2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)

Covid-19 Signal Analysis: Effect of Lockdown and Unlockdowns on Normalized Entropy in Italy

Francesco Benedetto
SP4TE laboratory
Roma Tre University
Rome, Italy
francesco.benedetto@uniroma3.it

Gaetano Giunta
SP4TE laboratory
Roma Tre University
Rome, Italy
gaetano.giunta@uniroma3.it

Chiara Losquadro
SP4TE laboratory
Roma Tre University
Rome, Italy
chiara.losquadro@uniroma3.it

Luca Pallotta
SP4TE laboratory
Roma Tre University
Rome, Italy
luca.pallotta@uniroma3.it

2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 978-1-7281-6215-7/20/$31.00 ©2020 IEEE DOI: 10.1109/BIBM49941.2020.9313396

Abstract—**Entropy concept is related to uncertainty and predictability of random time series. The estimated trend of such a parameter can provide useful information and possibly predict future behavior of a number of non-stationary noisy signals. The goal of this paper consists of analyzing the Covid-19 signal made by the number of registered infections in Italy during the first four months of the pandemic epidemy (March-June 2020). Finally, some considerations are drawn after matching historical dates of some Covid-19 related Acts made by the Italian Government (i.e., lockdown and unlockdowns). Based on the obtained results, we could conjecture that the provisions have inducted people to a common behavior concerning local mobility during the lockdowns and the progressive unlockdowns of the quarantine period in Italy.**

Keywords — Covid-19, Medical Signal Analysis, Time Series Analysis, Entropy.

I. Introduction

It is well known since the publication of the pivotal work of the mathematician C. Shannon "A Mathematical theory of Communications" in 1948 [1], that the concept of entropy is related to that of uncertainty. High values of the Shannon entropy result in unpredictable series, while lower values mean less uncertainty and hence a more predictable behavior of that series. This concept has been also applied to non-stationary signals as well. In practice, observing the trend of entropy versus the time, the level of entropy corresponds to "the predictivity of signals" [2]. Namely, the entropy value becomes smaller, the predictivity becomes higher on the entropy curves by time. In [3], an empirical method for evaluating the entropy of a financial series is proposed, namely the approximate entropy (ApEn). ApEn is able to obtain the entropy estimation by modifying an exact regularity statistic, namely the maximum entropy method (or Kolmogorov–Sinai entropy). In particular, the authors use the approximate entropy technique as a marker of market stability, with rapid increases possibly foreshadowing significant changes in a financial variable.

However, all the aforementioned works evaluate the entropy of historical data and, applying ex-post considerations, try to declare the predictability of the series, i.e., they implicitly assume that the series under investigation are characterized by a stationary behavior. This means that they suppose that the past statistical features of the analyzed series remain unaltered also in the future.

A novel signal processing method for the analysis of financial and commodity price time series was introduced to assess the predictability of financial markets [4]-[5]. Unlike a number of ex-post analyses, that technique is able to predict the entropy of the next future time interval of the time series under investigation by a least square minimization approach exploiting the maximum entropy method.

The goal of this paper consists of applying such technique, well suited for noisy non-stationary time series, to the analysis of Covid-19 time series of the number of registered infections in Italy during the first four months of the pandemic epidemy (March-June 2020). Moreover, based on the obtained results of the carried analysis, some considerations are finally drawn after matching historical dates of major Acts made by the Italian Government.

This conference paper is organized as follows. Work's material and methods are presented in sec. II, including a brief history of Covid-19 and acquisition of infection data in Italy, as well as a summary on the application of entropy estimates as a measure of information and randomicity of data under investigation. In sec. III, the numerical results of the carried analysis are reported and also discussed, matching historical dates of major provisions and the entropy trends of infections in Italian regions.

II. Material and Methods

A. Historical framework of Covid-19 epidemy in Italy

On December 31st, 2019, the Health Commission of Wuhan (China) informed the World Health Organization (WHO) about a cluster of unknown pneumonia cases in the province, later identified as the SARS-CoV-2. On 30 January 2020, as a result of this notification, the WHO declared a state of emergency.

On January 31st, 2020, the Italian Government, taking into account the epidemic rate of spread of the SARS-CoV-2, claimed the state of emergency in Italy and introduced the first preventive and containment measures to avoid the spread of the disease through the whole Italian country. On February 21st, 2020, a cluster of SARS-CoV-2 cases was identified in northern Italy; it forced the Italian Prime Minister to declare the state of quarantine on March 8th, 2020, first to Lombardy and 14 further provinces in the North of Italy, and then, after a few days, in all the Country.

