(t - (m + a(t)^T \gamma))}}$$

- trend lineari ma non costanti:

$$g(t) = (\kappa + a(t)^T \delta)t + (m + a(t)^T \gamma)$$

L'incertezza nella predizione dei trend futuri è misurata dall'assunzione che il tasso di cambiamenti in un intervallo T del futuro avrà la stessa frequenza e numerosità medie di un analogo periodo T nel passato.

Per definire le componenti stagionali si ricorre a funzioni periodiche. Considerando una serie di Fourier standard¹ si stimano i $2N$ parametri:

$$\beta = [a_1, b_1, \dots, a_N, b_N]^T$$

attraverso la costruzione di una matrice $X(t)$ di vettori stagionali per ogni istante t per i dati passati e futuri. Assumendo β distribuito normalmente si ottiene:

$$s(t) = X(t)\beta \quad (3)$$

Le feste ricorrenti sono trattate come mutuamente indipendenti. Viene richiesto all'analista di fornire una lista delle date che ritiene festività non già considerate e mediante una funzione indicatrice $Z(t)$ si associa ogni istante t al cambiamento nella previsione relativo all'evento in questione. Supponendo κ distribuito normalmente risulta:

$$h(t) = Z(t)\kappa \quad (4)$$

6 I dati

Al fine di raggiungere gli obiettivi esposti nel paragrafo 4 si sono utilizzate tre fonti principali:

Università degli studi di Milano-Bicocca : ha messo a disposizione le informazioni sui consumi energetici da gennaio 2018 al dicembre 2020;

Il materiale è costituito da una serie di file suddivisi per edificio (U1, U6) e anno (2018, 2019, 2020). Ognuno di questi identifica uno specifico mese e presenta informazioni riguardanti le rilevazioni effettuate giornalmente con un livello di dettaglio pari ad una misurazione ogni 15 minuti. Talvolta, emergono più file per un singolo mese con formati differenti (es. Marzo 2020) oppure un singolo mese descritto da più file separati (es. Agosto 2018).

Queste anomalie sono state trattate con particolare attenzione nel processo di unificazione dei dati. Per

¹ $s(t) = \sum_{n=1}^N (a_n \cos(\frac{2\pi n t}{P}) + b_n \sin(\frac{2\pi n t}{P}))$

il raggiungimento degli obiettivi si è deciso di procedere costruendo due dataset separati di 105406 e 108284 osservazioni e 8 variabili. Le variabili di interesse sono:

- **POD**: identificativo del contatore di energia;
- **DATA**: giorno in cui viene effettuata rilevazione;
- **ORA**: orario in cui viene effettuata la rilevazione (le rilevazioni sono effettuate ogni 15 minuti a partire dalla mezzanotte del primo giorno del mese);
- **FL_ORA_LEGALE**: denota se le rilevazioni sono riferite a un periodo di ora legale oppure a un periodo di ora solare;
- **CONSUMO_ATTIVA_PRELEVATA**: consumo energetico (espresso in kw) nei 15 minuti considerati;
- **CONSUMO_REATTIVA_INDUTTIVA_PRELEVATA**: consumo energetico (espresso in kvarh) prelevato ma non effettivamente impiegato per produrre lavoro utile;
- **POTENZA_MASSIMA**: limite massimo di potenza del contatore (espresso in kw) prelevabile dalla rete elettrica;
- **TIPO_DATO**: identifica la tipologia dei dati.

È stato ritenuto opportuno escludere le variabili che descrivono l'orario preciso della rilevazione e la potenza attiva (kw) in quanto presenti esclusivamente in due file.

segugio.it : ha permesso di ottenere informazioni sul costo della componente materia prima energia espressa in €/kWh trimestralmente da gennaio 2018 a ottobre 2020;[6]

Tali informazioni hanno permesso di costruire il dataset *pricing* costituito da 13 osservazioni e una sola variabile rappresentata dal prezzo della materia prima energia.

Politech technology srl : società utilizzata come benchmark per ricavare il fattore moltiplicativo utile a stimare il Carbon Footprint necessario per la costruzione della dashboard. [7]

6.1 Data Quality

In seguito si è passati alla valutazione dei dataset utilizzati in termini di qualità.

