

# BICOCCA-ENERGY ASSESSMENT: TRA PANDEMIA, SPRECHI E NUOVE PROPOSTE DI GESTIONE ENERGETICA

Autore 1: Emanuele, Marnati, e.marnati@campus.unimib.it  
Autore 2: Marco, Branciforti, m.branciforti@campus.unimib.it  
Autore 3: Valerio, Schips, v.schips@campus.unimib.it

## Indice

Indice .....	1	Identificazione di eventuali macro-anomalie .....	8
Indice Figure .....	2	3. Trasformazione Box-Cox .....	8
Sinossi .....	2	4. Differenze tra tecniche di analisi di serie storiche .....	9
Parole Chiave .....	2	5. Identificazione del modello ARIMA aderente al dato .....	9
Introduzione .....	2	6. Modello di regressione con temperatura, apertura/chiusura ed effetto Covid .....	9
Scopo e obiettivi .....	3	7. Modelli ARIMA e ARMAX .....	9
Aspetti metodologici .....	3	8. Decomposizione e ricomposizione della serie con MSTL .....	10
Dati .....	4	9. Modello ARIMA U1 e confronto .....	11
1. Data Ingestion and Aggregation .....	4	10. Modello NEURALE U6 e confronto .....	12
2. Feature Selection .....	4	Risultati .....	13
3. Feature Extraction .....	4	1. Integrazione dei missing values di U6 .....	13
4. Data Integration .....	5	2. Scomposizione e ricomposizione di U1 e U6 .....	13
5. Missing Values .....	5	2. Regressione e confronto tra ARIMA e ARMAX .....	14
6. Stima volumetrica degli edifici .....	5	3. Confronto con dati Covid e non Covid .....	14
Analisi del processo di trattamento dei dati .....	5	4. Proposta energy manager .....	14
1. Analisi grafiche e test .....	5	Conclusione .....	14
2. ....		Possibili sviluppi migliorativi .....	14

Riferimenti bibliografici .....	15
1. Bibliografia.....	15
2. Sitografia.....	15
3. Link pubblico al codice GitHub .....	15

## Indice Figure

Figura 1: Confronto U1-U6 .....	6
Figura 2: Seasonal plot U1 .....	6
Figura 3: Seasonal plot U6 .....	6
Figura 4: Subseries plot U1.....	7
Figura 5: Subseries plot U6.....	7
Figura 6: Residuals Analysis U1 .....	7
Figura 7: Residuals Analysis U1 .....	7
Figura 8: Partial AutoCorrelation Function U1 .....	7
Figura 9: Partial AutoCorrelation Function U6 .....	8
Figura 10: Tabella consumi mensili .....	8
Figura 11: Confronto ARIMA e ARMAX .....	10
Figura 12: Decomposizione multistagionale U1 18-19 .....	11
Figura 13: Decomposizione multistagionale U6 18-19 .....	11
Figura 14: Serie storica U6 trasformata .....	11
Figura 15: Forecast U1 con bande di confidenza 80% - 95%.....	12
Figura 16: Covid vs. Non Covid U1 ARIMA.....	12
Figura 17: Covid vs. Non Covid U6 Neural Network---	13
Figura 18: Covid vs. Non Covid U1 Neural Network---	13

## Sinossi

La pandemia provocata dal virus COVID 19 ha, come da disposizioni ministeriali, obbligato alla chiusura della maggior parte delle attività didattiche in presenza. L'ateneo Milano-Bicocca è comunque rimasto aperto per permettere la continuità delle attività fondamentali a garantire una fruizione della didattica anche a distanza. Con questa analisi statistica, che interesserà i dati energetici degli edifici Atlas (U1) e Agorà (U6) per il triennio 2018-2019-2020. Ci si pone in prima istanza l'obiettivo di comprendere le differenti modalità di consumo energetico dei due edifici e, alla luce di queste, individuare eventuali **inefficienze**. In seguito si è analizzato come la temperatura esterna, la capienza massima degli edifici e l'evento **Covid**, influiscano sui **consumi energetici**. Sarà ricercato un modello statistico che meglio riesca ad approssimare il dato, adoperando

una granularità prevalentemente giornaliera. Si è deciso di adottare più modelli di analisi di **serie storiche** (per lo più modelli ARIMA) per fornire previsioni. L'ipotesi al vaglio è che ci sia stato un **risparmio** inteso come calo fisiologico di consumo durante la pandemia. In seconda istanza si indagherà l'impatto in termini di risparmio economico degli efficientamenti energetici, nello specifico il teleraffreddamento, sui consumi degli edifici. Ciò sarà compiuto attraverso un processo di decomposizione delle serie storiche, manipolazione delle componenti e infine ricomposizione. In ultima istanza sarà presentato il tentativo di individuare, tramite dei modelli neurali, una stima economica approssimativa del minor costo di consumo dovuto alla pandemia. Sarà possibile, anche alla luce di come il consumo ha reagito all'evento pandemico, presentare all'**energy manager** dell'ateneo l'analisi e valutare insieme a lui una possibile proposta di miglioramento.

## Parole Chiave

Consumi; Covid; Energy Manager; Inefficienze; Serie Storiche.

## Introduzione

Ciò che viene implementato in maniera sequenziale grazie al software open source R è presentato con questa lista di operazioni e documentato nell'allegato apposito. Le librerie utilizzate in fase di analisi saranno: *readxl*, *dplyr*, *ggplot2*, *forecast* e *tseries*.

1. Preprocessing dei dati. Si è voluto semplificare il processo di integrazione delle numerose tabelle che compongono il dataset complessivo, attraverso la scrittura di 8 funzioni per il caricamento e l'aggregazione in R. Questo ha permesso di estrarre e comporre agevolmente l'insieme di dati nella struttura che più si addice al tipo di analisi effettuata (questa fase si ottiene dopo un pre-processing manuale per eliminare tabelle doppie e uniformare estensioni e nomenclatura). Con questo approccio il fruitore della presente ricerca è in grado di poter riprodurre il lavoro partendo dal dato grezzo, aggregandolo a proprio piacimento in modo da

ripercorrere l'analisi e quindi ottenere riscontri empirici nuovi che non sono stati analizzati in questo paper.

2. Esplorazione. Un'approfondita osservazione preliminare dei dati a disposizione ha fatto emergere la presenza di un dato mensile errato e inconsistente per il mese di giugno 2020: nello specifico si è notato come il contatore indicato per l'edificio U6 fosse in realtà il codice POD corrispondente a quello di U1. Non riuscendo a recuperare il dato mancante e non potendo operare alcun confronto temporale completo senza questo, si è deciso di reintegrare la serie con un metodo previsionale alternativo che sarà a breve mostrato (si veda paragrafo 5. *Missing Values*).

3. Modelli. Nel corpo centrale del paper, dopo analisi descrittive basate su una aggregazione mensile, si presenterà il modello ARIMA a partire dalla stima dei suoi parametri. In secondo luogo, si mostrerà come poter applicare serie storiche prelevate fonti di dati esterne nel medesimo modello. Sarà poi presentata una tecnica di decomposizione con un metodo MSTL e l'utilizzo di una tecnica per attribuire la componente stagionale di un edificio ad un altro. Lo scopo di questa parte, la più sperimentale, sarà quella di analizzare la differenza di consumo degli edifici. Si otterrà un tipo di serie che presenta una nuova stagionalità e permette una quantificazione economica del risparmio potenziale avuto grazie all'adozione del teleraffrescamento come per l'edificio U6.

4. Analisi sull'impatto del Covid. In conclusione, si è deciso di proporre un confronto tra i dati storici che sono ricollocabili al periodo Covid, da inizio pandemia Marzo 2020, con un dato previsionale ottenuto da un modello ARIMA e una rete neurale NNAR. In sostanza questa proposta permetterà di valutare in modo approssimativo i risparmi complessivi in termini di consumi avuti durante la pandemia.

