Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company’s public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Unknown uncertainties in the COVID-19 pandemic: Multi-dimensional identification and mathematical modelling for the analysis and estimation of the casualties

Onder Tutsoy\textsuperscript{a}, Kemal Balikci\textsuperscript{b}, Naime Filiz Ozdil\textsuperscript{a}

\textsuperscript{a} Adana Alparslan Turkes Science and Technology University, Adana, Turkey
\textsuperscript{b} Osmaniye Korkut Ata University, Osmaniye, Turkey

**Abstract**

Insights about the dominant dynamics, coupled structures and the unknown uncertainties of the pandemic diseases play an important role in determining the future characteristics of the pandemic diseases. To enhance the prediction capabilities of the models, properties of the unknown uncertainties in the pandemic disease, which can be utterly random, or function of the system dynamics, or it can be correlated with an unknown function, should be determined. The known structures and amount of the uncertainties can also help the state authorities to improve the policies based on the recognized source of the uncertainties. For instance, the uncertainties correlated with an unknown function imply existence of an undetected factor in the casualties. In this paper, we extend the SpID-N (Suspicious-Infected-Death with non-pharmacological policies) model as in the form of MIMO (Multi-Input-Multi-Output) structure by adding the multi-dimensional unknown uncertainties. The results confirm that the infected and death sub-models mostly have random uncertainties (due undetected casualties) whereas the suspicious sub-model has uncertainties correlated with the internal dynamics (governmental policy of increasing the number of the daily tests) for Turkey. However, since the developed MIMO model parameters are learned from the data (daily reported casualties), it can be easily adapted for other countries. Obtained model with the corresponding uncertainties predicts a distinctive second peak where the number of deaths, infected and suspicious casualties disappear in 240, 290, and more than 300 days, respectively, for Turkey.

* Corresponding author.
E-mail address: otutsoy@atu.edu.tr (O. Tutsoy).

https://doi.org/10.1016/jdsp.2021.103058
1051-2004 © 2021 Elsevier Inc. All rights reserved.

1. Introduction

CORONAVIRUS disease 2019 (COVID-19) is caused by acute respiratory syndrome coronavirus 2 (SARS-COV-2) endures to be threat for the societies since there is no confirmed cures yet \cite{1}. As of August 3, 2020 more than 26 million confirmed cases including more than 863,000 deaths have been reported in 119 countries \cite{2}. Together with the catastrophic impacts on the health systems, it has also detrimental effects on the economies causing the job losses, stock markets’ volatilities and uncertainties in businesses. These results have urged the state authorities to take actions against the pandemic \cite{3}. These actions include curfews and restrictions imposed on the vulnerable people and also closure of the schools and universities \cite{4}. However, recent casualties indicate that the number of the infected and death individuals grow rather than shrink after releasing the restrictions and curfews \cite{5}. This emerges the policy makers to have knowledge about the future character of the pandemic diseases under various kinds of uncertainties \cite{6}. Therefore, the key focus of this paper is to develop a comprehensive parametric model which takes into account the internal dynamics, coupling dynamics, non-pharmacological policies and unknown uncertainties. Our recently published paper developed SpID model \cite{7} considered second order internal dynamics together with the slightly coupled dynamics. The sub-models of the SpID model are second order which can represent the higher order dynamics such as the peak value of the casualties with a natural frequency and damping. The natural frequency is related to internal dynamics (i.e. a large natural frequency means a small rising time yielding a sharp increase in the casualties) and the damping is related to the coupling dynamics and external impacts (i.e. confining the spread of the virus). We also recently proposed SpID-N model covering the non-pharmacological policies as external inputs \cite{8}. Since there is no available data for the non-pharmacological policies, we have parametrized them by using the facts and insights about the pandemic diseases. This research pre-
sented in this paper extends the SpID-N model by incorporating it with the unknown multi-dimensional uncertainties. These uncertainties might occur due to unmeasured casualties, inconsistent non-pharmacological policies, or pharmacological policies varying from hospital to hospital.