- **Attendibilità**: tutte e tre le fonti godono di un'ottima attendibilità. I dati energetici sono rilevazioni dirette per mezzo di strumenti il cui eventuale errore di misura è irrilevante dato l'ordine di grandezza complessivo dei valori e dell'utilizzo che ne è stato fatto. **segugio.it** è un'affermata piattaforma di comparazione di tariffe e prodotti che collabora con numerose società. Infine, **Politech technology srl** è un'azienda italiana che da oltre 25 anni lavora nel campo dello stampaggio di materie plastiche ad iniezione la cui attenzione al tema ecologico è, sin dall'inizio, fondamento della loro missione;
- **Consistenza**: le variabili di interesse presentavano lo stesso formato in tutti i dataset disponibili quindi il livello di consistenza era più che buono. L'unico accorgimento da riportare è legato al mese di Giugno 2020 che risultava identico tra i due edifici indicando l'erroneità di uno dei due. Da approfondite analisi si è dedotto che i dati erano relativi all'edificio U1 portando a concludere l'assenza di quelli relativi ad U6;
- **Completezza**: all'interno dei dati energetici è stato trovato qualche valore mancante. Tra quelle di interesse, le variabili che hanno presentato questa anomalia sono state 'ORA' e 'CONSUMO_ATTIVA_PRELEVATA'. In entrambi i casi il problema è stato ovviato in fase di pulizia e successiva aggregazione. Inoltre, come espresso al punto precedente, mancavano totalmente i dati relativi al mese di Giugno 2020 per l'edificio U6.

6.2 Osservazioni

È inevitabile che proseguendo le analisi si presenteranno ulteriori punti di forza e debolezza, ad ora sono stati identificati i seguenti:

- **Punti di debolezza**: i dati grezzi hanno presentato i seguenti problemi:
 - * missing values
 - * valori estremi
 - * valori/periodi duplicati
 - * mancata gestione dell'ora legale e solare

Tali problematiche verranno esplorate e affrontate in maniera più estesa successivamente.

- **Punti di forza**: accuratezza delle rilevazioni in quanto derivano da un POD, inoltre l'elevata granularità dei dati permette di condurre svariate analisi in diversi campi.

7 Analisi e trattamento dei dati

Da un'analisi preventiva dei dati, sono emerse delle problematiche che compromettono una chiara lettura del fenomeno oggetto di studio:

1. Come mostra la Figura 1, le due serie storiche per gli edifici U1 e U6 sono perfettamente identiche nel mese di giugno 2020. In realtà, si è scoperto poi, grazie all'identificativo del POD che contraddistingue un edificio dall'altro, che i dati relativi al mese di giugno dell'edificio U1 sono stati copiati, presumibilmente per errore, nei dati riguardanti l'edificio U6.

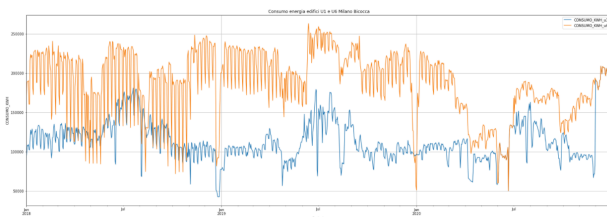


Figura 1: Consumo energia edifici U1 e U6 Milano Bicocca

2. Ci sono delle date, sette per edificio, che presentano consumi di energia conteggiati erroneamente dal POD: in particolare si è notato un azzeramento improvviso del consumo per alcuni intervalli di tempo e, sempre vicini a questi intervalli, dei valori che si discostano significativamente dal resto delle osservazioni. Inoltre, la data corrispondente al 25/10/2020 è stata registrata due volte per entrambi gli edifici
3. I giorni interessati dal cambio d'ora presentano, per ovvie ragioni, quattro rilevazioni in più o in meno e richiedono degli aggiustamenti per uniformare la serie storica.

Tali criticità verranno opportunamente trattate nel paragrafo che segue.

7.1 Data Cleaning

Da un'analisi esplorativa i problemi riscontrati sui dataset ottenuti sono stati i seguenti:

- **Missing values:** consumi energetici pari a 0 Kw. In particolare, il dataset U1 presenta 122 missing values e il dataset U6 presenta 216 missing values.

- **Valori estremi:** consumi energetici che si discostano in modo significativo dalla media. Tali valori emergono esclusivamente in quelle date in cui sono presenti consumi energetici pari a 0.

Queste anomalie sono attribuibili a problemi legati ai sensori del POD che talvolta, a causa di una saturazione del sistema, sono incapaci di gestire autonomamente i valori, pertanto si riavviano. Al fine di ottenere una serie storica ragionevole dei consumi si è deciso di imputare i missing values e di attribuire un nuovo valore agli outlier mediante la seguente regola decisionale:

CASO 1: valori estremi e/o missing values (k) concentrati in un piccolo intervallo temporale

- 1a** se k è compreso tra due valori non anomali: sostituirlo con la media aritmetica tra a e b

$$a \mathbf{k} b \quad a, b \in N \implies k = \frac{a+b}{2}$$

- 1b** se k è compreso tra due valori anomali: sostituire a , b con i primi valori non anomali adiacenti e k con la media aritmetica tra i nuovi valori di a e b

$$a \mathbf{k} B b \quad A, B \text{ valori anomali} \\ a, b \in N \implies \\ A = a \quad B = b \quad k = \frac{a+b}{2}$$

CASO 2: valori anomali (k) concentrati in un ampio orizzonte temporale

Giorno $X - 7$ dalle 9:00 alle 11:50:

4 **12** 4 56 67 32 44 9 20 ... 50

Giorno X dalle 9:00 alle 11:50:

4 **0** 0 0 0 0 0 0 ... 50

Giorno $X + 1$ dalle 9:00 alle 11:50:

4 **12** 4 56 67 32 44 9 20 ... 50

Stimare il valore mancante nel corrispettivo istante temporale, attraverso una media pesata tra i valori dello stesso istante del giorno successivo (s) e del medesimo giorno della settimana precedente (t), con pesi $s = 0.8$ e $t = 0.2$.