## Scopo e obiettivi

L'intento del lavoro è quello di appurare l'entità del consumo energetico dei due edifici presi in esame con un focus prima sulla fase pre-pandemica e poi pandemica (limitata al periodo 2020). Inoltre, si vuole integrare fonti informative esterne, come il meteo e la stima volumetrica, per poi eventualmente proporre un intervento strutturale al reparto che si occupa della gestione energetica all'interno dell'ateneo. La ricerca trova un naturale interlocutore nell'energy manager dell'ateneo o in colui che si occupa della gestione dei costi fissi e variabili degli edifici. La componente di costo fisso va inevitabilmente tenuta in considerazione in quanto, essendo l'università un edificio pubblico, sconta la presenza di regolamenti generali che poco interpretano le singole esigenze del singolo luogo. Il lavoro pone logicamente l'intento terziario di essere anche un monito per eventuali sprechi energetici qualora la fase pandemica si protragga o si ripresenti nel prossimo futuro.

## Aspetti metodologici

Il problema da analizzare si presenta certamente vasto e ricco di possibili integrazioni dati. Tuttavia, per avere una visione focalizzata sul dato energetico, si è deciso di integrare le variabili del dataset di partenza con alcune features in qualche modo legate al contesto del consumo: il meteo nelle sue componenti indicate in seguito, l'apertura dell'ateneo tramite l'identificazione dei giorni di festa e weekend (si veda calendario universitario in *Riferimenti bibliografici*) e l'effetto vero e proprio del Covid. Inizialmente era stata presa in considerazione anche una variabile che fosse in grado di distinguere "zona gialla", "zona arancione" o "zona rossa"; successivamente è stata volutamente scartata in quanto assume poco rilievo poiché questa disposizione è entrata in vigore a novembre e dunque avrebbe interessato una porzione limitata di dati. Nel caso della Lombardia la variabile avrebbe inoltre assunto valore costante di "zona rossa" per il 2020 quindi non avrebbe apportato alcuna ulteriore informazione.

Le variabili che incidono sul consumo e che non verranno prese in esame perché non disponibili

sono molteplici. Anzitutto l'università ha una duplice funzione: non solo è luogo di studio ma è al contempo luogo di lavoro per tutto il personale, primi tra tutti i professori insieme a tecnici e responsabili dell'amministrazione. Costoro, specie i professori, ricercatori e dottorandi, hanno visto il diritto al lavoro tutelato e dunque non è mai stato loro precluso l'accesso agli uffici. Questa considerazione potrebbe già di per sé spiegare il livello di consumo energetico in fase di lockdown totale. Ancor più rilevante è la conformazione planimetrica a partire da fondamenta, muri e architettonica degli edifici: l'edificio U1 è di fatto meno esposto alla luce naturale rispetto ad U6 e ciò comporterebbe un maggior consumo solo per riuscire a mantenere la temperatura costante al suo interno. Anche l'associare il consumo ad un particolare tipo di utenze è un lavoro arduo: alcuni edifici usano teleriscaldamento e altri sfruttano il gas; per il raffreddamento, alcuni edifici usano il teleraffrescamento (energia termica), altri usano energia elettrica. I vettori energetici usati per il riscaldamento e il raffrescamento sono quindi variegati e qualsiasi approccio deduttivo deve necessariamente essere cauto. Si è inoltre appurato come, per il raffreddamento, solo l'edificio U1 usa energia elettrica, mentre U6 utilizza energia termica. Il lavoro deve quindi soppesare queste componenti con quanta più confidenza possibile.

Non si hanno inoltre informazioni precise su: numero di uffici presenti, metri cubi rispettivi, ampiezza dei muri e loro coibentazione, ammodernamenti attuati nel corso degli anni, numero di finestre, meccanismo di termoregolazione, modalità di accesso a biblioteche di cui U6 ospita la principale, laboratori scientifici (presenti in U1) e informatici, mense e bar, segreterie sempre attive, bagni e altri tipi di infrastrutture che potrebbero o meno interessare il consumo. Tutto questo senza contare la semplice policy di gestione delle temperature, regolata internamente o a livello regionale piuttosto che ministeriale.

L'analisi scontrerà verosimilmente la mancanza di alcuni di questi dati, il che non necessariamente implica uno svantaggio. Tipicamente la vastità di dati definita "data flooding" porta a bias, inconsistenza e scarsa interpretabilità dei risultati. Nel piccolo di questa analisi, si potrebbe dunque

nascondere dell'informazione essenziale e attendibile e dunque di valore.

## Dati

### 1. Data Ingestion and Aggregation

Il primo step è stato quello di strutturare un metodo di aggregazione dati che fosse scalabile ed efficiente. I 72 file (12 mesi × 3 anni × 2 edifici) sono stati standardizzati, riordinati ed aggregati mediante l'utilizzo di 8 funzioni implementate con il software open source R. Tramite queste è possibile regolare tempo di riferimento, granularità di aggregazione ed edificio di interesse, riuscendo a ristrutturare a proprio piacimento i dati. Inoltre, è stata implementata anche la possibilità di integrare i valori meteo chiamando una semplice ulteriore funzione.

### 2. Feature Selection

Dell'ampio dataset fornito dall'ateneo si è deciso di prendere in considerazione il dato energetico aggregabile (CONSUMO\_ATTIVO\_PRELEVATO) e di non considerare la potenza massima e il consumo reattivo induttivo. Il campo POD è un valore costante che indica il codice del contatore di ogni edificio, dato utile solo all'identificazione dell'edificio. È presente anche un campo definito come "tipo\_dato" che riporta sempre lo stesso valore, probabilmente riferito all'ambito energetico, e che quindi viene scartato. La colonna contenente l'ora legale non viene integrata nel campo riferito all'orario in quanto i dati già scontano questo fenomeno. Si stima di fatto che l'ora legale (secondo l'azienda [Terna](#)) abbia permesso nel 2019 risparmi per il sistema elettrico nazionale pari a circa 100 milioni di euro e minori emissioni in atmosfera di CO<sub>2</sub> per 250 mila tonnellate. Anche i consumi dell'università sono stati sicuramente interessati da un risparmio in termini di consumo energetico che sarebbe analizzabile, ma non è oggetto di ricerca in questo paper.

### 3. Feature Extraction

Nel processo estrattivo dei dati si è deciso di operare nel seguente modo. In primo luogo, si è sfruttato il dato a granularità minima (15 minuti)

da cui è stata generata la colonna energia consumata, passando da kilowatt a kilojoule tramite la moltiplicazione della prima per il numero di secondi contenuto in 15 minuti, ovvero 900. In secondo luogo, si è aggregato il dato CONUSMO\_ATTIVA\_PRELEVATA a granularità minima e si sono ottenuti una serie di nuove variabili: consumo medio, consumo mediano, consumo massimo e consumo minimo. Si adopererà principalmente il valore medio e, per la tabella dei confronti energetici, si adopererà il valore di consumo energetico totale in kilojoule.

#### 4. Data Integration

Per poter rispondere alle domande e alle problematiche poc'anzi espresse, si è deciso di provvedere ad una integrazione dati che interessasse l'aspetto meteorologico, i cui dati sono stati ottenuti dalla fonte [meteostat](#). La scelta è ricaduta sui valori provenienti dai rilevatori situati presso Sesto San Giovanni. Infatti, la città è di fatto la più prossima e disponibile all'ubicazione dei due edifici presi in analisi, approssimativamente più vicina rispetto alla vastità della città metropolitana di Milano (gli edifici sono situati quasi al confine tra Milano e Sesto San Giovanni). I dati con granularità giornaliera sono composti da temperatura media, temperatura massima e temperatura minima, precipitazioni piovose, precipitazioni nevose, velocità del vento, direzione del vento, pressione atmosferica e temperatura del sole. È stata mantenuta solo la temperatura media perché unica variabile significativa individuata tramite il VIF test come non collineare alle altre.