The quarantine was a measure aimed at dealing with the COVID-19 emergency, known as Phase 1; this phase started on March 8th, 2020 and ended on May 3rd, 2020, enforcing Italy to a strict lockdown condition. This condition has forbidden any movement of Italian citizens, that were closely monitored. In particular, it has forced to stop all the public and private offices (except essential services, such as hospitals, food chain enterprises, and police), as well as all commercial activities were closed (except pharmacies and food shops).

Once the stabilization of the spread of COVID-19 occurred, the Phase 2 started on May 4th, 2020. The Phase 2 ended on June 14th, 2020 and it has been followed by the Phase 3 on June 15th, 2020. With the announcement of Phase 2 and Phase 3, Italian people went out progressively from quarantine (called ‘unlockdown’ by a recently suggested neologism [6]), while citizens’ mobility and commercial activities were gradually restored.

B. Dataset of infections in Italy

The Covid-19 data were numerically collected from the Italian ‘Dipartimento della Protezione Civile’ (Civil Protection department of Prime Ministry) that published daily updated statistics on the free public internet site of the Italian Ministry of Health [7], that presents the infection’s data of the Italian regions for each day from the beginning of the epidemy till now. For better understanding of timing of epidemic diffusion, the daily number of infections in Italy during the period March-June 2020 is reported in Figure 1.

The total number of infections (for the whole Italy and for each Italian region) versus day were processed by applying the same algorithm presented in [4], where the whole prediction estimator of entropy is diffusely detailed. That procedure has been here applied with a shorter timeline, different from the original one, tuned on financial time series. In fact, the estimates of entropy have been computed day by day within a sliding window of one week by using an autoregressive predictive model of signal autocorrelation made on three weeks of data sequence (unlike [4] where a window of one year and a model prediction of three years were set).

The goal of the analyses has been to evidence the daily relative trend of entropy for each Italian region. As a consequence, similarly to [4], data from all the region streams were normalized in terms of signal’s energy in order to evidence the day trend of entropy, regardless its absolute value to avoid the dependence on region’s size in terms of population. This approach has allowed to possibly correlate infection data with the effect of quarantine Acts, that can be regarded as impulsive events affecting the signal.