The COVID-19 virus encountered in bats in China, closely resembling the Betacoronavirus virus family including the SARS-COV and Middle East respiratory syndrome (MERS-COV) with 99.98% genome sequence similarity, where all can transmit from human-to-human [9]. However, even though they are in the same family, the COVID-19 exhibits previously unseen peculiar pathogenetic, epidemiological and clinical properties which are not utterly understood yet [10]. In addition, although the 58% and 70% of the SARS and MERS were contracted as a result of being hospitalized, respectively, the COVID-19 is extensively transmitted in the publics [10,11]. The key similarity among the SARS, MERS and COVID-19 is that the people with weak immune systems are more vulnerable to the pandemic [12]. Specifically, the COVID-19 stemmed death casualties indicate that the elderly people, male sex, obesity, diabetes, hypertension and ethnicity are the key factors in deaths [13]. Therefore, to protect the elderly people with the chronic diseases, curfews and restrictions have been implemented on them [14].

Precise modelling of the pandemics allows us to predict the possible future casualties under various uncertainties and non-pharmacological policies such as the school closures and curfews on the weekends [15]. Susceptible-Infected-Recovered (SIR) model is extensively considered to assess the pandemic diseases under various transmission and infectious rates [16]. Conducted recent researches have focused on estimating the future COVID-19 casualties of Pakistan [17], China [15], Italy and France [18], Spain [19] and the USA with the SIR model [20]. Susceptible-Exposed-Infected-Recovered (SEIR) model which considers the number of the exposed people is the enriched form of the pandemic modelling [21]. It predicts the exposed individuals under time-varying contact rates and other constant parameters in Wuhan city of China [22], Indonesia [23] and in Japan [24]. The SEIRD model with Designated (D) additional component takes into account the age groups classified as children, young adults, adults and elderly to consider the specific parameters such as different mobility [21]. Dynamic Panel SIR (DP-SIR) model, considering the non-pharmacological policies such as travel restrictions and mask wearing, is proposed and the parameters are determined from the partial data of 9 countries including England, Singapore, the USA and South Korea [20].

Modelling refers existence of a great amount of prior knowledge about the systems whereas the system identification attributes the process based on the observed input-output data such as the non-pharmacological policies as the inputs and the current COVID-19 casualties as the outputs [25]. In this context, a further alternative approach for modelling a pandemic disease is to system identification approaches which are able to consider various uncertainty structures. For the best of the authors’ knowledge, system identification-based MIMO pandemic disease modelling approaches covering the multi-dimensional unknown uncertainties and non-pharmacological policies have not been considered in the corresponding literature yet. We consider ARX (autoregressive with exogenous input) model structure with random uncertainty [26], ARMAX (autoregressive moving average with exogenous input) model structure with coloured uncertainty [27] and OE (output error) model structure with measurement type uncertainty [25].

We can summarize the key contributions of the paper as:

1. We extend the SpID-N model, which is a single-input-multiple-outputs (SIMO) model, as in the form of the MIMO structure with the multi-dimensional unknown uncertainties. All the uncertainties and their corresponding parameters are unknown and learning approaches for them are introduced.
2. We construct the ARX, ARMAX, and OE based extended SpID-N models regarding different uncertainty structures. The uncertainty structure reveals the character of the undetected casualties.
3. We perform the Least Square (LS) based optimization approaches with time-varying optimal rates to learn the unknown parameters and uncertainties of the MIMO SpID-N models.
4. We extensively analyse the developed models in terms of the mean errors, standard deviations and confidence intervals, and also provide the predicted future COVID-19 casualties.

In the rest of the paper, Section 2 initially reviews the SIR and SpID-N models and then introduces the extended SpID-N model. Section 3 constructs the ARX, ARMAX and OE based MIMO extended SpID-N models. Section 4 performs experiments, provides numerical simulation results, validates the obtained models and predicts the future COVID-19 casualties for Turkey.

2. Extended SpID-N model

This section initially reviews the SIR and SpID-N models and then introduces the extended SpID-N model. The key aim of this section is to provide insightful knowledge about the models that enables us to construct system identification-based approaches for the pandemic diseases, which can ease the model validation process.