- Consumi duplicati: il file excel presenta alcuni giorni che sono stati inseriti più volte (es. 25/10/20). Per tali giorni sono stati eliminati i consumi duplicati.

- Ora legale: le date del 27/10/18 e del 28/10/18 sono caratterizzate da un'ora in più mentre il 25/03/18, 31/03/19 e 29/03/20 presentano un'ora in meno.

Nelle date di ora legale, nell'intervallo che va dalle 2 alle 3 del mattino, sono presenti 2 volte gli stessi consumi, pertanto si è deciso di considerarli solo 1 volta. Nelle date di ora solare, non sono presenti i consumi nell'intervallo che va dalle 2 alle 3 del mattino, pertanto si sono stimati questi 4 valori attraverso una media dei valori presenti nella mezz'ora precedente e successiva.

$$a_1 \ a_2 \ a_3 \ a_4 \ k_1 \ k_2 \ k_3 \ k_4 \ b_1 \ b_2 \ b_3 \ b_4 \rightarrow a_1 \ a_2 \ a_3 \ a_4 \ \frac{a_3+a_4}{2} \ \frac{a_3+a_4}{2} \ \frac{b_1+b_2}{2} \ \frac{b_1+b_2}{2} \ b_1 \ b_2 \ b_3 \ b_4$$

k_1 e k_2 sono stati sostituiti dalla media aritmetica dei consumi dei due istanti precedenti. k_3 e k_4 sono stati sostituiti dalla media aritmetica dei consumi dei due istanti successivi.

Una volta terminata la fase di pulizia dei dati, si sono prese in considerazione solo 4 dimensioni che rappresentano i consumi energetici degli edifici oggetto di analisi da gennaio 2018 a dicembre 2020:

- ORA_u1
- DATA_u1
- CONSUMO_KWH_u1
- DATA_u6
- ORA_u6
- CONSUMO_KWH_u6

Le due variabili CONSUMO_KWH_u1 e CONSUMO_KWH_u6 indicano il consumo orario e sono state ottenute moltiplicando ogni valore dell'attributo CONSUMO_ATTIVITA_PRELEVATA per 15 (ovvero i minuti totali dell'intervallo di riferimento), poi aggregando per ora.

7.2 Stima dell'andamento Teorico dei consumi energetici durante il periodo pandemico

Al fine di stimare l'andamento teorico durante il periodo COVID (dal 01/03/2020), ossia come sarebbe stato senza pandemia, si è fatto ricorso a due differenti modelli ritenuti capaci di interpretare meglio i consumi per gli edifici considerati.

U1

Per quanto riguarda l'edificio U1 si è scelto di considerare la serie dei consumi energetici aggregata giornalmente. Osservando il grafico

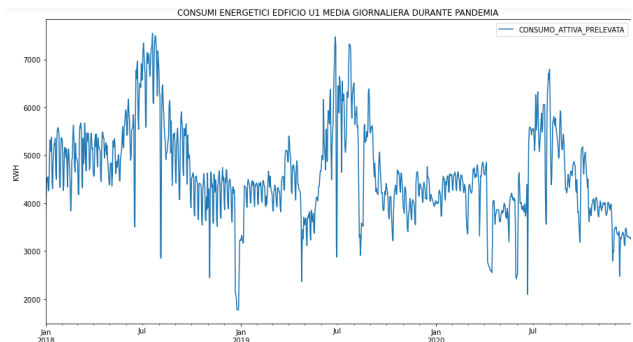


Figura 2: Consumo energetico giornaliero (edificio U1 Milano-Bicocca)

mostrato nella Figura 2, è possibile notare una lieve decrescita dei consumi durante il periodo della pandemia.