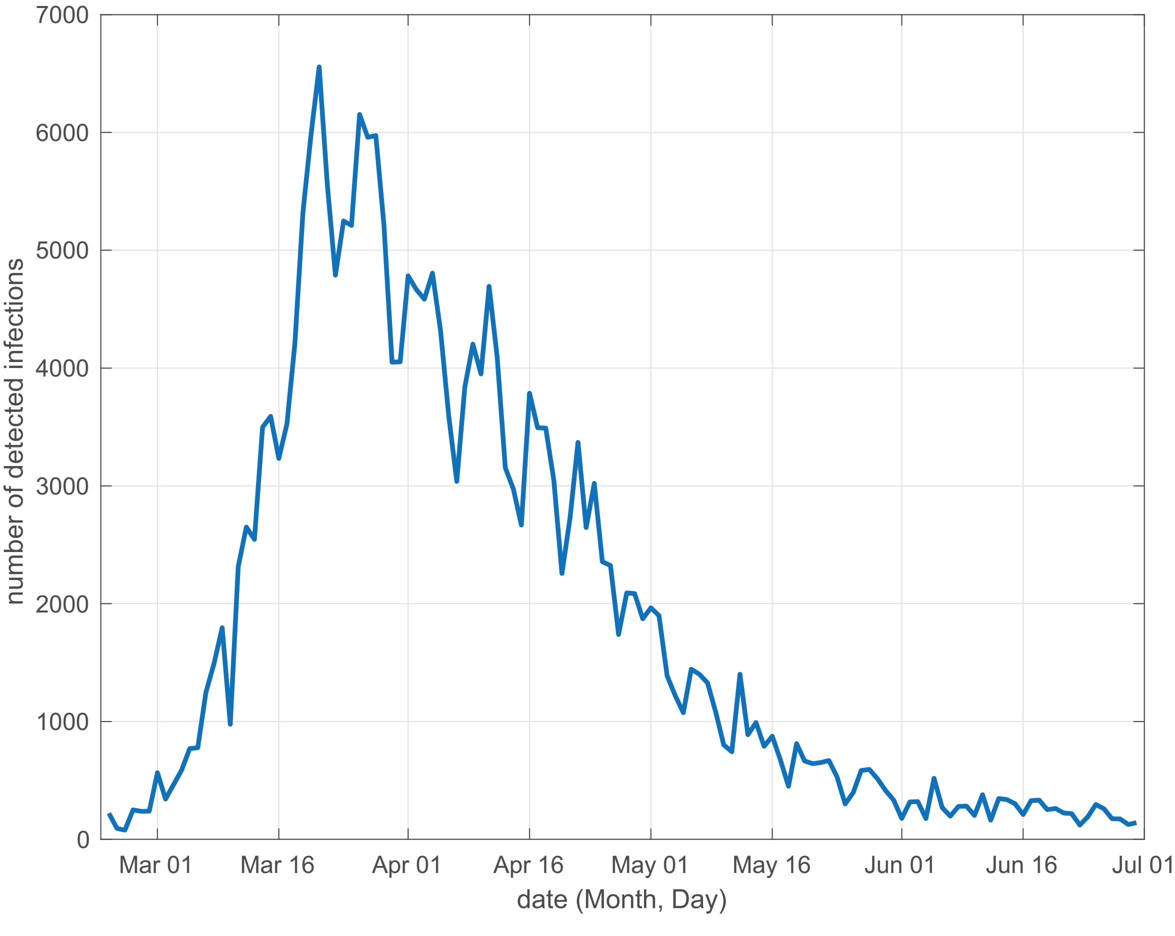

Figure 1. Daily number of detected infections in Italy (data from [7]).

C. Entropy estimate as a measure of series information

Entropy concept is usually related to infinite data series, corresponding to an infinitely accurate precision and resolution for entropy evaluation. However, practical data are finite time series data characterized by limited resolution. From Burg's approach, the correlation in a time series, related to the contents of information (entropy) can be estimated even when only a small number of data are available [8]-[9].

This is called the maximum entropy method (MEM). In addition, authors of [10]-[11] showed that MEM is equivalent to the least-squares method for fitting an autoregressive (AR) model (or all-pole model) to the given data.

MEM relates the entropy rate of a time series with its power spectral density (PSD). Hence, knowing the PSD of a series, allows us to know its entropy rate. A direct relationship holds for stationary Gaussian time series [12]-[13]. Nonetheless, one can consider the maximum entropy estimate as in [4] as a reasonable upper bound of entropy even when the series is neither Gaussian nor stationary.

The approximate entropy estimate H has the following expression [4]:

$$H = \frac{1}{2}\ln(2\pi\, e) + \frac{1}{4\pi}\int_{-\pi}^{\pi} \ln\left(2\pi\, PSD(\omega)\right) d\omega \quad (1)$$

Maximum entropy (1) happens for uncorrelated series [12]. Then, lower entropy values result in more predictable time series, while the value of entropy is not found to decrease for noise [14]. Hence, the entropy can be used as an indicator of the time series predictability.

Like in [4], the PSD was estimated by a predictive AR model fitted to the sample time auto-covariance of data. The entropy was scaled (normalized), assuming a unitary variance of the AR innovation process, for all the data sequence, regardless the amount of infections (that was very different between the Italian region), to evidence and match the entropy trends of each region.

III. Results and Discussion

The normalized entropy, estimated as above explained, consists of 22 plots (one for Italy and one for each Italian region, including Trent and South Tyrol autonomous provinces) versus time. In order to possibly correlate with the Prime Minister's Acts that were decided during the period, a detailed report of major provisions is presented in Table I.

For better reader's understanding, they are represented into six groups, viz. North-West, North-East, two for Central Italy, and two for South (including the two major islands), respectively in Figures 2-7. The plot referring to whole Italy is also included in all the six figures.

TABLE I
TIMELINE OF MAIN QUARANTINE ACTS IN ITALY FROM MARCH TO JUNE 2020.

Phase 1.