#### 5. Missing Values

Il dataset presenta essenzialmente valori mancanti dal mese di luglio 2020 in poi per quanto riguarda la potenza massima, ma essendo questo un dato poco rilevante, questa mancanza non porta ad alcun problema. Il principale buco temporale che ha un ruolo fondamentale nell'analisi è il mese di giugno 2020 per l'edificio U6. Nella fase di data ingestion sono stati erroneamente associati i dati dell'edificio U1. Nell'ipotesi di dati essenziali al monitoraggio dei consumi, si è deciso di colmare il buco temporale nel seguente modo: viene implementata una tecnica che sfrutta una sliding window, cioè una

finestra temporale che mantiene al proprio interno sempre lo stesso numero di osservazioni giornaliere. Questa consiste nel variare ad ogni step i dati su cui allenare il modello, aggiungendo la rilevazione successiva alla più recente all'interno dei dati di training e rimuovendone l'ultima, fornendo in questo modo una previsione dei successivi 7 giorni. Quindi ad eccezione dei primi 6, ogni giorno avrà 7 valori di previsione, ciascuno fornito da un modello allenato su dati leggermente differenti. Dato che l'incertezza della previsione aumenta al crescere dei giorni stimati dallo stesso modello, la previsione del primo giorno sarà più affidabile rispetto alla previsione dei giorni successivi. Si è deciso così di definire la previsione finale giornaliera come media pesata delle previsioni dello stesso giorno, in cui il peso di  $\hat{y}_{+n}$  sarà doppio rispetto al peso di  $\hat{y}_{+n+1}$ , in modo da rendere più robusta la stima.

#### 6. Stima volumetrica degli edifici

Si è deciso di adoperare gli strumenti topografici forniti da Google per potere ottenere una stima approssimativa dei due edifici in analisi. Per ottenere una misura di volume il più coerente possibile si è deciso di misurare le aree dei due edifici, tenendo conto dei 4 cortili interni di U6 (che sono stati sottratti alla stima) ma senza stimare il volume di ciò che sta sotto il livello del suolo e quindi senza valutare il consumo di parcheggi e passaggio intermedio tra edifici.

Allo stesso modo si è misurata l'altezza. Il volume di U6 è pari a 207104 m<sup>3</sup> e il volume di U1 è pari a 43959 m<sup>3</sup>. La proporzione dei due edifici è pari a 4,7.

### Analisi del processo di trattamento dei dati

#### 1. Analisi grafiche e test

Analizzando in prima battuta il grafico delle due serie storiche in *Figura 1* ci si rende conto di come esse non siano stazionarie, ma pare interessante intravedere come ci sia un pattern ricorrente nell'andamento di U1. Il meccanismo elettrico di raffreddamento comporta di fatto dei picchi considerevoli di consumo in prossimità dei mesi

estivi. Diverso pare essere per U6, il quale ha un pattern più uniforme sia nella stagione calda che invernale. I massimi picchi si sono avuti nell'estate del 2019 e questo fenomeno potrebbe essere spiegato sia dalle alte temperature sia nella crescita del numero di studenti, per Bicocca decisamente un fenomeno di rilievo. Le serie scontano evidentemente un rumore di fondo e una stagionalità molto marcata per U1, un po' più smooth per U6.

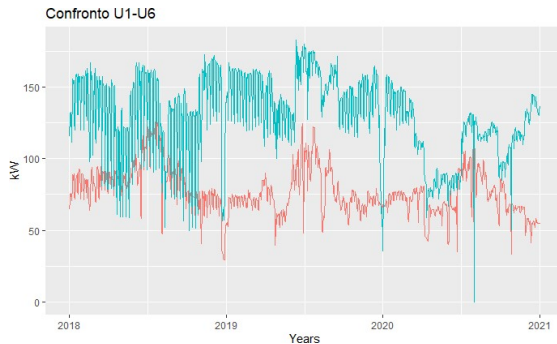


Figura 1: Confronto U1-U6

In primo luogo, è interessante notare la differenza in termini di consumo dei due edifici: il fenomeno non sorprende affatto in quanto U1 è decisamente più piccolo di U6 e di conseguenza il numero di uffici, accessi, aule ed altri fattori, è decisamente ridotto. Le due serie mostrano come il Covid abbia in linea generale diminuito i consumi, anche se parrebbe non sia stato un evento eccessivamente ribassista. Interessante notare come il rumore randomico attribuibile alle oscillazioni giornaliere di U6 sia calato con il Covid e come nell'anno 2020 questa serie presenti un andamento decisamente meno altalenante. Ciò è un'ulteriore evidenza di quanto accade, ovvero il consumo diminuisce e con esso anche la possibilità di valori estremi. Ogni picco anomalo verso il basso non è frutto di una avventata decisione da parte dell'energy manager, ma verosimilmente un calo di potenza o un guasto momentaneo. Se si analizzano le tre curve aggregando il dato mensile, sembra evidente come nei tre anni il comportamento dei due edifici sia stato diverso, come segue nelle figure qui riportate.

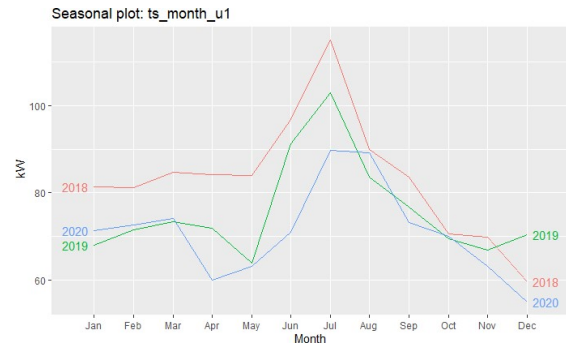


Figura 2: Seasonal plot U1

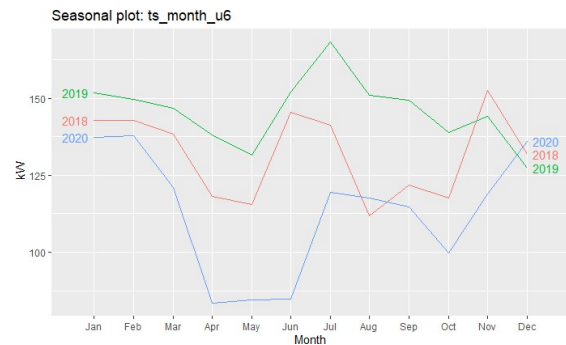


Figura 3: Seasonal plot U6

I due seasonal plot palesano un consumo elevato nei mesi estivi, specie per U1 come mostrato in *Figura 2*. Nei mesi invernali è U6 che, come mostrato in *Figura 3*, palesa un maggior consumo. La linea blu riferita all'ultimo anno mostra chiaramente l'impatto del Covid con la netta discesa di marzo. Un ulteriore metodo di analisi del fenomeno stagionale consiste nel prendere in esame l'andamento dello stesso mese nei 3 anni. Le figure 4 e 5 mostrano il risultato di questo tipo di approccio. Il range di variazione nei tre anni dell'edificio U1, rappresentato in *Figura 4* appare più moderato rispetto al range di variazione di U6 in *Figura 5*. In generale pare ci sia un netto calo nei mesi estivi del 2020 rispetto ai mesi degli anni precedenti. Inoltre, in U1 si registra un netto aumento dei consumi nel 2019 rispetto all'anno precedente, andamento completamente opposto rispetto all'altro edificio. Quest'ultima osservazione vale per tutti i mesi ad eccezione di dicembre, il quale assume un comportamento inverso in entrambi gli edifici.



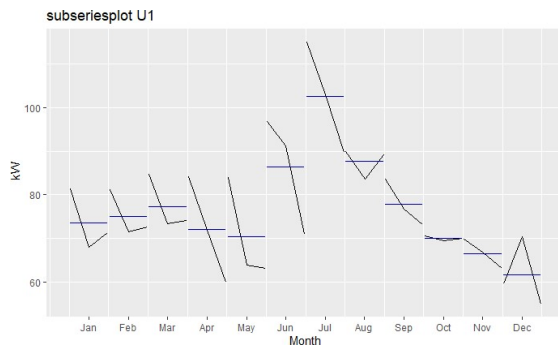


Figura 4: Subseries plot U1

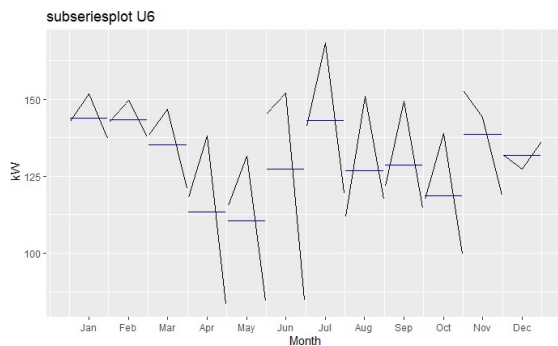


Figura 5: Subseries plot U6

In generale ciò che pare evidente è che i due edifici hanno un atteggiamento molto discostante, a parità di periodo dell'anno e consumo su più anni. Per U6 prevale una struttura "a punta" mentre per U1 prevale una struttura "a discesa".

