Results in Physics 26 (2021) 104391

Contents lists available at ScienceDirect

Results in Physics

journal homepage: www.elsevier.com/locate/rinp

Steady state Kalman filter design for cases and deaths prediction of Covid-19 in Greece

N. Assimakis [a], M. Adam [b], A. Ktena [a], C. Manasis [a,*]

[a] Energy Systems Laboratory, National and Kapodistrian University of Athens, Greece
[b] Dept. of Computer Science & Biomedical Informatics, University of Thessaly, Lamia, Greece

ARTICLE INFO

Keywords:
Prediction
Kalman filters
Steady state
Covid-19

ABSTRACT

In this work we study the applicability of the steady state Kalman filter in order to predict new cases and deaths of Covid-19. We use the actual observations of new cases and deaths. First, we deal with short term prediction, namely daily prediction. We propose the use of the golden steady state Kalman Filter, which is designed to have parameters related to the golden section. It was found that the proposed golden steady state Kalman Filter has a satisfactory behavior compared with the classical mean or average filter. Secondly, we deal with long term prediction, for example average prediction per quarantine period (14 days). We propose to process blocks of measurements of time window corresponding for example to the quarantine period in order to predict the average of cases and deaths using steady state Kalman Filter. It was found that the proposed golden steady state Kalman Filter produces more reliable predictions than the classical mean or average filter does. The use of steady state Kalman Filter for cases and deaths prediction of Covid-19 can be effective for resources and prevention measures planning.

Introduction

The ongoing pandemic of Covid-19 has driven researchers to direct their efforts towards the fight against the disease and the crisis management in health sector and border control. An important aspect in controlling the spread of Covid-19 and in resources and prevention measures planning, is the prediction of new cases and new deaths. Well established epidemiological models for long-term predictions, include variations of the Susceptible – Infectious – Recovered (SIR), such as the Susceptible – Infectious – Recovered – Dead (SIRD) [1] or the Susceptible – Exposed – Symptomatic infectious – Asymptomatic infectious – Quarantined – Hospitalized – Recovered cases – Dead (SESAQHRD) model, which focus on the basic reproduction number (R0), and the per day infection mortality and recovery rates [2]. The lack of a Covid-19 registry with traceable data obtained in a systematic manner and the lack of effective, well targeted or constant counter measures make the modeling of the Covid-19 outbreak a difficult task. Machine learning techniques have also been proposed [3], in order to face this difficulty. In [3], a moving average (MA) filter is used in order to forecast new daily cases and a Long Short Term Memory for Data Training-SAE (LSTM-SAE) network model was used to forecast the virus spreading. Time varying Kalman filters have been applied to daily predict new cases and deaths in [4]. An Autoregressive Integrated Moving Average (ARIMA) model is used in [5] to predict the epidemiological trend of the prevalence and incidence of Covid-19. Kalman Filter with the Autoregressive Integrated Moving Average (ARIMA) model is used in [6] to obtain forecasts of active cases, recoveries and deaths related to Covid-19 in Pakistan. Kalman Filters have also been proposed to estimate the future spread of Covid-19 [7] giving satisfactory results on short term estimates; their performance is poorer for long term forecasting.

The motivation of this work has been the development of tools for cases and deaths predictions, based on steady state Kalman Filter, that will allow the health system and authorities at the local or regional level to better prepare and manage the crisis. They may also be used in an early warning and emergency response system which collects ubiquitous data, such as smart city or smart building data, and combines it with historical statistical data and data collected at testing points (hospitals, borders).

The novelty of this work concerns (a) the design of a Finite Impulse Response (FIR) filter with coefficients related to the golden section (b) the design of a Steady State Kalman Filter with parameters related to the golden section, (c) the design of the FIR form of the Steady State Kalman Filter which does not require all the previous observations, (d) the applicability of the proposed filters in short term (daily) prediction as

* Corresponding author.

https://doi.org/10.1016/j.rinp.2021.104391
Received 5 January 2021; Received in revised form 24 May 2021; Accepted 25 May 2021
Available online 28 May 2021

well as in long term prediction (quarantine period).

In Section "Short term daily prediction" we present the classical mean or average filter as well as the proposed golden steady state Kalman filter for short term prediction, namely daily cases and deaths prediction for Covid-19; we also present simulation results. In Section "Long term average prediction" we present the use of steady state Kalman filter for long term cases and deaths average prediction, for example average prediction per quarantine period (14 days); we also present simulation results. The conclusions are summarized in Section "Conclusions".

Short term daily prediction

The aim is to predict the Covid-19 new cases and deaths on a daily basis using previous observations (cases/deaths). Thus we need a measurement driven scalar model, where the measurement $z(k)$ is the number of new cases or new deaths observed at time k and the state $x(k)$ corresponds to the number of new cases and deaths at time k. Then the aim is to predict the number of new cases and deaths $x(k+1/k)$ at time $k+1$ given the measurements till time$k \geq 0$, i.e. given the observations $z(0), z(1), \cdots, z(k)$. The observations are assumed to be zero for negative time: $z(k) = 0, k < 0$. The prediction is the output of a filter, which has the observations in its input, as it is shown in Fig. 1. In the remaining of this section, we are going to design various filters, including Kalman filter.

Mean filter (MF) or average filter (AF)

The basic idea is that the prediction of daily cases/deaths can be computed as the mean (average) of previous observations.