- *Phase 1.1: March 8th, 2020. In this phase, Lombardy and 14 other northern provinces were claimed as 'zone rosse' ('red zones') and forced them to a quarantine condition.*
- *Phase 1.2: March 9th, 2020. From this moment, also the rest of Italy was claimed as 'red zone'. The date had marked the prohibition of any movement of Italian citizens within the national territory, unless health, work or emergency reasons occurred. During this period, the stop of all commercial activities was set, all the educational institutions were closed and a home working condition (called 'smart working' in the Italian provision) was highly recommended to reduce workers' mobility of private enterprises and public offices. The mobility was restricted to the purchase of primary commodities exclusively.*
- *Phase 1.3: March 22nd, 2020. Despite the previous limitations, too many citizens have continued to free move, forcing an additional restriction by Law. The new notification aimed at closing parks and public gardens and avoiding outdoor sports. In addition, further commercial activities considered as unnecessary were closed, such as restaurant and bar activities. Moreover, citizens' mobility was limited to the own residence's municipality. Even visiting relatives not cohabitants was forbitten. Due to this Act, the quarantine procedure was fully implemented and forced a large part of population to stay at their own's home.*

Phase 2.

- *Phase 2.1: 4th, May 2020. The Italian citizens got the permission to visit their own strict relatives ('congiunti' in Italian, that a specific FAQ of Ministry, to avoid misunderstandings, limited to parents, brothers, sisters, life partners, friends, and few more), allowing to visit them during daytime only within the same region.*
  In addition, parks and public gardens reopened and individual sports were allowed again with a proper inter-personal distance.
- *Phase 2.2: May 18th, 2020. From this moment, most of commercial activities restarted and it was officially allowed to visit further people, in addition to relatives, but always within the same region.*
- *Phase 2.3: June 1st, 2020. The remaining commercial activities and restaurants reopened.*
- *Phase 2.4: June 3rd, 2020. From this moment, the interregional mobility and many recreational activities restarted.*

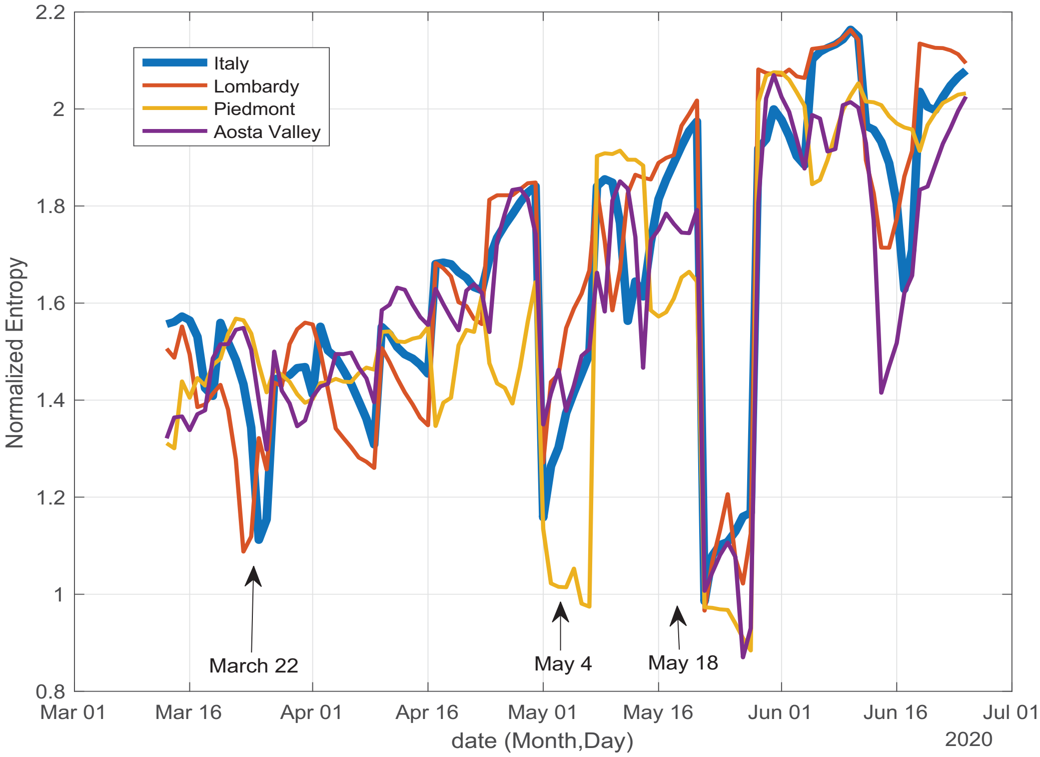

Figure 2. Normalized Entropy versus date for Italy and North-West regions.

