Problem and Goal

The company Terna S.p.A. is the transmission system operator (TSO) of the Italian power grid. Its task is to guarantee the right balance between the injections and withdrawals of the power grid. To accomplish its duty, Terna needs to buy services from the operators of the energy market. However, starting from the analysis of the available data, the problem is so complex, that it is often impossible to understand the choices operated by Terna on the offers' selections. In fact, it is not unusual to see significant fluctuations on the profits of the companies operating in the electrical sector. This is a relevant company risk. In the view of the foregoing, the scope of this research project is to forecast on the time series of energy's unbalances between the injections and withdrawals of the power grid, in order to put Terna and the energy market operators under a minor risk. In particular, we use three methods to forecast:

1. A method, here called *simplex projection* (SP), invented in the nineties by George Suigihara and Robert M. May in the context of dynamic systems.

2. A K-nearest neighbors method (KNN).

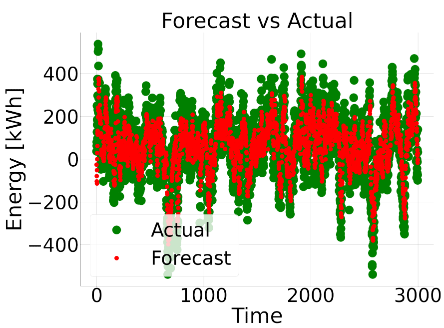
3. A recurrent neural network (RNN).

The tests on synthetic data

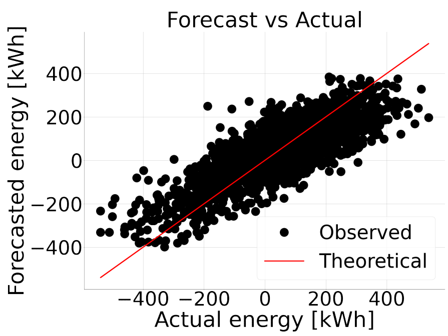
With the aim of evaluating the performances of the three methods, we tested them on three chaotic maps: the Logistic map, the Lorenz map and the Mackey-Glass map. The procedure consists in predicting a subset of the time series and measuring the accuracy of the predictions using the correlation coefficient (ρ) between the actual and predicted time series. For a chaotic time series the accuracy of the nonlinear forecast falls off when the prediction-time interval (T_P) increases at a rate which gives an estimate of the Lyapunov exponent. We did not find out whether one method of prediction is more effective than the others, however, we understood that the simplex-projection is only a more time-consuming version of the KNN and they display similar accuracy for every dataset tested.

The analysis of the energy's unbalances predictions

Terna has been collecting old time series of energy's unbalances every 15 minutes, and the values are labelled differently whether the energy consumed is on the northern or on the southern part of Italy.



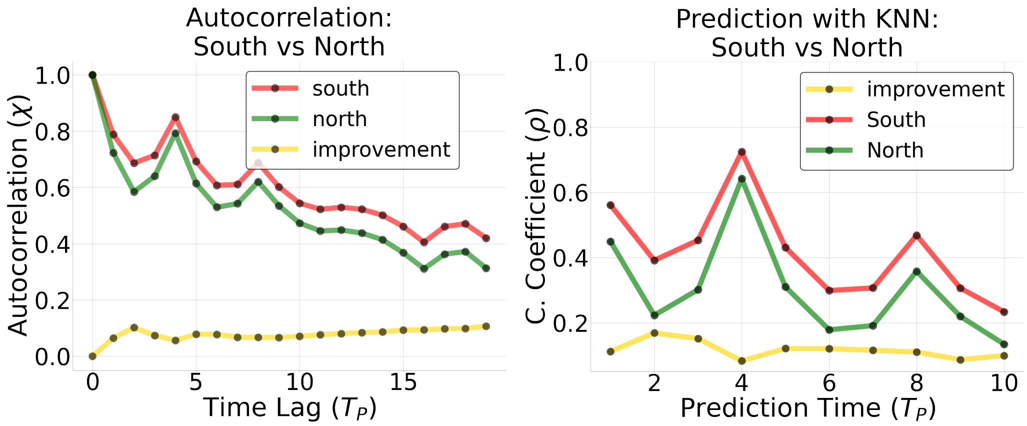
(a)



(b)

Figure 1: (a) This figure displays the predicted energy's unbalances and the observed ones in north Italy as function of time. (b) This figure shows on the y-axis the predictions of energy's unbalances and on the x-axis the actual energy unbalances of north Italy. In both figure the predictions are calculated 4 time steps in the future ($T_P = 4$) using a recurrent neural network.