The prediction is the mean (average) of M last observations.

$$x(k+1/k) = \sum_{i=0}^{M-1} m_i z(k-i) = \frac{1}{M} \sum_{i=0}^{M-1} z(k-i) \tag{1}$$

This is a Moving Average filter MA(M-1) with M equal coefficients:

$$m_i = \frac{1}{M}, i = 0, 1, \cdots, M-1 \tag{2}$$

Note that

$$\sum_{i=0}^{M-1} m_i = 1 \tag{3}$$

The MA filter is a Finite Impulse Response (FIR) filter. M is the order of the MA filter.

We name the filter defined by (1) with coefficients in (2) as Mean Filter (MF) or Average Filter (AF) due to the fact that the prediction is the mean (average) of the M last observations. M = 7 (1 week), M = 14 (2 weeks - quarantine time), M = 21 (3 weeks) are proposed in this case.

Golden FIR filter (GFIR)

The basic idea is that the prediction of daily cases/deaths can be computed using weighted previous observations.

The prediction is

$$x(k+1/k) = \sum_{i=0}^{M-1} d_i z(k-i) \tag{4}$$

The sum of the weights is

Fig. 1. Daily prediction using filter.

$$\sum_{i=0}^{M-1} d_i = 1 \tag{5}$$

The idea is to use M last observations.

Each observation will contribute to the prediction with weight a of the rest of the sum, where a is the golden section:

$$a = \frac{\sqrt{5}-1}{2} = 0.618 \tag{6}$$

which is the positive solution of the equation

$$a^2 + a - 1 = 0 \tag{7}$$

Then the weights are:

$$d_0 = a \cdot 1 = a$$

$$d_1 = a \cdot [1-a] = a \cdot a^2 = a^3$$

$$d_2 = a \cdot \left[1 - a - a^3\right] = a - a^2 - a^4 = a \cdot [1-a] - a^4 = a \cdot a^2 - a^4 = a^3 - a^4 = a^3 \cdot [1-a] = a^3 \cdot a^2 = a^5$$

...

$$d_{M-2} = a^{2M-3}$$

The last weight results from the restriction that the sum of the weights equals to 1.

$$\begin{aligned} d_{M-1} &= 1 - \left[a + a^3 + a^5 + \cdots + a^{2M-3}\right] = [1-a] - a^3 - a^5 - \cdots - a^{2M-3} \\ &= a^2 - a^3 - a^5 - \cdots - a^{2M-3} = a^2 \cdot [1-a] - a^5 - \cdots - a^{2M-3} \\ &= a^2 \cdot a^2 - a^5 - \cdots - a^{2M-3} = a^4 - a^5 - \cdots - a^{2M-3} = \cdots \\ &= a^{2M-4} - a^{2M-3} = a^{2M-4} \cdot [1-a] = a^{2M-4} \cdot a^2 = a^{2M-2} \end{aligned}$$

Hence the weights are related to the golden section and the coefficients of (4) are:

$$d_i = a^{2i+1}, i = 0, 1, \cdots, M-2 \tag{8}$$

$$d_{M-1} = a^{2M-2} \tag{9}$$

The derived filter is also a Moving Average filter MA(M−1) whose coefficients are:

$$a, a^3, a^5, a^7, \cdots, a^{2M-3}, a^{2M-2}$$

The contribution of the observations in the forecast decreases as the observations move away from the prediction time, in line with weights associated with the golden section.

This MA filter is a Finite Impulse Response (FIR) filter.

We name the filter defined by (4) with coefficients in (8)-(9) as Golden FIR filter (GFIR) due to the fact that the prediction results from a FIR filter with coefficients related to the golden section. M = 7 (1 week), M = 14 (2 weeks - quarantine time), M = 21 (3 weeks) can be used as in Mean Filter (MF) or Average Filter (AF).

Golden steady state Kalman filter (GSSKF)

The basic idea is that the prediction of daily cases/deaths can be computed using steady state Kalman filter.

The general discrete time scalar linear model used to formulate the Kalman Filter (KF) consists of the dynamic and the statistical model. The dynamic model expresses the relationship between state and the measurement and is described by the following state space equations:

$$x(k+1) = F(k+1,k)x(k) + w(k) \tag{10}$$

$$z(k) = H(k)x(k) + v(k) \tag{11}$$

for $k = 0, 1, \cdots$, with initial condition $x(0)$,

where $x(k)$ is the state vector, $z(k)$ is measurement vector, $w(k)$ is the state noise and $v(k)$ is the measurement noise at time k. The transition parameter $F(k+1,k)$ describes the relation between two succeeded states. The output parameter $H(k)$describes the relation between the state and the measurement at the same time. Assuming that the state corresponds to the measurement, we get

$$H = H(k) = 1 \tag{12}$$

Assuming that the state does change very slowly from day to day, we consider that it does not change from step to step, so we get

$$F = F(k+1,k) = 1 \tag{13}$$

The statistical model describes the nature of state and measurements. The basic assumption is that a) the state noise is a zero mean Gaussian process with variance $Q(k)$ and b) and the measurement noise is a zero mean Gaussian process with variance $R(k)$. The following assumptions also hold: (a) the initial value of the state $x(0)$ is a Gaussian random variable with mean x_0 and variance P_0, (b) the noise stochastic processes and the random variable $x(0)$ are independent.