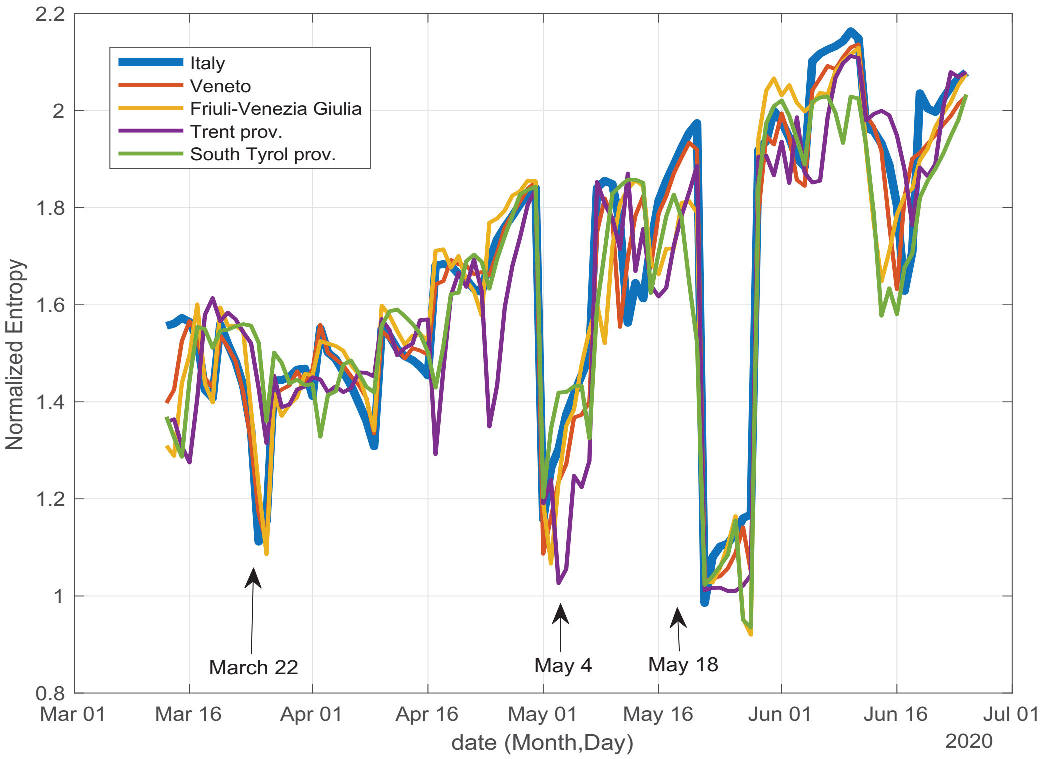

Figure 3. Normalized Entropy versus date for Italy and North-East regions.