Scomponendo la serie nelle sue tre componenti, si può trovare riscontro della precedente ipotesi. Da questa scomposizione, Figura 3, si può

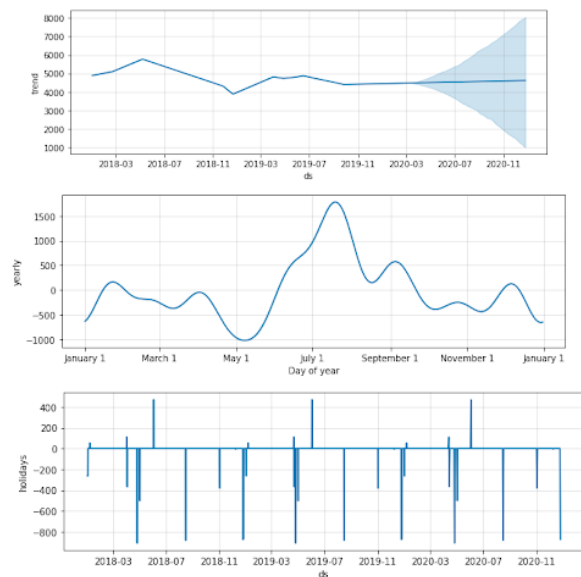


Figura 3: Scomposizione serie dei consumi energetici (edificio U1 Milano-Bicocca)

notare che durante il periodo pandemico:

1. Componente di trend: stabile rispetto ai mesi precedenti;
2. Stagionalità annuale: inizia un andamento decrescente nel mese di febbraio che si arresta a giugno presentando una considerevole flessione nel mese di maggio;

3. Componente vacanze: non è informativo poiché l'Università è chiusa nei giorni di vacanza.

Alla luce di quanto emerso fino ad ora, si è deciso di basare la stima dei valori futuri, sul modello "prophet", presente nel pacchetto python "fbprophet", presentato nel paragrafo 5.2. Tale scelta è ricaduta su questo modello poiché esso è in grado di rilevare automaticamente le variazioni tendenziali. Inoltre è molto flessibile, facilmente interpretabile e accurato.

Il dataset è stato suddiviso in due parti Training e Test set. Il Training set comprende tutti i consumi dal 01/01/18 al 29/02/20, mentre i restanti sono inclusi nel test set. Successivamente si è applicato il modello sul test set, come mostrato in Figura 4, dove i consumi sono stati stimati attraverso delle bande di previsione.

Alternativamente, per stimare tali dati senza

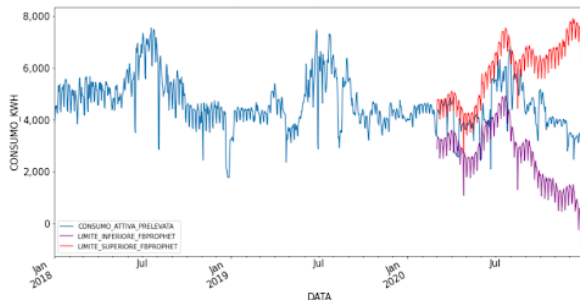


Figura 4: Stima dei consumi giornalieri da marzo 2020 per l'edificio U1

sfruttare un metodo blackbox, è possibile avvalersi della previsione rolling con i modelli ARIMA. In particolare, possiamo suddividere il dataset di addestramento in train e test set, utilizzando il train set per adattare il modello e generare una previsione per ogni elemento del test set.

Si nota che è necessaria una previsione mobile data la dipendenza delle osservazioni dalle fasi temporali precedenti. Un modo rozzo per eseguire questa previsione è ricreare il modello ARIMA dopo che è stata ricevuta ogni nuova osservazione (Figura 5), tenendo traccia manualmente di tutte le osservazioni in una lista chiamata "cronologia" che contiene i dati di addestramento al quale vengono aggiunte nuove osservazioni a ogni iterazione.

U6

Come emerso nel paragrafo 7, è stato necessario stimare, tramite un opportuno modello,

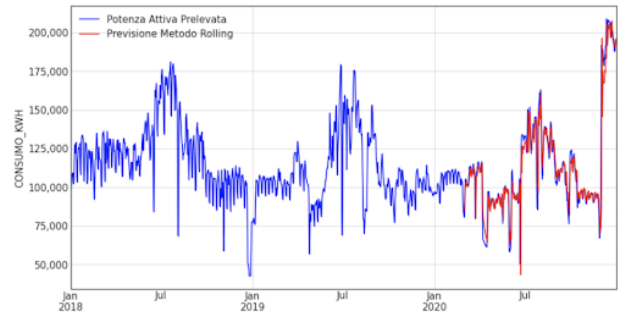


Figura 5: Previsione consumi U1 con metodo di previsione rolling

i consumi di giugno 2020 in quanto coincidenti con i valori dell'edificio U1 relativi allo stesso mese. Per procedere a tale stima si è fatto affidamento sul modello utilizzato precedentemente per l'analisi di U1. Analizzando l'anda-