2.1. The SIR model

Since the SIR model is extensively considered for the estimation of the future behaviour of the pandemic diseases, lately for the COVID-19 casualties, we first introduce the SIR model briefly. It consists of three continuous ordinary differential equations (ODE) given by

\[ \dot{S}(t) = -\beta S(t)I(t) \]  
\[ \dot{I}(t) = \beta S(t)I(t) - \gamma R(t) \]  
\[ \dot{R}(t) = \gamma R(t) \]

where

- \( S(t) \) represents the Susceptible (S) individuals,
- \( I(t) \) represents the Infected (I) individuals,
- \( R(t) \) represents the Recovered (R) individuals,
- \( \beta \) represents the transmission rate,
- \( \gamma \) represents the recovery rate.

The model assumptions for the SIR model are: 1) The population is fixed and an infected person leaves the susceptible group and returns back if recovered, 2) Age, sex, social status, and race do not affect the probability of being infected, 3) Nobody has inherited immunity, 4) The population is homogenously mix. The next sub-section presents the SpID-N model that we have developed recently [8].

2.2. The SpID-N model

The SpID-N model (Table 1) has single input \( u_k \) and multiple outputs which are \( S_p, I_k \) and \( D_k \). Therefore, it is a Single-Input-Multiple-Outputs (SIMO) model represented as

\[ S_{p,k+2} = a_1 S_{p,k+1} + a_0 S_p + b_3 I_k + c_1 u_k \]  

(4)
Table 1
Parameters of the SpID-N model.

| Parameter | Explanation |
|-----------|-------------|
| $Sp$      | Suspicious (Sp) individuals who are tested and/or quarantined |
| $I$       | Infected (I) individuals |
| $D$       | Death (D) due to the pandemic |
| $u_k$     | The input (external impacts) which represents the imposed curfews and/or restrictions called as the non-pharmacological policies N |
| $a_0$, $a_1$, $b_0$, $b_1$, $d_0$, $d_1$ | Unknown internal parameters associated with the $Sp_k$, $I_k$, and $D_k$ sub-models |
| $a_1$, $b_4$, $d_3$ | Unknown coupling parameters of the $Sp_k$, $I_k$, and $D_k$ sub-models |
| $c_1$, $c_2$, $c_3$ | Unknown parameters of the input |

Table 2
Parameters of the extended SpID-N model.

| Parameter | Explanation |
|-----------|-------------|
| $u_k$     | Input (external impacts/forces) for the curfews imposed on the people age over 65, age under 20 and people with chronic disease |
| $u_k^{wh}$ | Input for the curfews and/or restrictions implemented during the weekends and/or holidays |
| $u_k^{su}$ | Input for the closure of the schools and universities |
| $e_k^{Sp}$, $e_k^{I}$, $e_k^{D}$ | Uncertainties associated with the $Sp_k$, $I_k$, and $D_k$ sub-models |
| $a_{11}$, $a_{22}$, $a_{33}$ | Unknown internal parameters associated with the $Sp_k$, $I_k$, and $D_k$ sub-models |
| $a_{12}$, $a_{21}$, $a_{23}$, $a_{31}$, $a_{32}$ | Unknown coupling parameters of the $Sp_k$, $I_k$, and $D_k$ sub-models |
| $b_{ij}$ (i = 1.3, j = 1.3) | Unknown parameters associated with the inputs $u_k^{su}$ and $u_k^{wh}$ |
| $c_{ij}$ (i = 1.3, j = 1.3) | Unknown parameters associated with the uncertainties $e_k^{Sp}$, $e_k^{I}$, and $e_k^{D}$ |

\[
I_{k+2} = b_1 I_{k+1} + b_0 I_k + a_3 Sp_k + d_3 D_k + c_2 u_k \\
D_{k+2} = d_1 D_{k+1} + d_0 D_k + b_4 I_k + c_3 u_k
\]  

The next subsection introduces the extended SpID-N model proposed in this paper.

2.3. The extended SpID-N model

The extended SpID-N model (Table 2), which is more inclusive than its former versions, has Multiple-Inputs-Multiple-Outputs (MIMO) with unknown $e_k$ uncertainties represented directly in discrete form as
\[
Sp_{k+1} = a_1 Sp_k + a_2 I_k + a_3 D_k + b_11 u_k^{I} + b_12 u_k^{wh} + b_13 u_k^{su} + c_{11} e_k^{Sp} + c_{12} e_k^{I} + c_{13} e_k^{D} \\
I_{k+1} = a_21 Sp_k + a_22 I_k + a_23 D_k + b_21 u_k^{I} + b_22 u_k^{wh} + b_23 u_k^{su} + c_{21} e_k^{Sp} + c_{22} e_k^{I} + c_{23} e_k^{D} \\
D_{k+1} = a_31 Sp_k + a_32 I_k + a_33 D_k + b_31 u_k^{I} + b_32 u_k^{wh} + b_33 u_k^{su} + c_{31} e_k^{Sp} + c_{32} e_k^{I} + c_{33} e_k^{D}
\]