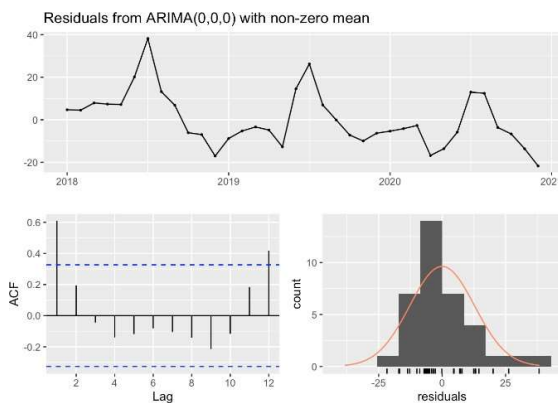


Figura 6: Residuals Analysis U1

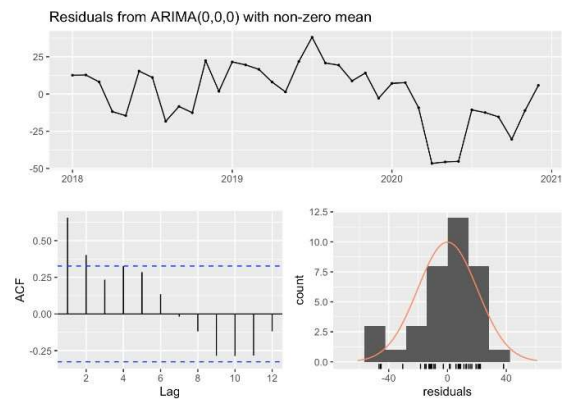


Figura 7: Residuals Analysis U1

Nella Figura 6 e nella Figura 7 viene invece riportata una analisi grafica dei residui di un modello ARIMA(0,0,0) e della valutazione della presenza di autocorrelazione. Il risultato è ottenuto adoperando i dati con aggregazione mensile per potere avere una chiave di lettura più fruibile rispetto al dato aggregato giornalmente. Si nota come la distribuzione dei residui sia solo lontanamente approssimabile ad una condizione normale nel caso di U1 e sia lievemente più normale nel caso di U6. I lag per l'aggregazione mensile esprimono un grado di auto correlazione di 1 e 2 rispettivamente. L'autocorrelazione parziale in Figura 8 e in Figura 9 indica invece un primo grado. Si prenderà atto del risultato grafico riferito alla granularità mensile ma, tuttavia, non si adopereranno Acf e Pacf per la determinazione del lag da adoperare nel modello ARIMA (che sarà stimato per granularità giornaliera e con la funzione di R chiamata *auto.arima*).

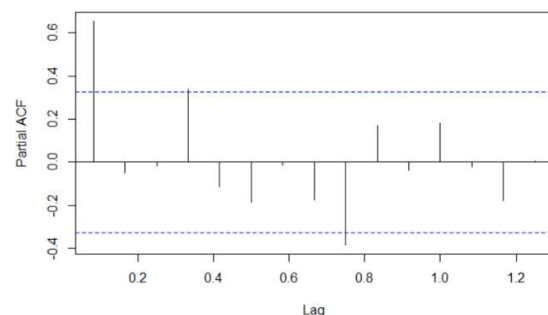


Figura 8: Partial AutoCorrelation Function U1

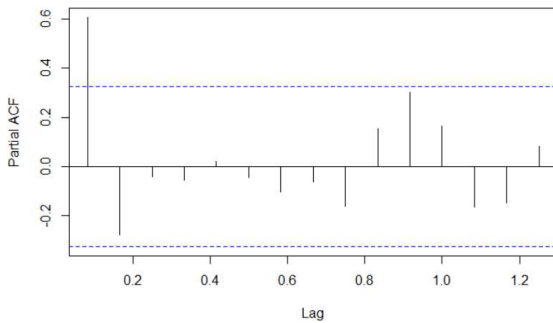


Figura 9: Partial AutoCorrelation Function U6

Dai grafici riportati inerenti alla due serie storiche, è evidente come la stazionarietà manchi, la realtà del dato consumo non può che mancare di questa peculiarità. Per completezza dell'analisi, si implementa un Augmented Dickey Fuller Test per avere un riscontro oggettivo che conferma inequivocabilmente il rifiuto dell'ipotesi nulla di stazionarietà della serie.

## 2. Identificazione di eventuali macro-anomalie

La prima analisi condotta riguarda le anomalie a livello macro dell'andamento dei consumi dei due edifici. Per poter avere una visione d'insieme, si prende a confronto il dato di consumo energetico cumulato di ogni mese e si calcola la percentuale di crescita o decrescita del consumo, facendo un confronto tra i 3 anni.

Mese	Q1	Energia Consumata - 2018	Energia Consumata - 2019	Energia Consumata - 2020	Energia Consumata - 2021	Energia Consumata - 2022	Energia Consumata - 2023	Delta U1	Delta U6	Delta U1	Delta U6
Media		239.221.805 kJ	345.978.344 kJ	359.366.310 kJ	340.808.425 kJ	347.476.485 kJ	397.800.532 kJ	-6,9%	13,8%	-6,9%	13,8%
Gennaio		218.152.349 kJ	342.590.185 kJ	351.051.161 kJ	408.244.321 kJ	351.191.321 kJ	347.721.623 kJ	-10,0%	9,0%	9,0%	-6,0%
Febbraio		194.291.281 kJ	341.772.351 kJ	372.515.141 kJ	382.177.488 kJ	381.808.281 kJ	345.110.411 kJ	-11,0%	6,0%	6,0%	-7,0%
Marzo		224.444.402 kJ	360.347.481 kJ	385.441.411 kJ	361.538.891 kJ	361.538.891 kJ	361.538.891 kJ	-10,0%	6,0%	6,0%	-10,0%
Aprile		227.088.881 kJ	358.595.971 kJ	386.402.061 kJ	358.024.481 kJ	358.404.441 kJ	316.402.251 kJ	-14,0%	10,0%	-14,0%	10,0%
Maggio		224.788.881 kJ	358.688.131 kJ	371.578.241 kJ	382.581.211 kJ	381.538.131 kJ	282.248.141 kJ	-26,0%	13,0%	-26,0%	13,0%
Giugno		251.091.581 kJ	371.247.121 kJ	236.512.581 kJ	381.801.891 kJ	381.801.891 kJ	216.546.711 kJ	-42,0%	4,0%	-42,0%	4,0%
Luglio		307.488.481 kJ	378.578.521 kJ	279.444.421 kJ	485.476.211 kJ	341.325.181 kJ	310.947.801 kJ	-35,0%	10,0%	-35,0%	10,0%
Agosto		340.891.741 kJ	290.581.951 kJ	224.581.441 kJ	442.258.801 kJ	238.871.881 kJ	311.546.701 kJ	-7,0%	10,0%	6,0%	-10,0%
Settembre		224.891.301 kJ	323.584.771 kJ	384.651.301 kJ	387.551.201 kJ	381.521.481 kJ	297.511.841 kJ	-24,0%	22,0%	-24,0%	10,0%
Ottobre		184.421.211 kJ	321.510.271 kJ	385.541.401 kJ	371.571.241 kJ	381.581.441 kJ	271.564.421 kJ	-28,0%	2,0%	-28,0%	10,0%
Novembre		180.428.181 kJ	385.577.781 kJ	371.581.061 kJ	371.441.771 kJ	381.581.061 kJ	308.511.021 kJ	-19,0%	10,0%	-19,0%	10,0%
Dicembre		159.836.221 kJ	351.562.181 kJ	358.612.321 kJ	341.844.311 kJ	347.131.601 kJ	344.541.211 kJ	15,0%	10,0%	10,0%	6,0%