The discrete time Kalman filter [8] is the most well-known algorithm that solves the filtering problem, producing the state estimation $x(k/k)$ and the corresponding estimation error variance $P(k/k)$ as well as the state prediction $x(k+1/k)$ and the corresponding prediction error variance $P(k+1/k)$.

The state and measurement noise variances $Q(k)$ and $R(k)$ concern a time period before the prediction time. In fact, $Q(k)$ is the variance of the difference of two succeeded observations and $R(k)$ is the variance of the measurements for a time period before the prediction time. In [4] it was proposed that the noise variances can be computed on line for the time interval from the beginning till the prediction time or for a fixed backwards time interval before the prediction time. Then the noise variances are time varying and the Time Varying Kalman Filter is derived:

$$K(k) = \frac{P(k/k-1)}{P(k/k-1)+R(k)} \tag{14}$$

$$x(k/k) = [1-K(k)]x(k/k-1) + K(k)z(k) \tag{15}$$

$$P(k/k) = [1-K(k)]P(k/k-1) \tag{16}$$

$$x(k+1/k) = x(k/k) \tag{17}$$

$$P(k+1/k) = Q(k) + P(k/k) \tag{18}$$

for $k = 0, 1, \cdots$, with initial conditions $x_0 = x(0/-1)$ and $P_0 = P(0/-1)$

where$K(k)$is the Kalman Filter gain.

A variation results assuming that the noise variances are constant, i.e. $Q = Q(k)$ and $R = R(k)$. The noise variances can be computed off-line for the time interval from the beginning of the disease till the beginning of producing predictions or for a fixed initial time interval determined by the previous pandemic wave. Then the noise variances are time invariant and the Time Invariant Kalman Filter is derived.

In the following, we propose a model with constant noise variances $Q = R = 1$. In this case, the Kalman Filter parameters are $F = H = Q = R = 1$ resulting in the random walk system.

Then the prediction is

$$x(k+1/k) = \frac{1}{P(k/k-1)+1}x(k/k-1) + \frac{P(k/k-1)}{P(k/k-1)+1}z(k) \tag{19}$$

and the prediction error variance is

$$P(k+1/k) = 1 + \frac{P(k/k-1)}{P(k/k-1)+1} \tag{20}$$

The prediction error variance can be computed iteratively by the previous equation, known as the Riccati equation [9]. The Riccati equation can be written as:

$$P(k+1/k) = \frac{2P(k/k-1)+1}{P(k/k-1)+1} \tag{21}$$

The prediction error variance tends to the steady state prediction error variance P [9] that satisfies the algebraic Riccati equation is:

$$P = \frac{2P+1}{P+1} \tag{22}$$

The algebraic Riccati equation can be written as:

$$P^2 - P - 1 = 0 \tag{23}$$

The positive solution is the steady state prediction error variance P:

$$P = \frac{1+\sqrt{5}}{2} = \frac{1}{a} \tag{24}$$

where a is the golden section in (6).

Then, from (14) and (24) the steady state Kalman filter gain is

$$K = \frac{P}{P+1} = \frac{\frac{1}{a}}{\frac{1}{a}+1} = \frac{1}{1+a} = a \tag{25}$$

Thus the steady state Kalman filter gain is $K = a$ and is equal to the golden section.

Then the Steady State Kalman Filter is derived:

$$x(k+1/k) = Ax(k/k-1) + Bz(k) \tag{26}$$

for $k = 0, 1, \cdots$, with initial condition $x_0 = x(0/-1)$,

where

$$A = [1-K] = 1-a = a^2 \tag{27}$$

$$B = K = a \tag{28}$$

as derived by the equalities in (25) and (7).

Thus, the Golden Steady State Kalman Filter (GSSKF) has been derived:

$$x(k+1/k) = a^2 x(k/k-1) + az(k) \tag{29}$$

for $k = 0, 1, \cdots$, with initial condition $x_0 = x(0/-1)$.

The parameters of this Steady State Kalman Filter are related to the golden section.

We are able to use zero initial condition $x_0 = x(0/-1) = 0$ or to use the last measurement before the prediction beginning as initial condition $x_0 = x(0/-1) = z(-1)$.

This is an Autoregressive Moving Average filter ARMA(1,0) with coefficients:

$b_0 = a$ and $a_1 = -a^2$

We name the filter defined in (29) as Golden Steady State Kalman Filter (GSSKF) due to the fact that the prediction results from the implementation of a Steady State Kalman Filter with parameters related to the golden section.

Golden FIR steady state Kalman filter (GFIRSSKF)

The basic idea is that the prediction of daily cases/deaths can be computed using the FIR form of the steady state Kalman filter.

The FIR form of the Steady State Kalman Filter (SSKF) can be implemented when $A < 1$ [10]. Thus, the FIR form of the Golden Steady State Kalman Filter (GSSKF) can be implemented due to the fact that

$$A = a^2 = 1 - a < 1 \tag{30}$$

Without loss of generality, let $x(0/-1) = 0$.