The detailed derivation of the internal and coupling dynamics are available in our recent paper [7], and derivation of the non-pharmacological policies are available in our recent work [8]. This current paper focuses on identification and modelling of the multi-dimensional unknown uncertainties in the three parametric SpID-N model structures.

The SpID-N model consists of the death ($D$) sub-model rather than the recovered ($R$) sub-model as in the SIR model, because the next target of the research is to develop artificial intelligence based policy making algorithms aiming to minimize the pandemic casualties, where the extended SpID-N model is acting as the background parametric model and manipulated by the generated policy for learning the long-term optimal policy.

The assumptions for the extended SpID-N model are: 1) A person is in suspicious group if any engagement has occurred with an infected or death person and exhibiting the symptoms, 2) A person leaves the group after the medical testing or after 14 days quarantine, 3) A person becomes infected if the medical test is positive, 4) Age, sex, social status and race do not affect the probability of being infected, 5) No inherited immunity exists. These unknown internal, coupling and policy parameters will be learned in Section 3 by using the LS based ARX, ARMAX and OE optimization algorithms. In the next sub-section, we summarize the properties of the SIR, SpID, and extended SpID-N models and also provide insights about them.

The extended SpID-N model assumes that the virus is transmitted from human to human and does not consider the bats as the infection source which is extensively modelled and analysed recently in [29,30]. However, since the extended SpID-N model constructs uncertainty structures for the unknown sources in the casualties, depending on the character of the infection source, which can be the bats, the model can improve the estimates by updating its parameters with the available data.

2.4. Comparison of the SIR, SpID-N and extended SpID-N models

We can summarize the key properties of the three models as in Table 3.

We can highlight the key insights of the models as

**Insight 1:** The SIR model is represented with continuous ODEs. However, the SpID-N and extended SpID-N models are represented with the difference equations since the pandemic data are discrete. They are constructed in discrete forms, not converted from the continuous form.

**Insight 2:** The SpID-N model has second order sub-models whereas the extended SpID-N model has first order sub-models. Our previous research [8] has confirmed that the second order dynamics associated with the overshoot, damping, natural frequency occur due to applied non-pharmacological policies rather than the internal dynamics of the pandemic disease.

**Insight 3:** The SIR model has partially coupled dynamics since $I(t)$ sub-model is linearly coupled with $R(t)$ through $\gamma R(t)$ in Equation (2).

**Insight 4:** The SpID-N model has moderately coupled dynamics since the $Sp$ and $D$ sub-models in Equations (4) and (6), respectively, do not directly contribute to each other.

**Insight 5:** The SpID-N and extended SpID-N models have unknown parameters learned with the optimization algorithms.

**Insight 6:** The SpID-N model is partially stochastic since it has bounded uncertainties in the non-pharmacological policy.

**Insight 7:** The extended SpID-N model is stochastic since it covers the unmeasured uncertainties which can be utterly random, function of the internal dynamics or function of an unknown property, which are the key concerns of this paper.

The next subsection briefly introduces the non-pharmacological policies acting as the inputs of the extended SpID-N model.
2.5. The non-pharmacological policies N

It is a fact that the casualties of the pandemic diseases are closely related to the applied non-pharmacological policies. Therefore, these policies should be incorporated into the model as the external inputs. Nevertheless, the corresponding data for these policies are not available. Henceforth, we need to derive background mathematical models to obtain the corresponding data, which can closely imitate the response of the non-pharmacological policies. In this sub-section, we consider the non-pharmacological policies covering the curfews and/or restrictions on people 1) with chronic disease, 2) people age over 65, 3) people under age 20, 4) restrictions on the weekends and holidays, 5) closure of the schools and universities.