Figura 10: Tabella consumi mensili

Ciò che si riscontra dalla Figura 10, in particolare dal confronto dei delta di aumento o discesa del consumo, è che nel 2019 c'è stato un calo sostanziale del consumo che si riscontra solo nel caso di U1 e non U6. Questo dato è tuttavia interpretabile solo cautamente, in quanto una aggregazione di questo tipo comporta inevitabilmente una perdita di informazioni. Il dato che in realtà è di vero interesse è il delta 2020, il quale nel caso di U6 palesa una fortissima

discesa che nel mese di giugno si attesta attorno al 50%. U1 mostra invece una decrescita non molto consolidata per lo stesso mese, ma in linea con i dati di U6. Dunque, alcune considerazioni che vengono da sé, riguardando la reattività dell'edificio alla pandemia e al parziale fermo dei servizi che ne deriva, sono: U6 grazie alle tecnologie di termo-regolazione permette un maggior controllo dei consumi, U1 mostra inefficienze sotto questo punto di vista. Difficile da spiegare il motivo delle differenze di consumo tra il 2018 e il 2019, tuttavia il dato per U6 rispecchia una distribuzione più conforme all'aumento del numero di iscritti e dei servizi erogati. Alla luce del dato volumetrico che si è stimato, è interessante notare come l'edificio U1 dovrebbe dimostrare un livello di consumo energetico di circa 4-5 volte minore (stando ai rapporti determinati tra capienze massime o volumi) del suo parente maggiore U6, a parità di tutti i possibili fattori che potrebbero essere considerati. Quindi questo primo risultato già di per sé è indicatore di una qualche anomalia energetica per U1, edificio di costruzione antecedente al parente maggiore e per il quale il consumo medio mensile pare essere poco più della metà.

## 3. Trasformazione Box-Cox

In alcune delle tecniche che saranno presentate sono state proposte delle trasformazioni. Nel caso specifico si sarebbe potuto attuare una trasformazione per calendario in cui ad esempio si assegna un certo ruolo ai fine settimana, oppure un population adjustment dove si pondera per una misura di flusso di individui o in base alla capienza dell'edificio, o ancora una trasformazione matematica. La prima ipotesi è stata scartata in quanto i fine settimana sono ancora giorni abbastanza operativi per le università sebbene gli studenti tipicamente non abbiano lezione. La seconda ipotesi sarà mostrata nel capitolo di scomposizione e ricomposizione della serie mentre ciò di cui si scriverà ora è una trasformazione di Box-Cox. Solitamente utilizzata per rendere più simmetrica la variabile target, questa è basata sulla stima di un parametro  $\lambda$  ottenuto tramite il metodo della massima verosimiglianza, funzione utilizzata per minimizzare la misura RMSE (Root Mean Square Error). Definito tale parametro, la trasformazione avviene come segue:



$$w_t = \begin{cases} \log(y_t) & \text{if } \lambda = 0; \\ (y_t^\lambda - 1)/\lambda & \text{otherwise.} \end{cases}$$

Un buon parametro è quello che rende la dimensione della variazione stagionale più o meno la stessa per tutta la serie, poiché questo rende il modello di previsione più semplice. Calcolata con l'apposita funzione nel caso di U1  $\lambda$  assume valore  $1.410841 \approx 1$ , per U6 assume un valore approssimabile allo 0. Questo è un ulteriore riscontro di come U1 abbia una stagionalità più incostante in variazione rispetto ad U6, la tesi è la medesima dei paragrafi precedenti.

#### 4. Differenze tra tecniche di analisi di serie storiche

Uno dei principali problemi/opportunità dell'analisi storica è che non c'è un criterio assoluto di analisi di serie storiche, tant'è vero che oggi ci sono anche degli approcci neurali che permettono di analizzare egregiamente dati storici come i modelli Neural Network skip layer o LSTM, specie con molti importanti di dati. Tuttavia, se il tema principale è la previsione, i modelli ARIMA sono preferibili, mentre se il tema è la decomposizione, come nel caso di questa ricerca, sono invece i modelli strutturali ad essere presi in considerazione dato che sono anche dei buoni previsori. In questo lavoro si è cercato di spaziare tra vari metodi preferendo i primi: manipolare componenti e scomporre multistagionalità è infatti utile con questo tipo di set, quindi si è deciso di adoperare oltre al modello ARIMA, il modello MSTL per la decomposizione e un modello di rete neurale autoregressiva.

#### 5. Identificazione del modello ARIMA aderente al dato

Istruendo un modello autoarima, in grado di individuare autonomamente i parametri, si è ottenuto come modello per la serie U6 un ARIMA(5,1,2). I tre valori, identificati solitamente con le lettere p, d, e q, indicano: ordine della parte autoregressiva, grado del primo differenziale coinvolto e ordine della parte di media mobile. Addestrando invece per U1, con componente  $\lambda$  pari a zero, il modello risulta essere un ARIMA(3,1,1)(0,1,0). Quest'ultima contiene anche

i parametri P, D, Q nella seconda parentesi, segno che il modello abbia rilevato una componente di stagionalità non stazionaria.

#### 6. Modello di regressione con temperatura, apertura/chiusura ed effetto Covid

Un'analisi interessante è stata lo studio dell'impatto dei dati riferiti di meteo, chiusura/apertura ateneo e presenza/assenza del Covid. Tuttavia, si è dovuto inizialmente valutare l'integrazione di questo dato in analisi nel caso in cui si fosse ritenuto davvero esplicativo. Il modello regressivo è stato ottenuto eliminando tutte le componenti meteo inadatte dopo aver appurato la mancanza di collinearità tra variabili. I modelli risultano significativi per entrambi gli edifici, lo stesso vale per tutte le variabili esplicative, tranne per la temperatura per U6. Di fatti avendo un sistema di teleraffrescamento, l'edificio è meno sensibile alle temperature. Come atteso, i fattori inseriti nel modello non spiegano l'intera variabilità del fenomeno del consumo energetico, ma comunque una buona parte: l'indice R quadro risulta pari a circa il 41% per entrambi, risultati buoni per modelli abbastanza semplici come questi. In particolare, si stima che l'apertura di U6 comporti un aumento del consumo medio giornaliero di circa a 28 kW mentre per U1 si stima che l'aumento sia sensibilmente inferiore, cioè pari a circa 9 kW. Si stima inoltre che la presenza Covid per U1 comporti un consumo medio giornaliero minore, pari a circa 10kW, per U6 un consumo medio giornaliero minore, pari a circa 31 kW. U1 risente anche dell'aumento della temperatura, che per ogni aumento unitario favorisce l'aumentare del consumo di 1 kW (fattore invece non significativo per U6). I due modelli mostrano come U1 risenta meno dell'influenza Covid e dei giorni di apertura/chiusura. Data la significatività delle variabili proposte e la loro discreta potenzialità esplicativa, si è deciso di considerarle anche per un modello ARIMA.

#### 7. Modelli ARIMA e ARMAX

L'idea è quella di essere davanti a valori di serie storiche che potrebbero non essere spiegati unicamente dai valori passati della serie, ma che

molto probabilmente risentono dell'influenza esterna. Per riuscire a catturare adeguatamente la relazione esistente con questi valori si utilizza un modello di tipo ARMAX (AutoRegressive Moving Average with eXogenous terms), cercando perciò di inserire serie storiche diverse da quella data, definite variabili esogene, ma che si ritiene possano essere utili per la previsione di quest'ultima. Il modello che si crea è quindi:

$$x_i = a_1x_{i-1} + a_2x_{i-2} + \dots + c_1z_{i-1} + c_2z_{i-2} + \dots + b + \varepsilon_i$$

Dove  $z_1, \dots, z_n$  è un'altra serie storica (o più di una).