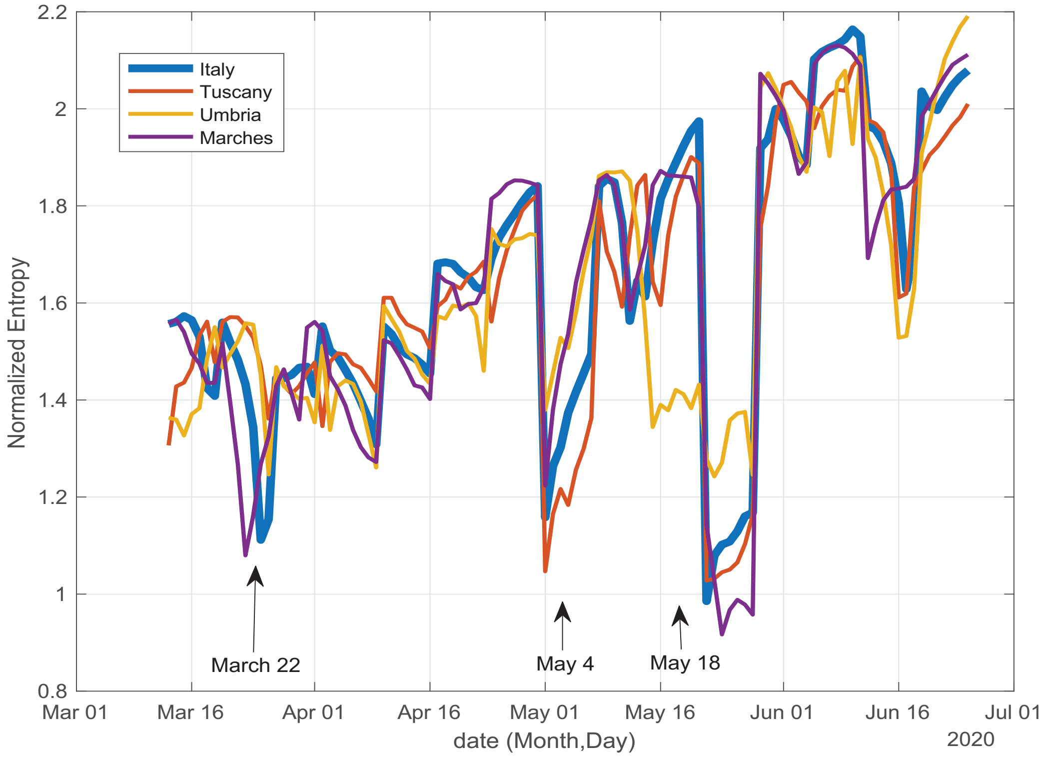

Figure 4. Normalized Entropy versus date for Italy and upper Central regions.

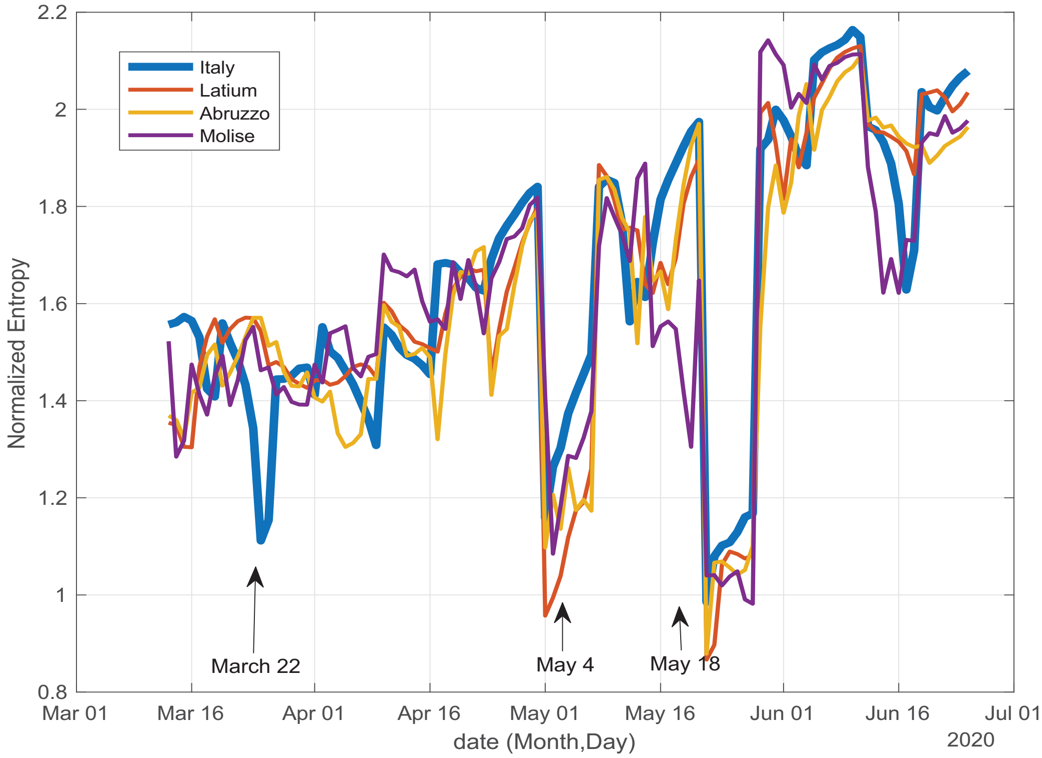

Figure 5. Normalized Entropy versus date for Italy and lower Central regions.