Then

$$x(1/0) = a^2 x(0/-1) + az(0) = az(0)$$

$$x(2/1) = a^2x(1/0) + az(1) = a^3z(0) + az(1)$$

$$x(3/2) = a^2x(2/1) + az(2) = a^5z(0) + a^3z(1) + az(2)$$

$$x(4/3) = a^2x(3/2) + az(3) = a^7z(0) + a^5z(1) + a^3z(2) + az(3)$$

…

$$x(M+1/M) = a^{2M+1}z(0) + a^{2M-1}z(1) + \cdots + a^3z(M-1) + az(M)$$

Working as in [10], let

$$a^{2M+1} < \varepsilon \tag{31}$$

then

$$x(M+1/M) = a^{2M-1}z(1) + \cdots + a^3z(M-1) + az(M)$$

$$x(M+2/M+1) = a^{2M-1}z(2) + \cdots + a^3z(M) + az(M+1)$$

$$x(M+3/M+2) = a^{2M-1}z(3) + \cdots + a^3z(M+1) + az(M+2)$$

…

Then

$$x(k+1/k) = a^{2M-1}z(k-(M-1)) + a^{2M-3}z(k-(M-2)) + \cdots + a^3z(k-1) + az(k)$$

It is obvious that this pattern holds for $k \geq M$. Recall that the observations are assumed to be zero for negative time; then this pattern holds for $k \geq 0$. Thus, the FIR Steady State Kalman Filter (FIRSSKF) is derived:

$$x(k+1/k) = \sum_{i=0}^{M-1} b_i z(k-i) \tag{32}$$

where

$$b_i = a^{2i+1}, i = 0, 1, \cdots, M-1 \tag{33}$$

The coefficients of the filter are related to the golden section.

This is a Moving Average filter MA(M-1) with coefficients:

$$a, a^3, a^5, a^7, \cdots, a^{2M-3}, a^{2M-1}$$

The MA filter is a Finite Impulse Response (FIR) filter.

We name the filter defined by (32) with coefficients in (33) as Golden FIR Steady State Kalman Filter (GFIRSSKF) due to the fact that the prediction results from the implementation of a FIR Steady State Kalman Filter with coefficients related to the golden section.

Proposed M values are: M = 4 ($a^{2M+1} = a^9 = 0.0132$), M = 7 ($a^{2M+1} = a^{15} = 7.3314 \cdot 10^{-4}$), M = 14 ($a^{2M+1} = a^{29} = 8.6968 \cdot 10^{-7}$), M = 21 ($a^{2M+1} = a^{43} = 1.0316 \cdot 10^{-9}$).

Note that GFIRSSKF differs from GFIR in the last coefficient.

The FIR implementation of the Steady State Kalman Filter requires the knowledge of a subset of previous time measurements to calculate the state estimate; there is no need of any previous estimates calculation [10]. Thus, the basic advantage of the FIR implementation of the Steady State Kalman Filter is that there is no need of all the previous observations.

Simulation results

The results presented are based on the publicly available data for Greece.

The data used concern the new cases and the new deaths in Greece for the time interval from Feb 26, 2020 till June 14, 2020 [11].

The output of each filter (prediction) is corrected computing the ceiling of the prediction, in order to produce integer predictions (pessimistic scenario).

Fig. 2 depicts the daily prediction of new cases in Greece using the Mean Filter (MF) or Average Filter (AF) of order M = 7 and M = 14, as well as using the Golden Steady State Kalman Filter (GSSKF).

Table 1 presents the mean percent absolute error and the mean absolute error prediction for new cases in Greece using MF, GFIR, GFIRSSKF with M = 4, M = 7, M = 14, M = 21 and GSSKF. The **mean absolute error** is computed as the average of the absolute difference between observation and prediction. The **mean percent average absolute error** is computed as the absolute difference between the average observation and average prediction divided by the average observation and multiplied by 100.

Table 2 presents the mean percent absolute error and the mean absolute error prediction for new deaths in Greece using MF, GFIR, GFIRSSKF with M = 4, M = 7, M = 14, M = 21 and GSSKF.

Long term average prediction

The basic idea is to process blocks of measurements of a time window corresponding for example to the quarantine period in order to predict the average of cases and deaths using steady state Kalman Filter. Then long term prediction is derived, in the sense of predicting the average of states. So, the cases/deaths prediction concerns the prediction at time moment equal to the middle of the time window, that can be selected to correspond to a week, or a quarantine period. The length of the time window is denoted by N, i.e. the average prediction is derived every N (forward) days, for example 7 days (1 week), 14 days (quarantine period).

Assume the time invariant model with parameters $F = H = Q = R = 1$.

Then

$$x(k+1) = x(k) + w(k) \tag{34}$$

$$z(k) = x(k) + v(k) \tag{35}$$

Let us now use the **average** values of the state and measurement variables for every N (forward) time instants, defining them respectively as:

$$\overline{x_N}(j) = \frac{1}{N}\sum_{i=0}^{N-1} x(jN+i), j = 0, 1, \cdots \tag{36}$$

$$\overline{z_N}(j) = \frac{1}{N}\sum_{i=0}^{N-1} z(jN+i), j = 0, 1, \cdots \tag{37}$$

From (34) we get

$$x(k+1) = x(k) + w(k)$$

$$x(k+2) = x(k+1) + w(k+1) = x(k) + w(k) + w(k+1)$$

$$x(k+3) = x(k+3) + w(k+2) = x(k) + w(k) + w(k+1) + w(k+2)$$

…

$$x(k+\mathrm{N}) = x(k) + w(k) + w(k+1) + w(k+2) + w(k+(N-1))$$

i.e.