2.6. Curfews on the people with chronic disease, age over 65 and age under 20

Governments or local authorities implement policies for the people who are vulnerable to the outbreaks and also for the people who spread the viruses even though they are not infected. Since the curfews implemented on people with the chronic disease, age over 65 and age under 20 have same characters such as imposed for a coincided duration of time, we will sum their response (impacts) and consider as one input of the extended SpID-N model. This will allow us to avoid null space in parameter optimization performed in Section 3.

It is reported that the symptoms of being infected can be indicated in 14 days where the possible peak might appear around day 7 as shown in Fig. 1.

We can comment on Fig. 1 slightly different as,

A curfew or restriction imposed for a day will contribute decidedly in 14 days with a possible transient ascent and a transient descent part. The transient ascent part of the response can be modelled as

\[ u_k^s = n^s (1 - \alpha^k) \]  \hspace{1cm} (10)

where

- \( u_k^s \) is the response of the curfew,
- \( n^s \) is the scaling factor of the number of the people under curfew,
- \( \alpha \) is the discount factor of the impact,
- \( k \) is the sample of discrete time (here \( k \) is the days)

Note that for \( \alpha = 0.71 \) and \( k = 7 \), the response (impact) in Fig. 1 gradually becomes maximum. We incorporate the random non-parametric uncertainty \( \sigma_k^s \) in Equation (10) to reflect the unmeasured quantities as

\[ u_k^s = n^s (1 - \alpha^{k - k_i} + \sigma_k^s), \hspace{0.5cm} \text{for } k = k_i, \ldots, k_n \]  \hspace{1cm} (11)

where \( k_i \) represents the start day of the curfew, \( k_n = k_i + 7 \) for this case.

With respect to the transient descent part of the response in Fig. 1, we can construct a mathematical model as

\[ u_k^d = n^d \left( \alpha^{k - k_i - 1} + \sigma_k^d \right), \hspace{0.5cm} \text{for } k = k_n + 1, \ldots, k_t \]  \hspace{1cm} (12)

where \( k_i = k_n + 7 \).

Partial models given by Equations (11) and (12) represent the uncertain response of only one day curfew, similar to the impulse response. To generalize the model for a duration of time in which the solution acts like a step response do

- Substitute \( k_n = k_1 + k_2^d \), where \( k_1^d \) is the start day and \( k_2^d \) is the duration of the curfew, in Equation (11). This modification allows us to represent the transient part and also the steady-state part of the response.
- Substitute \( k_i = k_n + 14 \) for the transient descent part of the response.

The next sub-section presents the non-pharmacological model of the curfews imposed on the weekends and holidays.

2.7. Curfews on the weekends and holidays

Since this type of curfews are implemented at certain intervals, it has piecewise impulses expressed as
\[ u_{i,k}^{wh} = n^{wh} \left(1 - \alpha^k - k_{i,k} + \sigma_{i,k}^{wh}\right) \delta_i \]

\[
\begin{cases}
    k = k_{i,k}^n, \ldots, k_n \\
    \delta_i = 0, \quad \text{curfew} \\
    \delta_i = 1, \quad \text{without curfew}
\end{cases}
\]

where

- \( \delta_i \) is the impulse representing the existence of the curfews,
- \( u_{i,k}^{wh} \) is the response of the curfew on the weekends and holidays,
- \( n^{wh} \) is the scaling factor of the number of the people under curfews on the weekends and holidays,
- \( \alpha \) is the discount factor of the impact,
- \( \sigma_{i,k}^{wh} \) is the random uncertainty in the response,
- \( k_n = k_{i,k}^n + k_{n}^{wh} \) where \( k_{i,k}^n \) is the start day and \( k_{n}^{wh} = 7 \) is the half duration of the impact.

The transient descent part is expressed as

\[ u_{i,k}^{wh} = n^{wh} \left(\alpha^k - k_{i,k} - 1 + \sigma_{i,k}^{wh}\right) \delta_i \]

\[
\begin{cases}
    k = k_{i,k} + 1, \ldots, k_t \\
    \delta_i = 1, \quad \text{curfew} \\
    \delta_i = 0, \quad \text{without curfew}
\end{cases}
\]

where \( k_t = k_n + 14 \). The summed responses \( u_{k}^{wh} \) is

\[ u_{k}^{wh} = \sum_{i=k-14}^{k} u_{i,k}^{wh} \tag{15} \]

The next sub-section forms the model for the schools and universities closure.