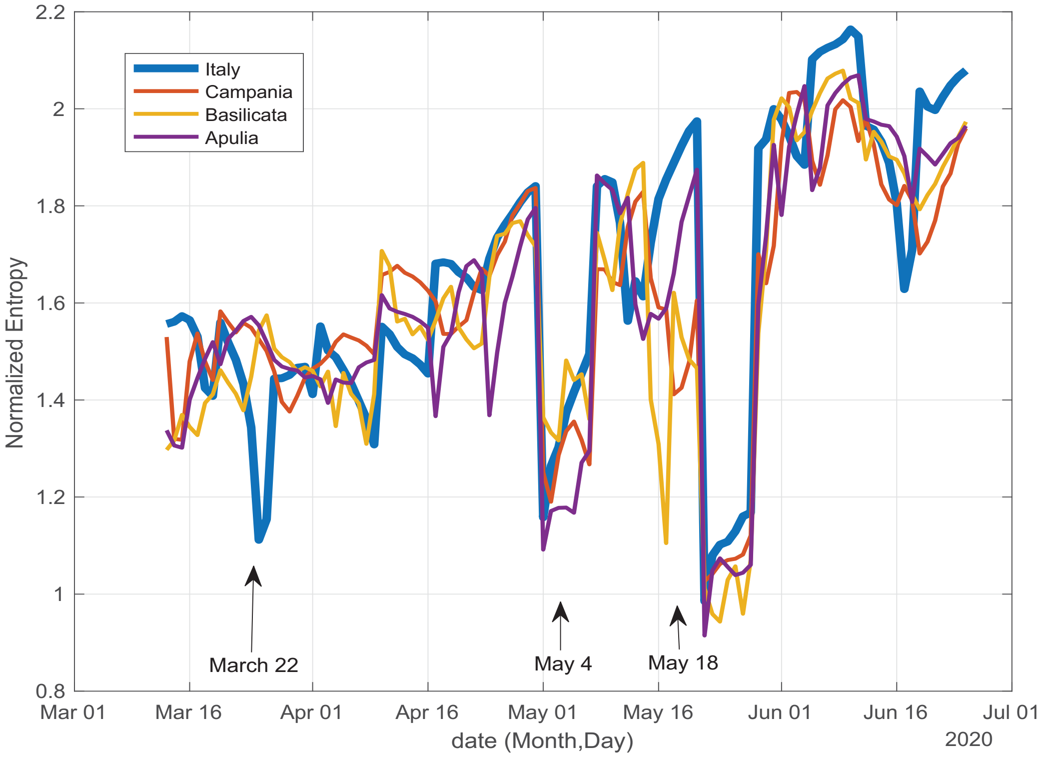

Figure 6. Normalized Entropy versus date for Italy and upper South regions.

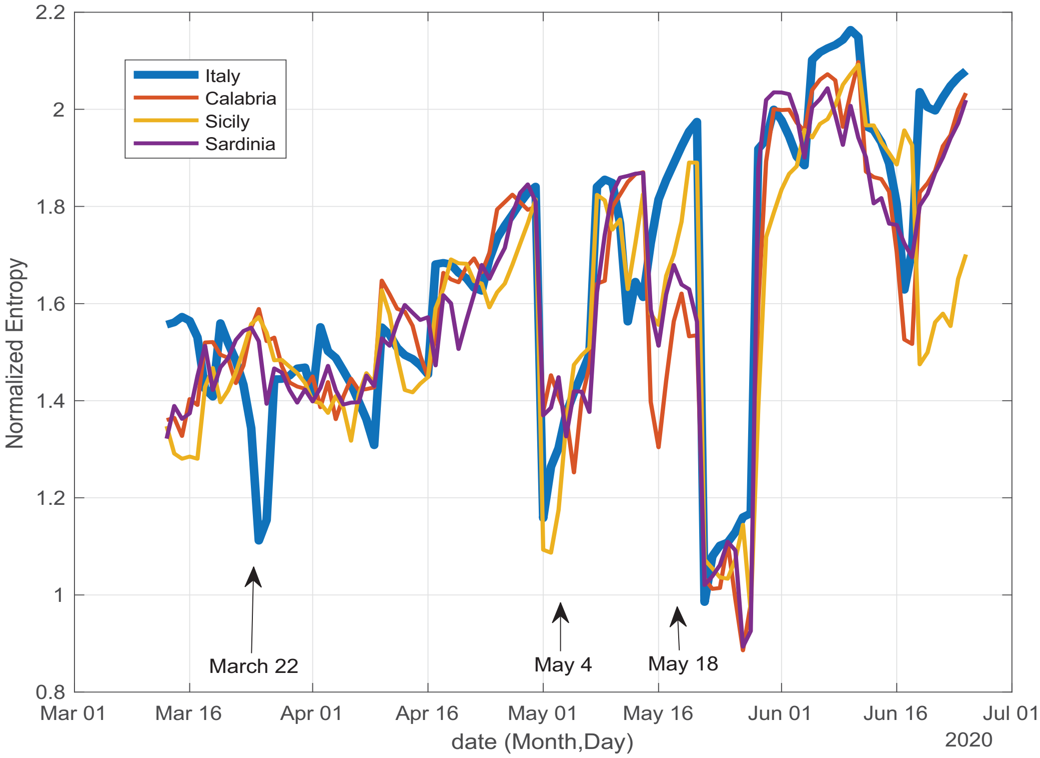

Figure 7. Normalized Entropy versus date for Italy and lower South (with major Islands) regions.