Per avere un metro di giudizio sulle performance del nuovo modello, viene effettuato un confronto con il modello ARIMA semplice, tramite l'osservazione di alcuni indici. Per il confronto si è deciso di adoperare solo l'edificio U1. Nella seguente *Figura 11* si riportano alcuni valori di bontà di stima per i due modelli precedenti:

Serie	RMSE	sigma^2	AIC	MAE
ARMAX	8.519	109	5520.8	4.169
ARIMA	8.919	120.1	5653.3	4.505

*Figura 11: Confronto ARIMA e ARMAX*

Oltre la varianza, le due misure più comunemente utilizzate si basano sugli errori assoluti o sugli errori quadratici. Quando si confrontano i metodi di previsione applicati a una singola serie temporale, il MAE (Mean Absolute Error) è comunemente usato in quanto facile da capire e da calcolare. Un metodo di previsione che minimizza il MAE porterà a previsioni della mediana, mentre minimizzando l'RMSE porterà a previsioni della media. Di conseguenza, anche il RMSE è ampiamente utilizzato, nonostante sia leggermente più difficile da interpretare. Un ulteriore indice è il criterio di informazione di Akaike (AIC), definito come:

$$AIC = -2\ln(L) + 2k$$

In cui  $L$  è la massima verosimiglianza, cioè il parametro con la più alta probabilità di

rappresentare correttamente la relazione tra input e output, mentre  $k$  è il numero di parametri.

Insieme al BIC, sono indici che servono a valutare un modello: a livello interpretativo l'AIC cerca di selezionare il modello che descrive nel modo più adeguato una realtà ignota e multidimensionale. Ciò significa che la realtà non è mai nel set di modelli candidati che vengono considerati. Il BIC invece cerca di trovare il modello vero tra un numero finito di candidati. A livello matematico invece, la differenza sostanziale tra l'AIC e il BIC è che il secondo penalizza maggiormente modelli più complessi. I modelli efficienti saranno quelli che minimizzeranno l'AIC o il BIC. È importante notare che questi criteri di informazione tendono a non essere buone guide per selezionare l'ordine appropriato di differenziazione ( $d$ ) di un modello, ma solo per selezionare i valori di  $p$  e  $q$ . Questo perché il differenziamento cambia i dati su cui viene calcolata la verosimiglianza, rendendo i valori AIC tra modelli con diversi ordini di differenziazione non comparabili. Per quanto descritto, la valutazione che ne segue è che il modello ARMAX risulta essere migliore rispetto al modello ARIMA semplice. La deduzione nasce dall'osservazione di valori più bassi di errore e dal valore dell'AIC. Quest'ultima fa inoltre denotare un miglioramento del modello dopo l'inserimento delle variabili esogene.

## 8. Decomposizione e ricomposizione della serie con MSTL

In questa fase si è implementato un modello di decomposizione multistagionale chiamato MSTL, in cui le componenti stagionali sono stimate in modo iterativo utilizzando STL. In generale, STL è un metodo versatile e robusto per decomporre le serie temporali. STL è l'acronimo di "Seasonal and Trend decomposition using Loess", dove Loess è una funzione che aiuta a stimare relazioni non lineari. Questo metodo gestisce qualsiasi tipo di stagionalità: la componente può cambiare nel tempo e il tasso di cambiamento può essere controllato dall'utente, come anche la smoothness.

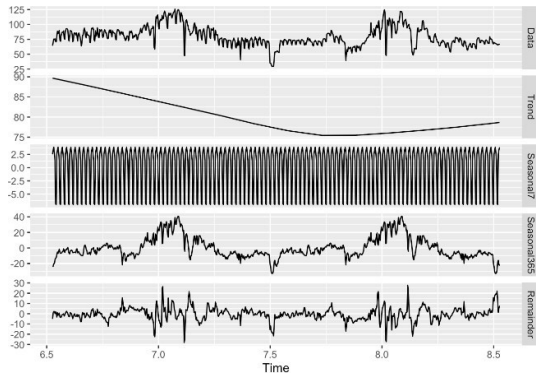


Figura 12: Decomposizione multistagionale U1 18-19

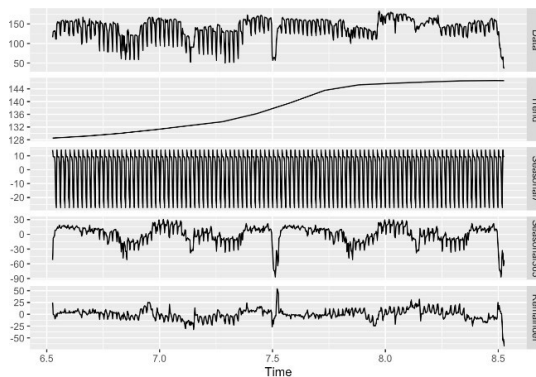


Figura 13: Decomposizione multistagionale U6 18-19

Si è deciso di decomporre la serie storica di U1 e U6 utilizzando come stagionalità 365 giorni e 7 giorni come nel grafico in Figura 12 e Figura 13. Si estrae poi la stagionalità 356 e si procede standardizzando l'effetto della stagionalità sui consumi pro capite, dividendo per 1300 persone (capienza massima di U1). Si è deciso di applicare questa tecnica su dati pre-Covid prendendo solo i primi due anni di entrambe le serie U1 e U6. A questo punto si estraggono i periodi stagionali che vanno da maggio ad ottobre, periodi nei quali sono impiegati gli impianti di raffreddamento, in modo tale da avere dati che possono intercettare la sostanziale differenza tra i due edifici e quindi, più nello specifico, il meccanismo di teleraffrescamento mancante in U1. In seguito, si innesta la stagionalità U1 da maggio ad ottobre nella serie di U6, moltiplicando dunque il dato di consumo pro capite per il numero di persone a massima capienza, ovvero 5311. A questo punto si ottiene una nuova stagionalità di U6 che viene ricomposta con questa variazione estiva, la Figura 14 riporta graficamente il risultato ottenuto.

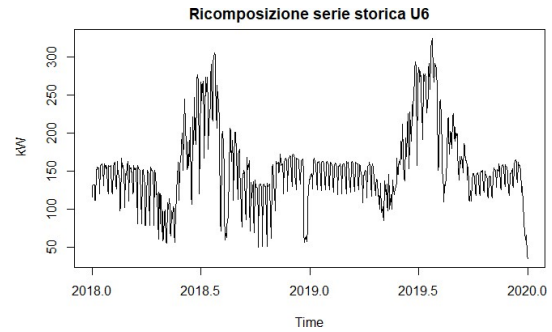


Figura 14: Serie storica U6 trasformata

Avendo un dato espresso in kilowatt per U6 e considerando la media giornaliera del consumo attivo, è possibile di fatto effettuare un confronto con il consumo della serie originale: 2429663.761 kWh per U6, mentre il consumo ipotetico con la stagionalità di U1 innestata e dunque senza meccanismi di termoregolizzazione sarebbe di 2714987.257 kWh. Il costo medio considerato fa riferimento ad una bolletta con 6 kW di potenza massima erogabile cioè 0.51 €, la Bicocca ha sicuramente un PME maggiore e quindi un costo maggiore. Calcolando il costo della differenza di queste due quantità, valutando al massimo del costo medio della bolletta e considerando costi con accise e iva si può calcolare un costo risparmiato di circa 136955 € in due anni, quindi 68477 € cioè l'equivalente di circa 23 rette annuali di fascia massima ogni anno.

## 9. Modello ARIMA U1 e confronto

Per poter stimare un modello che possa essere usato come paragone dell'evento non-Covid, si decide applicare un modello ARIMA prevedendo un periodo estremamente lungo, con il solo fine di avere una chiave di lettura molto approssimativa dell'evento e considerare un ipotetico confronto.