$$x(k+\mathrm{N}) = x(k) + \sum_{i=0}^{N-1} w(k+i) \tag{38}$$

Then

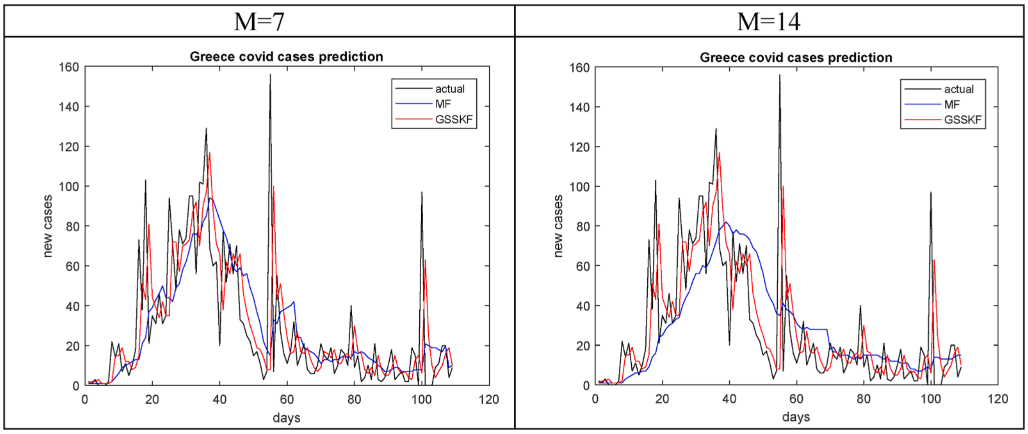

Fig. 2. Daily cases prediction using MF and GSSKF.

Table 1
Daily error cases prediction in Greece using MF, GFIR, GSSKF, GFIRSSKF.

mean percent average absolute error				
	MF	**GFIR**	**GSSKF**	**GFIRSSKF**
M = 4 (4 days)	0.4469	1.2448	1.3406	0.7980
M = 7 (1 week)	0.0958	1.4044	1.3406	1.2448
M = 14 (2 weeks - quarantine time)	1.4044	1.3406	1.3406	1.3406
M = 21 (3 weeks)	3.0322	1.3406	1.3406	1.3406
mean absolute error				
	MF	**GFIR**	**GSSKF**	**GFIRSSKF**
M = 4 (4 days)	16.0000	15.4404	15.5596	15.3670
M = 7 (1 week)	15.4954	15.5229	15.5596	15.5505
M = 14 (2 weeks - quarantine time)	17.3578	15.5596	15.5596	15.5596
M = 21 (3 weeks)	19.7523	15.5596	15.5596	15.5596

Table 2
Daily error deaths prediction in Greece using MF, GFIR, GSSKF, GFIRSSKF.

mean percent average absolute error				
	MF	**GFIR**	**GSSKF**	**GFIRSSKF**
M = 4 (4 days)	20.8791	27.4725	28.5714	26.3736
M = 7 (1 week)	22.5275	28.5714	28.5714	28.5714
M = 14 (2 weeks - quarantine time)	24.1758	28.5714	28.5714	28.5714
M = 21 (3 weeks)	25.2747	28.5714	28.5714	28.5714
mean absolute error				
	MF	**GFIR**	**GSSKF**	**GFIRSSKF**
M = 4 (4 days)	1.2766	1.3617	1.3617	1.3404
M = 7 (1 week)	1.2447	1.3617	1.3617	1.3617
M = 14 (2 weeks - quarantine time)	1.3830	1.3617	1.3617	1.3617
M = 21 (3 weeks)	1.4681	1.3617	1.3617	1.3617

$$\overline{x_N}(j+1) = \frac{1}{N}\sum_{i=0}^{N-1} x((j+1)N+i) =$$
$$\frac{1}{N}\sum_{i=0}^{N-1} x(jN+i+N) =$$
$$\frac{1}{N}\sum_{i=0}^{N-1}\left\{x(jN+i) + \sum_{\ell=0}^{N-1} w(jN+i+\ell)\right\} =$$
$$\frac{1}{N}\sum_{i=0}^{N-1} x(jN+i) + \frac{1}{N}\sum_{i=0}^{N-1}\sum_{\ell=0}^{N-1} w(jN+i+\ell) =$$
$$\overline{x_N}(j) + \frac{1}{N}\sum_{i=0}^{N-1}\sum_{\ell=0}^{N-1} w(jN+i+\ell)$$

Hence

$$\overline{x_N}(j+1) = \overline{x_N}(j) + \overline{w_N}(j) \tag{39}$$

where

$$\overline{F} = 1 \tag{40}$$

and

$$\overline{w_N}(j) = \frac{1}{N}\sum_{i=0}^{N-1}\sum_{\ell=0}^{N-1} w(jN+i+\ell) \tag{41}$$

Using (41) and $Q = 1$ we have:

$$\overline{Q} = \frac{1}{N}\sum_{i=0}^{N-1}\sum_{\ell=0}^{N-1} 1 = \frac{1}{N}\sum_{i=0}^{N-1} N = \frac{N}{N}\sum_{i=0}^{N-1} 1 = N \tag{42}$$