### 2.8. Schools and universities closure

Closure of the schools and universities refrain the students from mass gatherings, which impedes the spread of the virus. Since it is not a curfew, students continue to engage with each other, but less frequently. Therefore, its response \( u_{k}^{su} \) has a transient part which is same as the impulse response given by Equation (11). However, it has shorter transient descent part because it is expected to students re-involve each other in a short time until other curfews. Its transient response is interrupted and saturated around \( u_{sat} \) as

\[ u_{k}^{su} = n^{su} \left(u_{sat} + \sigma_{k}^{su}\right) \quad \text{for} \quad k = k_{i,k}^{su}, \ldots, k_n \tag{16} \]

where

- \( n^{su} \) is the scaling factor of the number of the students,
- \( \sigma_{k}^{su} \) is the random uncertainty in the response,
- \( k_k = k_{i,k}^{su} + k_{n}^{su} \) where \( k_{i,k}^{su} \) is the start day and \( k_{n}^{su} \) is the duration of the closure.

It can be seen from Fig. 2 that the response \( u_{k}^{su} \) becomes zero at \( k_n \), in which the curfews have been initiated and taken over the impacts of the schools and universities closure. The next section introduces the three well-known model structures for the MIMO SpID-N model.

### 3. Uncertainty structures for the extended SpID-N model

Developed models should cover the uncertainties intrinsic to the specific pandemic disease and also countries or regions. However, for the pandemic diseases such as the COVID-19 both the uncertainties and their structures are not available. In this part of the paper, three system identification approaches having different uncertainty structures are modified for the MIMO SpID-N model to reveal its uncertainty characteristics.

#### 3.1. MIMO representation of the extended SpID-N model

We represent the extended SpID-N model, consisting of the coupled three sub-models given by Equations (7), (8), and (9), in terms of a single discrete state space form as

\[\begin{bmatrix}
  S_{k+1} \\
  I_{k+1} \\
  D_{k+1}
\end{bmatrix} = \begin{bmatrix}
  a_{11} & a_{12} & a_{13} \\
  a_{21} & a_{22} & a_{23} \\
  a_{31} & a_{32} & a_{33}
\end{bmatrix} \begin{bmatrix}
  S_k \\
  I_k \\
  D_k
\end{bmatrix}
+ \begin{bmatrix}
  b_{11} & b_{12} & b_{13} \\
  b_{21} & b_{22} & b_{23} \\
  b_{31} & b_{32} & b_{33}
\end{bmatrix} \begin{bmatrix}
  u_k \\
  u_{k}^{wh} \\
  u_{k}^{su}
\end{bmatrix}
+ \begin{bmatrix}
  c_{11} & c_{12} & c_{13} \\
  c_{21} & c_{22} & c_{23} \\
  c_{31} & c_{32} & c_{33}
\end{bmatrix} \begin{bmatrix}
  e_{k}^S \\
  e_{k}^D
\end{bmatrix} \tag{17}
\]

Equation (17) can be simply expressed as

\[ x_{k+1} = Ax_k + Bu_k + Ce_k \tag{18} \]

where \( x_k = [S_k \ I_k \ D_k]^T \), \( x_{k+1} = z^T x_k \) with \( z \) discrete delay operator, \( u_k = [u_k \ u_{k}^{wh} \ u_{k}^{su}]^T \), \( e_k = [e_{k}^S \ e_{k}^D]^T \), and \( A, B, C \) are the corresponding unknown parameter matrices. Now we will introduce the system identification approaches for the MIMO SpID-N model in terms of the known bases and unknown parameters.