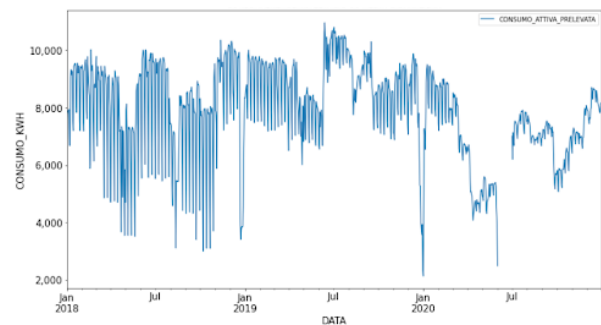


Figura 6: Consumo energetico giornaliero (edificio U6 Milano Bicocca)

mento della serie aggregata giornalmente visibile nella Figura 6, si nota che sembrano esserci diversi picchi che potrebbero indicare un qualche grado di stagionalità. Osservando i grafici di autocorrelazione acf e pacf, Figura 7, questa ipotesi sembra essere confermata in corrispondenza dei picchi ogni 6 lag. Inoltre, il test di Dickey-Fuller rifiuta l'ipotesi di stazionarietà in media, accettando l'ipotesi nulla di non stazionarietà in media. Poiché le stime ottenute tra-

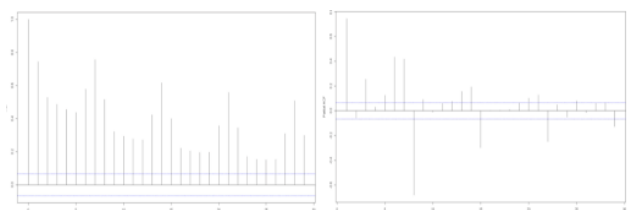


Figura 7: Grafici di autocorrelazione (ACF) e autocorrelazione parziale (PACF)

mite i classici modelli stagionali SARIMAX, una volta reso stazionario il modello, non risultano essere troppo affidabili in quanto tendono alla media dopo poche previsioni, si è optato per

un modello che utilizza un approccio differente: “prophet”.

Le stime ottenute che poi verranno utilizzate per la successiva analisi e la dashboard sono riportate nella Figura 8. Al fine di sviluppare il

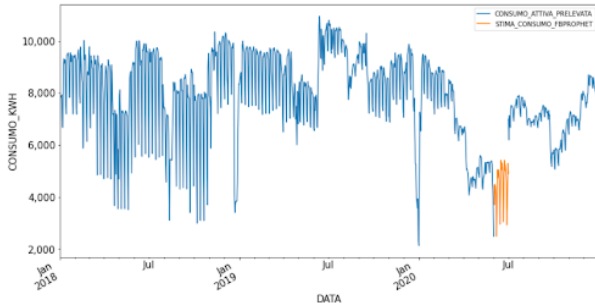


Figura 8: Stima dei consumi di giornalieri di giugno 2020 per l'edificio U6

modello si è deciso di procedere con un'aggregazione settimanale dei dati, inoltre, si suddiviso il dataset in training set, contenente i consumi energetici dalla 1^a settimana (01/01/2018) alla 149^a settimana (02/11/20) e test set, contenente i valori rimanenti. Osservando il grafi-

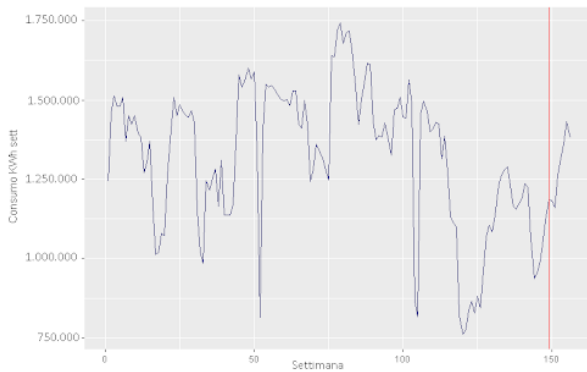


Figura 9: Consumo energetico settimanale (edificio U6 Milano Bicocca)

co mostrato nella Figura 9, è possibile ipotizzare che la serie necessita una differenziazione affinché sia garantita la stazionarietà. Analizzando i correlogrammi emerge una forte presenza di autocorrelazione. Infatti, come si osser-

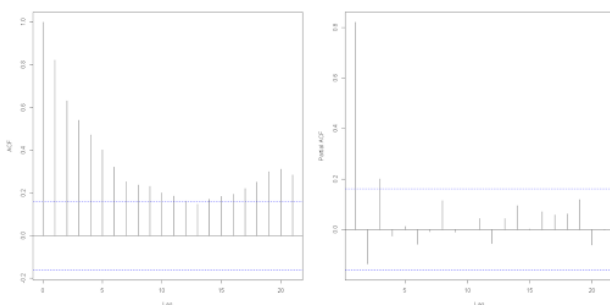


Figura 10: ACF e PACF serie dei consumi differenziata

va dai grafici nella Figura 10, i valori sono tutti al di fuori dalle bande di confidenza. Applicando il test di Dickey-Fuller si ha una conferma di quanto appena visto accettando l'ipotesi nulla con un p-value pari a 0.521; dunque, potrebbe rivelarsi utile applicare una differenza di un periodo.