Also

$$\overline{z_N}(j) = \frac{1}{N}\sum_{i=0}^{N-1} z(jN+i) = \frac{1}{N}\sum_{i=0}^{N-1}\{x(jN+i) + v(jN+i)\}$$
$$= \frac{1}{N}\sum_{i=0}^{N-1} x(jN+i) + \frac{1}{N}\sum_{i=0}^{N-1} v(jN+i) = \overline{x_N}(j) + \frac{1}{N}\sum_{i=0}^{N-1} v(jN+i)$$

Hence

$$\overline{z_N}(j) = \overline{x_N}(j) + \overline{v_N}(j) \tag{43}$$

where

$$\overline{H} = 1 \tag{44}$$

and

$$\overline{v_N}(j) = \frac{1}{N}\sum_{i=0}^{N-1} v(jN+i) \tag{45}$$

Using (45) and $R = 1$ we have:

$$\overline{R} = \frac{1}{N}\sum_{i=0}^{N-1} 1 = \frac{1}{N}\sum_{i=0}^{N-1} 1 = \frac{N}{N} = 1 \tag{46}$$

Hence the parameters of the state space model (per time instant i for the average of time window of length N) are: $\overline{F} = \overline{H} = 1, \overline{Q} = N, R = 1$.

Then the algebraic Riccati equation is:

$$\overline{P} = N + \frac{\overline{P}}{\overline{P} + 1} \tag{47}$$

and it is written as:

$$\overline{P}^2 - N\overline{P} - N = 0 \tag{48}$$

The positive solution is the steady state prediction error variance P:

$$\overline{P} = \frac{N + \sqrt{N^2 + 4N}}{2} \tag{49}$$

Then, the steady state Kalman filter gain is

$$\overline{K} = \frac{\overline{P}}{\overline{P} + 1} = \frac{N + \sqrt{N^2 + 4N}}{N + 2 + \sqrt{N^2 + 4N}} = \frac{-N + \sqrt{N^2 + 4N}}{2} \tag{50}$$

and

$$\overline{A} = \left[1 - \overline{K}\right] = \frac{2}{N + 2 + \sqrt{N^2 + 4N}} = \frac{N + 2 - \sqrt{N^2 + 4N}}{2} \tag{51}$$

Hence, the average prediction is

$$\overline{x_N}(j+1/j) = \overline{A}\,\overline{x_N}(j/j-1) + \overline{K}\,\overline{z_N}(j) \tag{52}$$

We name the filter defined by (52) as Average Golden Steady State Kalman Filter (AGSSKF) due to the fact that the filter results from GSSKF used for daily prediction.

Note that average prediction is derived every N (forward) days and the steady state Kalman filter parameters depend on N. In the case where N = 1 (1 day) the AGSSKF is GSSKF. N = 7 (1 week), N = 14 (2 weeks - quarantine time) can be used.

For the FIR filters MF, GFIR, GFIRSSKF, form (1), (4), (32) we get the daily prediction as

$$x(k+1/k) = \sum_{i=0}^{M-1} c_i z(k-i) \tag{53}$$

with coefficients $c_i = m_i$, $c_i = d_i$, $c_i = b_i$, $i = 0, 1, \cdots, M-1$ defined in (2), (8)-(9) and (33) respectively.

We define the **average** values of the measurement variables for every set of N measurements:

$$\overline{z_N}(j) = \frac{1}{N}\sum_{i=0}^{N-1} z(jN+i), j = 0, 1, \cdots \tag{54}$$

Then the **average prediction** for the FIR filters MF, GFIR, GFIRSSKF is

$$\overline{x_N}(j+1/j) = \sum_{i=0}^{M-1} b_i \overline{z_N}(j-i), j = 0, 1, \cdots \tag{55}$$

with coefficients defined in (2), (8)-(9) and (33) respectively.

Simulation results

The data used concern the new cases and the new deaths in Greece for the time interval from Mar 16, 2020 till Oct 11, 2020 [11]. The prediction period is of 210 days (i.e. 30 weeks or 15 quarantine periods).

Fig. 3 depicts the daily prediction of new cases in Greece using the Mean filter (MF) or Average filter (AF) of order M = 7 and M = 14, as well as using the Average Golden Steady State Kalman Filter (AGSSKF) with N = 1.

Fig. 4 depicts the average prediction of new cases in Greece using the Mean filter (MF) or Average filter (AF) of order M = 7 and M = 14, as well as using the Average Golden Steady State Kalman Filter (AGSSKF) with N = 7 (week) and N = 14 (quarantine period).

Table 3 presents the mean absolute error prediction for new cases in Greece using MF, GFIR, GFIRSSKF with M = 7, M = 14, as well as AGSSKF with N = 1 (day), N = 7 (week) and N = 14 (quarantine period).

Table 4 presents the mean absolute error prediction for new deaths in Greece using MF, GFIR, GFIRSSKF with M = 7, M = 14, as well as AGSSKF with N = 1 (day), N = 7 (week) and N = 14 (quarantine period).

Conclusions

The aim of this paper was to develop prediction filters in order to predict the cases and deaths of Covid-19 in a daily basis (short term prediction) or in a quarantine period basis (long term prediction).

Aiming the cases and deaths of Covid-19 daily predictions we have used a) the classical mean or average filter, b) the proposed Golden FIR filter (GFIR) with coefficients related to the golden section, c) the developed the Golden Steady State Kalman Filter (GSSKF) with parameters related to the golden section and d) the Golden FIR Steady State Kalman Filter (GFIRSSKF) with the advantage that is that there is no need of all the previous observations.