The plots of all regions look very similar along the time. In fact, the correlation coefficients between each pair of regions were computed, resulting very high (mean: 0.88 with a standard deviation of 0.06).

The effects of some Acts seem to visually correlate with the plots, especially around March 22nd, May 4th, and after May 18th (see again Figs. 2-7). In these particular periods, the normalized entropy goes down for some days, then goes up by a sudden change of the trend.

There is no proved explanation of such a trend. Nevertheless, some hypothesis can be assessed in order. We can argue that the strict limitation of citizens' mobility of Phase 1.3 (strict lockdown at home), as well as the partial re-openings ('unlockdowns') of Phase 2.1 (allowed visits to parents and strict relatives) and 2.2 (allowed visits to friends in the same regions), have inducted to people a common behavior concerning local mobility, that has produced an increasing of predictability of infection data trend in all the regions, corresponding to a reduction of variability of such data.

After some days, we can suppose that people's mobility became more various and random, since people came back to move asynchronously each other. As a consequence, after a short drop, normalized entropy went back to standard values.

IV. Conclusion

This paper has presented an attempt of interpretation of the Covid-19 signal made by the number of registered infections in Italy during the first four months of the pandemic epidemy.

For such a purpose, the entropy concept, related to uncertainty and predictability of random time series, has been adopted. In fact, the estimated trend of such a parameter can be capable to provide useful information and possibly predict future behavior of a number of non-stationary noisy signals.

Based on the obtained results, some considerations are finally drawn after matching historical dates of major Acts made by the Italian Government, showing a singular correlation with people's mobility restrictions and releases.

References

[1] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27, 379–423, 623–656, 1948.

[2] Y. Takizawa, A. Fukasawa, Non-stationary signal spectrum analysis improving maximum entropy estimation error, in: Int. Conf. on Acoustics, Speech, and Signal Proc., ICASSP-89, vol. 4, pp. 2206–2209, 1989.

[3] S.M. Pincus, E.K. Kalman, Irregularity, volatility, risk, and financial market time series, Proc. Natl. Acad. Sci. 101 (38), 13709–13714, 2004.

[4] F. Benedetto, G. Giunta, L. Mastroeni, "A maximum entropy method to assess the predictability of financial and commodity prices," Digital Signal Processing, vol. 46, pp. 19-31, 2015.

[5] F. Benedetto, G. Giunta, L. Mastroeni, "On the predictability of energy commodity markets by an entropy-based computational method," Energy Economics, vol. 54, pp. 302-312, 2016.

[6] https://www.collinsdictionary.com/submission/22332/unlockdown

[7] http://www.salute.gov.it/portale/nuovocoronavirus/dettaglioContenutiNuovoCoronavirus.jsp?area=nuovoCoronavirus&id=5351&lingua=italiano&menu=vuoto

[8] J.P. Burg, Maximum entropy spectral analysis, in: Proc. 37th Ann. Meet. Soc. Explor. Geophys., 1967.

[9] J.P. Burg, Maximum entropy analysis, Ph.D. dissertation, Stanford Univ., Stan- ford, CA, 1975.

[10] Y. Inouye, Maximum entropy spectral estimation for regular time series of degenerate rank, IEEE Trans. Acoust. Speech Signal Process. 32 (4), 733–740, 1984.

[11] A. Van den Bos, Alternative interpretation of maximum entropy spectral analysis, IEEE Trans. Inf. Theory 17 (4), 493–494, 1971.

[12] B.S. Choi, T.M. Cover, An information-theoretic proof of Burg's maximum en- tropy spectrum, Proc. IEEE 72 (8), 1094–1096, 1984.

[13] B.S. Choi, On the relation between the maximum entropy probability density function and the autoregressive model, IEEE Trans. Acoust. Speech Signal Pro- cess. 34 (6), 1659–1661, 1986.

[14] Y. Takizawa, A. Fukasawa, Non-stationary signal spectrum analysis improving maximum entropy estimation error, in: Int. Conf. on Acoustics, Speech, and Signal Proc., ICASSP-89, vol. 4, pp. 2206–2209, 1989.