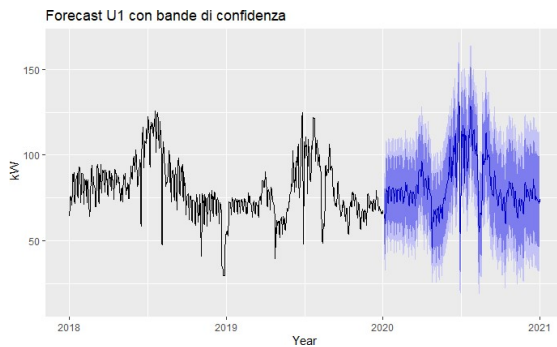


Figura 15: Forecast U1 con bande di confidenza 80% - 95%

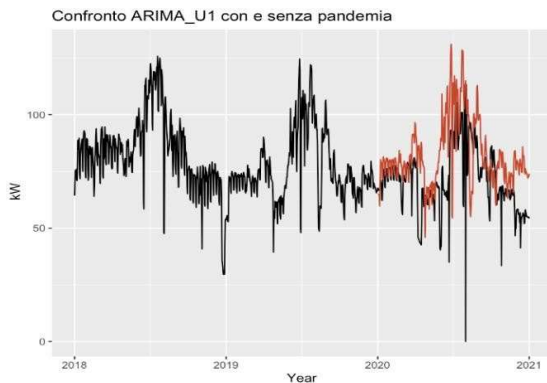


Figura 16: Covid vs. Non Covid U1 ARIMA

Per non trarre in inganno il fruitore si mostra prima in *Figura 15* la previsione con bande di confidenza con intervalli dell'80% e 95% infine il confronto tra il dato previsto non covid e il dato covid in *Figura 16*. Il grafico e la previsione sono stati effettuati con tutti i dati giornalieri fino a febbraio 2020 di modo da allenare il modello sul periodo pre-pandemico e fare una previsione molto "azzardata" sul periodo restante. Come prevedibile, mostrato a livello di banda di confidenza multilivello, la previsione strutturata in questo modo lascia molto a desiderare. Però illustra uno scenario particolarmente interessante su come i livelli di consumo nel 2020 sarebbero potuti tornare estremamente elevati come nel 2018, per poi scontare il caldo della bella stagione insieme all'aumento di studenti e del corpo docenti (trend consolidato in Bicocca da tempo).

## 10. Modello NEURALE U6 e confronto

Le reti neurali artificiali sono metodi di previsione che si basano su modelli matematici, i quali cercano molto lontanamente di replicare meccanismi simili a quelli del cervello umano,

come la funzione del neurone e il passaggio dell'informazione. Essi permettono lo sviluppo di relazioni non lineari complesse tra la variabile di risposta e i suoi predittori.

Nello specifico il modello che si adopera è il cosiddetto NNAR, ovvero autoregressione a rete neurale, la cui base teorica è la seguente: con i dati delle serie temporali, i valori ritardati possono essere usati come input per una rete neurale, proprio come con un modello di autoregressione. Il modello implementato è una rete feed-forward con uno strato nascosto. Per fare un esempio, un modello NNAR(9,5) è una rete neurale con cinque neuroni nello strato nascosto che utilizza le ultime nove osservazioni come input per la previsione dell'output. In presenza di dati stagionali è utile aggiungere anche gli ultimi valori osservati della stessa stagione come input. Gli autori adopereranno il modello tramite funzione `nnetar()`, il quale adatta un modello  $NNAR(p,P,k)_m$  in cui i parametri vengono generati in maniera autonoma. Un modello  $NNAR(p,P,0)_m$  è equivalente ad un modello  $ARIMA(p,0,0)(P,0,0)_m$  ma senza le restrizioni sui parametri che garantiscono la stazionarietà. In estrema sintesi, in un ambito di un obiettivo volto alle previsioni, la rete viene applicata iterativamente per fornire un valore il più accurato possibile. Per la previsione al tempo  $t + 1$  si usano semplicemente gli input storici disponibili, per le previsioni successive si adoperano le previsioni stimate fino a quel momento insieme ai dati storici.

Il modello implementato per U6, mostrato in *Figura 17*, è un NNAR(27,1,14) ovvero 27 osservazioni usate come input e, 14 neuroni nascosti e P pari a 1 in quanto è una serie stagionale. Il modello mostra il seguente output: si è ottenuta una media di 20 reti neurali, ognuna delle quali è una rete 28-14-1 con 421 pesi. Grafico e previsione sono stati effettuati con i dati fino a febbraio 2020, di modo che la linea rossa possa rappresentare la situazione ipotetica senza pandemia e quella nera il dato vero. L'algoritmo consiste in una media dei risultati di un numero finito di reti addestrate e per via della casualità dei pesi di ciascuna rete, il modello presenterà output differenti. Per tal motivo si è deciso di impostare

un seme che permetta di poter replicare l'output ottenuto.

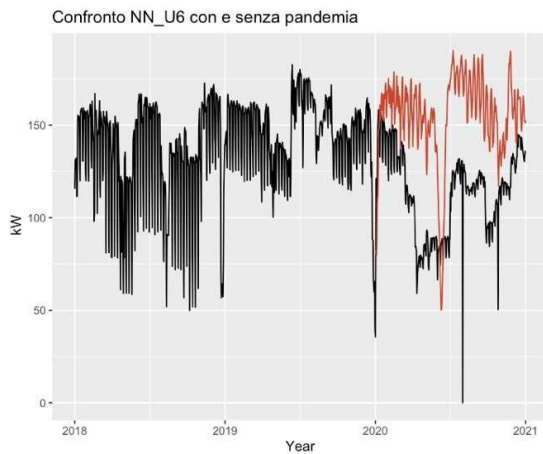


Figura 17: Covid vs. Non Covid U6 Neural Network

Questo metodo è poco accurato per motivi di orizzonte temporale. L'aspetto che è bene cogliere è quanto sarebbe stato effettivamente il consumo ipotetico in una situazione non pandemica.

Ai fini del confronto tra i due edifici si è poi deciso di riapplicare il metodo neurale anche ad U1, ottenendo (sempre con le dovute limitazioni del caso) la seguente previsione in Figura 18.

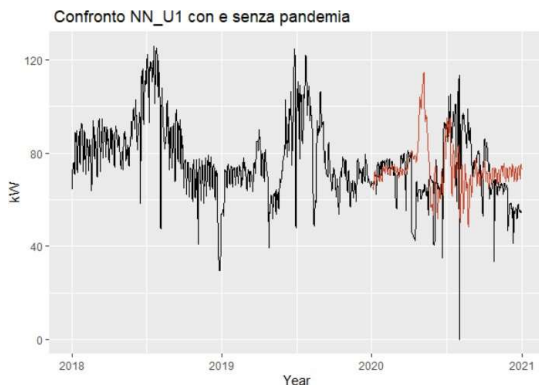


Figura 18: Covid vs. Non Covid U1 Neural Network

Si tratterà nel prossimo capitolo del risultato ottenuto dai confronti.

## Risultati

### 1. Integrazione dei missing values di U6

Il primo risultato notevole è proprio il modo in cui si è sopperito alla mancanza di dati di U6 per il mese di giugno. Inizialmente è stato fatto un tentativo di "rolling" a un mese cercando di prevedere i giorni successivi come una media ponderata di 10 predizioni. Valutando il tentativo su dati già in possesso, si è dimostrato molto scarso. Il nuovo metodo implementato tramite una sliding window è un approccio più efficiente che con buona confidenza appare discretamente indicato per predire un numero di giorni limitato, nel caso specifico 30, utilizzando i dati del problema. Ciò che si ottiene, ben intuibile dal codice allegato, è che i tentativi che utilizzano finestre temporali piccole sono, oltre che dispendiosi di tempo, anche estremamente fallaci. Il risultato che viene integrato nel dato originale è stato di fatto ottenuto con una finestra temporale pressoché completa di tutti i gironi antecedenti il mese da prevedere. Quindi il metodo avrebbe forse dovuto scontare una migliore comprensione della componente di rumore bianco.