The basic results are:

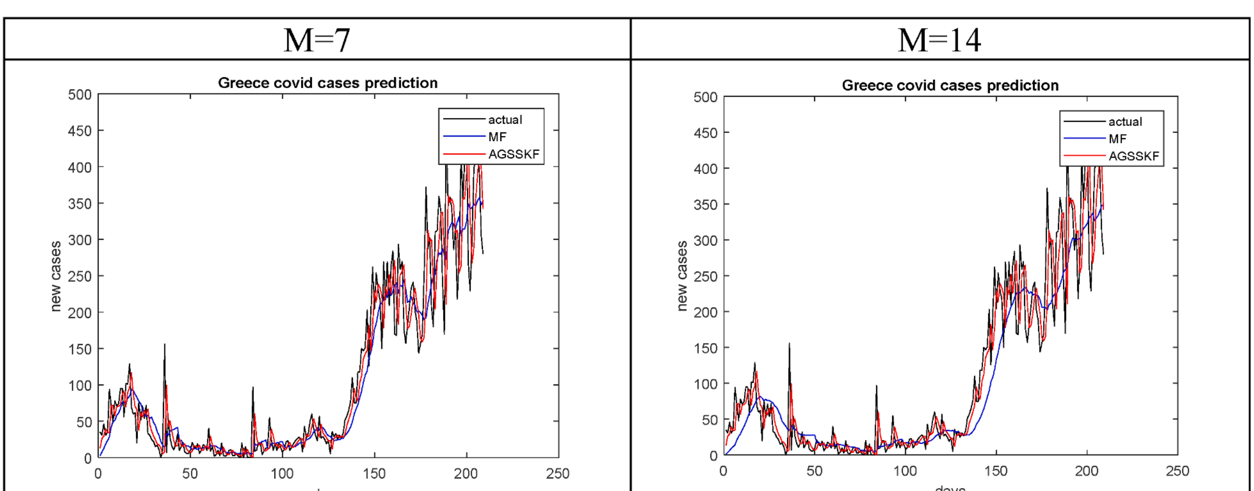

Fig. 3. Daily prediction of new cases in Greece using MF and AGSSKF.

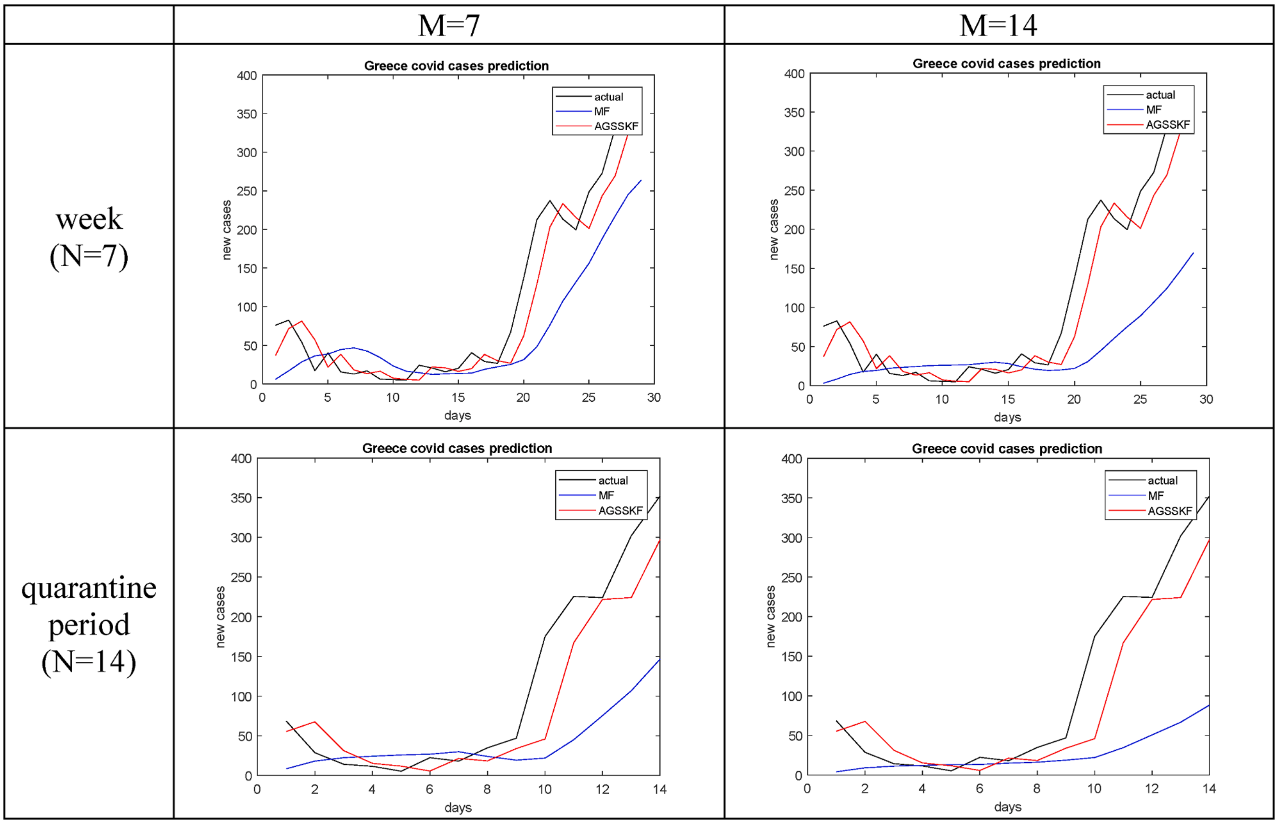

Fig. 4. Average prediction of new cases in Greece using MF and AGSSKF.