Applicando la funzione del software R “auto.arima()”, contenuta nel pacchetto “forecast”, che individua tra tutti i possibili modelli quello con AICc minore, ci viene suggerita l'applicazione di un ARIMA(1,1,2), tutto a conferma delle ipotesi precedenti. Il modello stimato può essere esplicitato formalmente come segue:

$$\Delta y_t = a_1 \Delta y_{t-1} + \epsilon_t + b_1 \epsilon_t + b_2 \epsilon_t$$

dove $\Delta y_t = (y_t - y_{t-1})$

con coefficienti:

AR1	MA1	MA2
0.53928	-0.533742	-0.344133

É utile sottolineare come i parametri siano tutti significativi in quanto il valore assoluto del rapporto tra coefficiente e standard error risulta maggiore di due. Altre informazioni di diagnostica e di metriche di performance verranno fornite nel sottoparagrafo che segue. Le stime ottenute vengono invece riportate nel grafico rappresentato nella Figura 11.

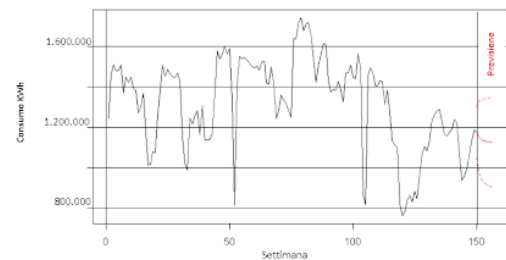


Figura 11: Consumo energetico settimanale con previsioni per la seconda ondata pandemica

7.2.1 Risultati

Considerate le stime ottenute nel paragrafo precedente, è fondamentale valutare la bontà dei modelli con un'opportuna metrica, in particolare si è fatto affidamento su un indice che considera l'errore di previsione. Quest'ultimo

viene valutato come la differenza tra la quantità effettiva di Kwh e il valore stimato ($E_t = D_t - P_t$):

MAPE (Errore percentuale medio assoluto)

: media aritmetica dei rapporti tra il valore assoluto degli errori di previsioni e la domanda effettiva verificatasi in n intervalli di uguale durata fino ad arrivare al periodo t rispetto al quale è stata effettuata la previsione.

$$MAPE_t = \frac{\sum_{i=0}^{n-1} \frac{E_{t-i}}{D_{t-i}}}{n} \quad (5)$$

Valutando la performance del modello stimato per U1, si ha:

- MAPE=15.88 quindi, in media, la differenza tra valori previsti e teorici è del 15.88%

Per quanto riguarda la stima con il metodo rolling la valutazione della performance restituisce:

- MAPE=5.84 quindi, in media, la differenza tra valori previsti e teorici è del 5.84%

Considerando ora il modello stimato per U6 è fondamentale, prima di valutare la metrica della performance, analizzare l'incorrelazione e la normalità dei residui.

Partendo dall'autocorrelazione essi risultano incorrelati poiché la statistica test di Durbin-Watson cade nella regione di accettazione dell'ipotesi nulla (incorrelazione dei residui). Valutando invece la normalità, considerato il test statistico di Jarque-Bera, emerge che i residui non seguono una distribuzione normale.

Nonostante ciò, è importante notare che tali modelli presentano una certa robustezza alla non normalità. Inoltre, è possibile affermare che, osservando il grafico del qq plot (Figura 12), c'è un leggero scostamento dei punti dai quantili teorici sulle code della retta, questo può significare la presenza di diversi outlier che impattano sui diversi test di normalità, sensibili a tali valori. Una volta effettuata tale analisi è possibile valutare la performance del modello sul test set analizzando il MAPE:

- MAPE=6.85 quindi, in media, la differenza tra valori previsti e teorici è del 6.85%

In conclusione, confrontando le stime teoriche con i consumi effettivi degli edifici non si evince una sostanziale differenza durante la pandemia; uno dei possibili motivi è attribuibile

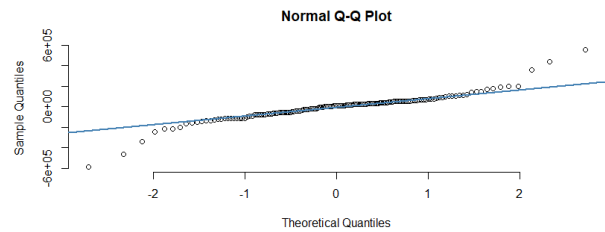


Figura 12: QQ plot dei residui del modello ARIMA(1,1,2)

al velo di incertezza dello shock causato dal Covid-19.