### 2. Scomposizione e ricomposizione di U1 e U6

Ci sono diverse considerazioni da fare alla luce di quanto ottenuto. Infatti, nei precedenti capitoli si è stimato che la proporzione volumetrica di U6 è pari a circa 4.7 quella di U1, mentre il rapporto delle capienze massime è di circa 4.1, molto in linea con quanto visto a livello volumetrico. Dunque, si possono stimare alcuni valori di risparmio per U1 nell'ipotesi in cui si avesse l'opportunità di installare un sistema di teleraffrescamento, cifre che si aggirano intorno ai 15000€ - 17000€ l'anno (la stima è stata ottenuta dividendo il risparmio di U6 per le due proporzioni). La stima è molto ribassista perché calcolata a prezzo di mercato per cittadini privati, non per istituzioni pubbliche che hanno impianti enormi e servizi a costi meno calmierati rispetto a quelli per privati. Ci sarebbe anche da valutare l'IVA che nel caso di istituzioni pubbliche andrebbe a regime di split payment.



Questo dato palesa come sarebbe possibile risparmiare in termini di consumo uniformando l'infrastruttura di raffreddamento dei due edifici. Per ovvi motivi gli autori non hanno contezza di quanto possa essere il costo di ammodernamento degli edifici e si rendono conto di quanto questo tipo di conclusione possa essere forzata.

## 2. Regressione e confronto tra ARIMA e ARMAX

Ciò che otteniamo è un modello che si adatta meglio del modello privo di variabili esogene, nello specifico: il calcolo previsionale con variabili esterne alla serie storica in analisi è un metodo che permette di migliorare significativamente il modello. Infatti, i risultati ottenuti tramite la regressione multipla si possono interpretare come un monito al fatto che l'edificio potrebbe adattare meglio la propria impiantistica e i propri consumi per essere meno incline ai cambiamenti di utilizzo dell'edificio. Al solito questa deduzione potrebbe essere del tutto errata in mancanza di ulteriori informazioni.

## 3. Confronto con dati Covid e non Covid

Si è illustrato come il calo dei consumi dei due edifici non sia stato il medesimo e quindi come l'edificio U6 abbia reagito meglio alla pandemia al netto anche delle informazioni che non si possiedono. Facendo una stima del tutto indicativa e priva di pretese, la quale è stata presentata negli ultimi due capitoli dei paragrafi precedenti, si è potuto calcolare le differenze di consumi ottenute tra la fase Covid e la fase non-Covid. Ebbene, i numeri calcolati usando le reti neurali palesano un risparmio complessivo calcolato per U6 di circa 283462 € e per U1 di 59554 €. Le previsioni indicano ancora, quasi come fosse una sorta di "Canone di Policeto", la proporzione di 4.76 ad indicare il rapporto tra le due grandezze previste. Dunque, l'università degli studi Milano Bicocca ha certamente avuto un risparmio notevole sui costi energetici e il risparmio si è figurato proporzionale a quanto sono le grandezze volumetriche.

## 4. Proposta energy manager

Le analisi temporali indicano quindi come l'ammodernamento pre-pandemico degli edifici avrebbe potuto comportare un risparmio monetario in termini di consumo anche alla luce di quanto visto con i dati metereologici. Lo stesso fatto che la variabile Covid e la variabile dummy inerente all'effettiva apertura/chiusura degli edifici siano più impattanti per U6 e meno per U1, porta all'evidenza di un edificio che reagisce meglio alla temperatura esterna ed un altro che invece ne risente poco, banale ma chiaro segnale di spreco energetico o possibilità di miglioramento della gestione dei consumi.

## Conclusione

La conclusione, che per sua natura comporta asserire una qualche evidenza empirica, in realtà sarebbe più onestamente sintetizzabile con una domanda: quanta ricchezza (anche in termini di non inquinamento ambientale) si sarebbe potuta generare dal risparmio nei consumi e quanto "cashback" sarebbe stato possibile reinvestire nell'università? L'analisi che viene riportata in questa sede avrebbe comportato come risposta: probabilmente molta! Avere un impianto più reattivo al cambiamento atmosferico, alla decrescita di utilizzo e con un tipo di tecnologia di mantenimento della temperatura come i così detti "cappotti termici", avrebbe certamente permesso all'università di poter stanziare qualche borsa di studio in più oppure scontare il risparmio direttamente alla retta degli studenti, cosa che non è stata fatta. In secondo luogo, distribuire il personale e gli studenti sfruttando gli edifici a minore impatto di consumo (come U6) e limitare il più possibile l'accesso agli edifici ad alto impatto (come U1), avrebbe permesso un ulteriore risparmio anche in fase pandemica.

## Possibili sviluppi migliorativi

Gli sviluppi futuri potrebbero essere molteplici: in primis integrare i dati con maggior successo rispetto a quanto implementato finora, riuscendo a reperire più informazioni amministrative e tecniche oppure coinvolgendo esperti di dominio energetico, ambientale e architettonico. In secondo luogo, avere un dato più preciso e



granulare in termini spaziali. I due edifici hanno solo due centraline, una ciascuno, quindi un buon obiettivo sarebbe quello di capire veramente dove sia localizzato lo spreco ed evitare di generalizzare con il rischio di sbagliare palesemente.

Gli autori si ritengono comunque soddisfatti di quanto si è cercato di implementare e, al di là delle evidenze talvolta da prendere con la dovuta parsimonia, uno dei riscontri effettivi è il seguente: l'analisi di dati complessi, senza la consulenza di un esperto di dominio e con strumenti informativi limitati, è un compito arduo che richiede anni se non decenni di esperienza in vari domini e richiede inoltre una mentalità comprensiva ed elastica. Sarebbe certamente interessante poter avere un dialogo con qualche professionista dell'energia che possa o meno confermare la tesi che trapela da questo lavoro.

Come recita la famosa euristica attribuita al danese Niels Bohr, premio Nobel per la Fisica nel 1922: "È difficile fare previsioni, soprattutto sul futuro".

Grazie per il tempo dedicato alla lettura.

## Riferimenti bibliografici

### 1. Bibliografia

Fattore M. (2020) UniMib. Fundamentals of time series analysis, for the working data scientist (DRAFT)

Ignoto. lavori\_u6\_-\_inail.pdf disponibile su [https://www.unimib.it/sites/default/files/allegati/lavori\\_u6\\_-\\_inail.pdf](https://www.unimib.it/sites/default/files/allegati/lavori_u6_-_inail.pdf)

Iannantuoni G. (2020) UniMib. GESTIRE L'EMERGENZA PER RIPARTIRE INSIEME disponibile su <https://www.unimib.it/coronavirus-informazioni-e-servizi>

Ricci V. (2015) UniBa. Analisi Delle Serie Storiche con R.

### 2. Sitografia

I seguenti siti sono stati consultati a più riprese durante i mesi di maggio e giugno 2021.

<https://www.unimib.it/ateneo/energy-management>

<https://otexts.com/fpp2/>

<https://nwfsc-timeseries.github.io/atsa-labs/index.html>

<https://meteostat.net/en>

<https://www.terna.it/it/media/comunicati-stampa/dettaglio/Terna--con-ora-legale-benefici-elettrici-ambientali-ed-economici-per-il-sistema-energetico-italiano#:~:text=Dal%202004%20al%202020%20Terna,e%20720%20milioni%20di%20euro.>

[https://www.sociologia.unimib.it/sites/st14/files/Allegati/Dipartimento\\_Sociologia\\_e\\_Ricerca\\_Sociale\\_Calendario\\_didattico\\_AA\\_2018\\_2019.pdf](https://www.sociologia.unimib.it/sites/st14/files/Allegati/Dipartimento_Sociologia_e_Ricerca_Sociale_Calendario_didattico_AA_2018_2019.pdf)

<https://www.sociologia.unimib.it/sites/st14/files/Allegati/Calendario%20didattico%20a.a.%202019-2020.pdf>

[https://www.unimib.it/sites/default/files/Allegati/Calendario\\_Accademico\\_20-21.pdf](https://www.unimib.it/sites/default/files/Allegati/Calendario_Accademico_20-21.pdf)

<https://www.unimib.it/amministrazione-trasparente/altri-contenuti/ateneo-cifre/dati-sulle-infrastrutture>

<https://taqialabolletta.it/quanto-costa-un-kwh/>

### 3. Link pubblico al codice GitHub

[Open Link GitHub](#)