The accuracy of the energy's unbalances' predictions shows two maximums at $T_P = 4, 8$ (indeed 1 and 2 hours in the future) and, while the goodness of the predictions made by the recurrent neural network are rather similar whether made in north or in south Italy, we saw that the other two methods display better predictions on the south. We suppose that a cause of this phenomenon might be that the time series of south unbalances is more autocorrelated.



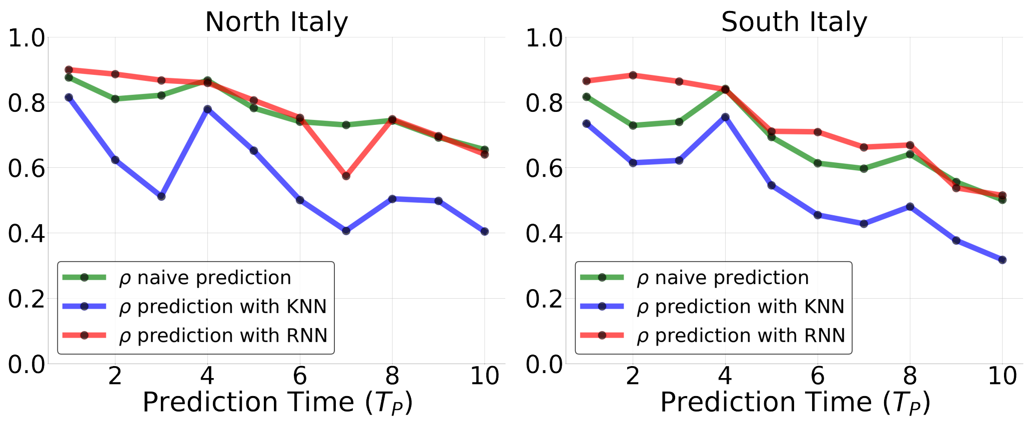
(a) (b)

Figure 2: (a) This figure shows the autocorrelations of the energy's unbalances time series in north and south Italy. In yellow is shown the difference between the autocorrelations of the two parts of Italy. (b) On the right the correlation coefficient between the actual and predicted time series of the energy's unbalances as T_P varies (using the method of KNN). In yellow is shown the difference between the correlations of the two parts of Italy

To understand if the results obtained are satisfying, we implemented a naive method of prediction. If the naive method's predictions are more trustworthy, we will conclude that our methods are not effective. Given a time series $x(t)$, the naive method consists in predicting the time series at time $x(t+T_P)$ with the value $x(t)$.

$$x(t+T_P) \longrightarrow x(t)$$

This is the most reasonable method we can implement when our time series is a random walk and we do not know anything abut the future steps, which, by definition are completely random. This is exactly what we did on the energy's unbalances time series. We saw that the accuracy displayed by the KNN and the SP changes accordingly to the dataset used. For some dataset the KNN and SP's accuracy is dominated by the naive method's one, for other dataset it is the naive method's accuracy to be dominated by the other two's. On the other hand, the predictions made using the RNN were more stable when we changed the dataset, and they were more accurate than the predictions of the naive method. The figure 3 shows the prediction's accuracy as function of the prediction-time interval (T_P). The predictions' accuracy of the method of Suigihara and May is not shown because it is very similar to the KNN method's one.



(a) (b)

Figure 3: In these figures are compared the correlation coefficients between the actual and predicted time series calculated with the algorithm of KNN, with a RNN and with the naive method. The analysis shows that the KNN's accuracy is always dominated by the one of the other two methods. For almost every T_P the RNN displays the best accuracy among all the methods tested. Moreover, unlike the K-nearest neighbors and simplex projection, the RNN's accuracy does not follow the same trend of the autocorrelation of the series: it is overall decreasing for every T_P.

Conclusions

- The KNN and SP's methods are not useful to predict the energy's unbalances.
- The RNN could be useful to predict the energy's unbalances.