7.3 Dashboard di monitoraggio costi e consumi

In questo ultimo paragrafo di analisi, ci si è concentrati sullo sviluppo di una dashboard interattiva che permette il monitoraggio degli andamenti dei consumi energetici, individuando delle apposite stime che identificano il costo economico e l'impronta di CO_2 rilasciata dall'edificio nei diversi mesi dell'anno considerato.

Prima di procedere alla costruzione dei grafici interattivi è stato necessario stimare sia il prezzo per kWh sia il fattore moltiplicativo che fornisce una rappresentazione dell'impatto dei consumi sull'ambiente.

Costi

Come già accennato nel paragrafo 6, per la stima del prezzo si è fatto riferimento al sito web segugio.it che riporta in forma tabellare i differenti costi della componente materia prima energia nei diversi trimestri dal 2018 al 2020, Figura 13. È opportuno far notare che vengo-

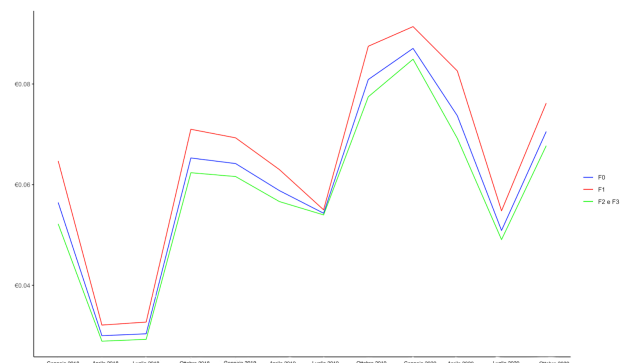


Figura 13: Andamento trimestrale del prezzo dell'energia

no messe in relazione 3 colonne, riportanti 3 differenti fasce di prezzo:

- F0: prezzo in €/kWh nella fascia unica giornaliera.
- F1: prezzo in €/kWh nella fascia diurna.
- F2 e F3: prezzo in €/kWh nella fascia notturna e weekend.

Tenendo conto che il prezzo è fortemente dipendente dall'operatore e dal contratto che ha stipulato l'università, si è preferito prendere in considerazione solo i costi presenti nella fascia unica. Infine, per ottenere la nuova *feature*, si è svolta una semplice moltiplicazione tra i consumi dei differenti trimestri ed i prezzi associati.

CO2

Per stimare le emissioni di CO_2 associate ai consumi energetici, invece, si è dovuto svolgere un lavoro qualitativo, al fine di trovare un fattore di conversione, da kWh a $kg\ CO_{2eq}/kWh$, che rendesse il più attendibile possibile tale stima. A tal proposito, si è scelto di analizzare diversi report di sostenibilità aziendali, al fine di identificare l'azienda che in Lombardia fosse più adatta a considerarsi come benchmark. Dal momento che nessuna azienda manifesta un esplicito collegamento diretto con l'Università Milano-Bicocca, la scelta è ricaduta sulla Polytech Technology Srl in quanto, in territorio lombardo, presenta dei valori in merito di sostenibilità e salvaguardia per l'ambiente, fortemente allineati a quelli dell'Ateneo. Infatti la Politech Technology srl, sin dalla sua fondazione, ha posto al centro della propria mission il riciclaggio dei materiali plastici e il continuo aggiornamento delle macchine di produzione volte ad un sempre maggiore risparmio energetico. Inoltre, nel 2017 ha commissionato uno studio di valutazione sulla quantità di CO_2 e gas equivalenti emessi nell'atmosfera dalla propria attività con lo scopo di compensare questa produzione attraverso la forestazione di terreni nell'area di Brescia. Da questo report è risultato un fattore di conversione per l'energia elettrica pari a $0.4121kgCO_{2eq}/kWh$. [8]

In conclusione, si è stimata la nuova ed ultima *feature* convertendo i consumi mediante il fattore appena trovato.

Aggiunte queste ultime informazioni al dataset iniziale si è deciso di utilizzare un software di visualizzazione, Tableau, per costruire un'infografica interattiva per gli edifici in analisi.