Table 3
Daily and average mean absolute error for cases prediction in Greece using MF, GFIR, AGSSKF, GFIRSSKF.

		MF	GFIR	AGSSKF	GFIRSSKF
N = 1	M = 7 (1 week)	28.6822	27.1767	27.1842	27.1859
(1 day)	M = 14 (2 weeks)	33.3551	27.1842	27.1842	27.1842
N = 7	M = 7 (1 week)	52.7544	27.5232	23.7433	27.5460
(1 week)	M = 14 (2 weeks)	72.1650	27.5346	23.7433	27.5346
N = 14	M = 7 (1 week)	74.9832	43.5088	32.2581	43.5281
(2 weeks)	M = 14 (2 weeks)	83.5886	43.5112	32.2581	43.5112

Table 4
Daily and average mean absolute error for deaths prediction in Greece using MF, GFIR, AGSSKF, GFIRSSKF.

		MF	GFIR	AGSSKF	GFIRSSKF
N = 1	M = 7 (1 week)	1.2057	1.3103	1.3103	1.3102
(1 day)	M = 14 (2 weeks)	1.2666	1.3103	1.3103	1.3103
N = 7	M = 7 (1 week)	1.3547	0.8196	0.8013	0.8192
(1 week)	M = 14 (2 weeks)	1.6615	0.8196	0.8013	0.8196
N = 14	M = 7 (1 week)	1.6159	1.0620	0.8615	1.0620
(2 weeks)	M = 14 (2 weeks)	1.4103	1.0612	0.8615	1.0612

- prediction using MF becomes worse as M increases
- prediction using SSKF becomes better as M increases; there is no prediction improvement after some M
- the FIR form of the SSKF requires the knowledge only of a subset of previous time measurements to calculate estimate
- cases prediction using SSKF is better than MA prediction of order M = 14 in Greece
- the mean absolute error in Greece new deaths prediction is of the order of 1 death
- It was found that the proposed golden steady state Kalman Filter has a satisfactory behavior compared with the classical mean or average filter

Aiming the cases and deaths of Covid-19 average predictions we proposed to process blocks of measurements of time window corresponding for example to the quarantine period. We designed the Golden Steady State Kalman Filter (GSSKF) with parameters related to the golden section.

The basic results are:

- cases and deaths average prediction using SSKF is better than MA prediction
- MA prediction becomes worst as the time window increases
- the mean absolute error in Greece new deaths average prediction is of the order of 1 death
- the proposed golden steady state Kalman Filter produces more accurate predictions than the classical mean or average filter does

Thus, the developed golden steady state Kalman Filter can be used in daily or average prediction producing reliable predictions. Hence, the use of steady state Kalman Filter for cases and deaths prediction of Covid-19 can be effective for resources and prevention measures planning.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Anastassopoulou C, Russo L, Tsakris A, Siettos C, Othumpangat S. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. PLoS ONE 2020;15 (3):e0230405. https://doi.org/10.1371/journal.pone.0230405.
[2] Nabi KN. Forecasting COVID-19 pandemic: A data-driven analysis. Chaos, Solitons Fractals 2020;139:110046. https://doi.org/10.1016/j.chaos.2020.110046.
[3] Pereira IG, Guerin JM, Silva Júnior AG, Garcia GS, Piscitelli P, Miani A, et al. Forecasting Covid-19 Dynamics in Brazil: A Data Driven Approach. Int J Environ Res Public Health 2020;17(14):5115. https://doi.org/10.3390/ijerph17145115.
[4] Assimakis N, Ktena A, Manasis C, Mele E, Kunicina N, Zabasta A, et al. Using the time varying Kalman filter for prediction of Covid-19 cases in Latvia and Greece. 2020 IEEE 61th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON). 2020.
[5] Domenico B, Giovanetti M, Lazzaro V, Silvia A, Massimo C, Application of the ARIMA model on the COVID-2019 epidemic dataset, Data in Brief 105340, 2020.
[6] Aslam M. Using the kalman filter with Arima for the COVID-19 pandemic dataset of Pakistan. Data in Brief 2020;31:105854. https://doi.org/10.1016/j.dib.2020.105854.
[7] Koushlendra, K.S., Prachi, D., Manish, K.B., Kalman Filter Based Short Term Prediction Model for COVID-19 Spread, medRxiv preprint doi: https://doi.org/10.1101/2020.05.30.20117416, 2020.
[8] Kalman RE. A new approach to linear filtering and prediction problems. J Bas Eng, Trans ASME Ser D 1960;8(1):34–45.
[9] Anderson BDO, Moore JB. Optimal Filtering. New York: Dover Publications; 2005.
[10] Assimakis N, Adam M. FIR implementation of the steady state Kalman filter. Int J Signal Imaging Systems Eng (IJSISE) 2008;1(3/4):279–86.
[11] Daily Overview, Information about Covid-19 in Greece, https://covid19.gov.gr/covid19-live-analytics/. [Accessed on July 15 and November 2].