7.3.1 Risultati

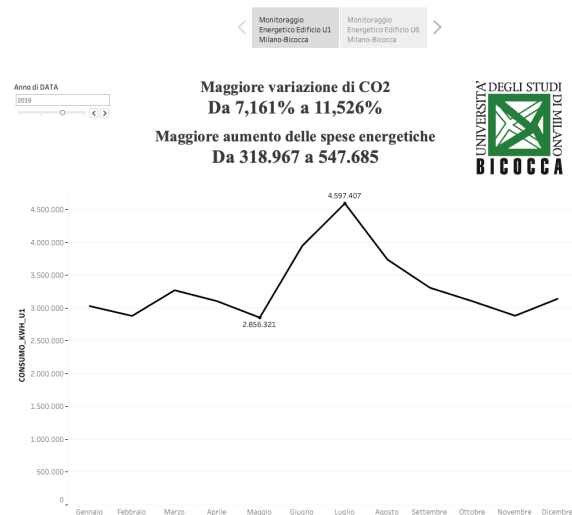


Figura 14: Dashboard di monitoraggio di costi e consumi energetici

Come si può osservare dalla Figura 14, il grafico presenta l'andamento dei consumi mensili al variare degli anni, inoltre, segnala all'interno del periodo considerato qual è stato il maggior aumento in termini di emissioni e di costi all'interno dell'anno. È interessante notare come queste variazioni, per entrambi gli edifici, si presentino sempre con dei picchi nei mesi di luglio per i periodi dal 2018-2019, mentre nel periodo pandemico non risulta esserci nessun picco particolarmente alto. Tuttavia, come ci si potrebbe aspettare, in tale anno i consumi sono leggermente più elevati per gennaio e febbraio. In seguito verranno effettuate delle ipotesi per provare a spiegare i motivi dei picchi appena individuati.

8 Conclusione e possibili sviluppi

Grazie alle analisi svolte e ai risultati ottenuti in questo elaborato, è possibile avere un'idea più chiara dell'impatto dei consumi energetici sull'Ateneo.

La cosa che fin da subito colpisce il lettore è che durante il Covid-19 e quindi la DAD, entrambi gli edifici avrebbero dovuto avere consumi molto più ridotti rispetto a quelli mostrati nei grafici di analisi, evidenziando un'inversione dell'andamento dal periodo di chiusura dell'Università. Infatti, guardando i consumi di U6 si vede chiaramente che dallo scoppio della pandemia c'è stato un calo nei consumi, mentre per quanto riguarda U1 un decremento dei consumi non

è così visibile. Questa non evidente diminuzione di consumi di U1 potrebbe essere dovuta a vari fattori, ad esempio, impianti più vecchi rispetto a U6, POD non funzionanti correttamente o non affidabili per quanto riguarda la rilevazione. Essendo un edificio più piccolo dell'altro, l'eventualità che non siano state attuate strategie di consumo mirato nel periodo Covid-19 non ha permesso di ottimizzare al meglio la potenza assorbita.

Per quanto riguarda l'edificio principale dell'Università degli studi di Milano-Bicocca in entrambe le ondate pandemiche si nota un calo dei consumi. Nonostante ciò, non ci riteniamo soddisfatti della posizione dell'Ateneo[3] perché ad oggi dobbiamo fare sempre più attenzione al tema della sostenibilità.

A nostro parere l'Università dovrebbe attuare strategie mirate per la riduzione dei consumi; una soluzione potrebbe essere un piano basato su investimenti, con lo scopo di aggiornare tutti i macchinari che assorbono troppa potenza ma che sono poco efficienti, in modo da ridurre una parte dei costi e consumi fissi. Ad esempio, installando impianti maggiormente efficienti di produzione di energia frigorifera, strategia suggerita dall'evidente consumo elevato nei mesi di luglio.

In conclusione, oltre a cercare maggiori sforzi per ridurre i consumi, sarebbe utile migliorare gli strumenti di monitoraggio degli impianti, attraverso una conoscenza dell'operatore con cui l'università ha sottoscritto il contratto per l'energia e dei fattori di conversione che permetterebbero di ottenere il Carbon Footprint effettivo. Inoltre, si potrebbero perfezionare i modelli di stima mediante tecniche di intelligenza artificiale come le reti neurali, ad esempio il modello "Long Short-term Memory Network".

Tutto ciò allo scopo di riuscire a capire meglio quali siano i veri benefici di queste azioni, sia da un punto di vista ambientale che economico.

9 Bibliografia

Riferimenti bibliografici

- [1] National Bureau of Statistics of China, *China Statistical Yearbook in 2008*, Beijing, 2009
- [2] Christopher Uhl and Amy Anderson, *Green Destiny: Universities Leading the Way to a Sustainable Future*

- [3] <http://greenmetric.ui.ac.id/rankings/overall-rankings-2020>, consulted on 10/06/2021
- [4] Università degli Studi Milano-Bicocca, *REPORT di SOSTENIBILITÀ 2020*
- [5] Università degli Studi Milano-Bicocca, *Energy management*, <https://www.unimib.it/ateneoenergy-management>
- [6] <https://tariffe.segugio.it/guide-e-strumenti/domande-frequenti/quanto-costa-un-kwh-di-energia-elettrica.aspx>, consulted on 16/06/2021
- [7] <https://www.politec-srl.com>, consulted on 16/06/2021
- [8] Ing. Chiara Caterini, *CALCOLO DELLA CARBON FOOTPRINT PER L'ABBATTIMENTO DELLE EMISSIONI TRAMITE PIANTUMAZIONE*, 2017, pag. 10