#### 3.2. The ARX based extended SpID-N model

The ARX model has additive \( e_k \) uncertainty in the form of white noise with no definite mean and bounded variance represented as

\[ x_{k+1} = Ax_k + Bu_k + e_k \tag{19} \]

The ARX based SpID-N estimator is

\[ \hat{x}_{k+1} = \hat{\theta}^T \phi_k + \hat{e}_k \tag{20} \]

where \( \hat{x}_{k+1} \) is the estimated output (casualties), \( \hat{e}_k \) is the estimated uncertainty, \( \hat{\theta} \) is the unknown parameter matrix and \( \phi_k \) is the known basis vector formed as
\[
\hat{\theta} = \begin{bmatrix} \hat{A}; \hat{B} \end{bmatrix}^T
\]
\[
\phi_k = \begin{bmatrix} x_k; u_k \end{bmatrix}
\] (21)

The temporal estimation error \( \delta_k \) between the real system in Equation (19) and the estimated model in Equation (20) is

\[
\delta_k = x_{k+1} - \hat{x}_{k+1} = (A - \hat{A}) x_k + (B - \hat{B}) u_k + \epsilon_k - \hat{\epsilon}_k
\] (23)

To determine the unknown parameter matrix, the recursive least squares (LS) with gradient descent approach is expressed as

\[
\hat{\theta}_{k+1} = \hat{\theta}_k - \eta \frac{\partial \delta_k^2}{\partial \theta_k} = \hat{\theta}_k + \eta \delta_k \phi_k
\] (24)

where \( \eta \) is the parameter update rate defined as \( 0 < \eta < 2/\|\phi_k\|^2 \).

Since \( \hat{\epsilon}_k \) is not available, we apply the iterative LS approach given by

\[
\hat{\epsilon}_{k+1} = \hat{\epsilon}_k - \eta^{\text{un}} \delta_k
\] (25)

where \( \eta^{\text{un}} \) is the update rate of the uncertainty. Configuration of the constructed ARX estimator is shown in Fig. 3.

Since the uncertainty of the model in Equation (19) is not scaled up with a parameter matrix/vector, the unknown parameter matrix in Equation (21) and the basis in Equation (22) are implicit function of the uncertainty \( \epsilon_k \). However, the ARMAX based extended SpID-N estimator, introduced next, is an explicit function of the uncertainty.

3.3. The ARMAX based extended SpID-N model

The ARMAX model considers a particular structure \( \hat{C} \) for the unknown uncertainty represented as

\[
\hat{x}_{k+1} = \hat{A} x_k + \hat{B} u_k + \hat{C} \hat{\epsilon}_k
\] (26)

where \( \hat{A}, \hat{B}, \hat{C} \), and \( \hat{\epsilon}_k \) are the estimates forming an unknown parameter matrix \( \hat{\theta} \) and a basis consisting of known \( x_k \) and \( u_k \) data, but unknown disturbance \( \hat{\epsilon}_k \) expressed as

\[
\hat{\theta} = \begin{bmatrix} \hat{A}; \hat{B}; \hat{C} \end{bmatrix}^T
\]
\[
\phi_k = \begin{bmatrix} x_k \ u_k \ \hat{\epsilon}_k \end{bmatrix}^T
\] (27)

The parameter update rule is same as in Equation (24). The temporal estimation error is the difference between the real system in Equation (18) and the estimated model in Equation (26) is

\[
\delta_k = x_{k+1} - \hat{x}_{k+1} = (A - \hat{A}) x_k + (B - \hat{B}) u_k + C \epsilon_k - \hat{C} \hat{\epsilon}_k
\] (29)

The estimated uncertainty is obtained by applying gradient descent on the temporal error in Equation (30) as

\[
\hat{\epsilon}_{k+1} = \epsilon_k - \frac{\partial \delta_k^2}{\partial \epsilon_k} = \hat{\epsilon}_k + \eta^{\text{un}} \delta_k \hat{C}
\] (30)

The parameter update rate \( \eta^{\text{un}} \) is \( 0 < \eta^{\text{un}} < 2/\|\hat{C}\| \). Configuration of the constructed ARMAX estimator is shown in Fig. 3. The ARMAX model can represent the coloured disturbance which is autocorrelated, but the OE model, introduced next, assumes additive uncertainty.

3.4. The OE based extended SpID-N model

The OE based extended SpID-N model estimator is

\[
\hat{x}_{k+1} = \hat{A} x_k + \hat{B} u_k + \hat{C} \hat{\epsilon}_k
\] (31)

The disturbance \( \hat{\epsilon}_k \) is multiplied by the \( \hat{A} \) matrix representing the internal and coupling dynamics of the SpID-N model. Its unknown parameter matrix \( \hat{\theta} \) and the known basis vector \( \phi_k \) are

\[
\hat{\theta} = \begin{bmatrix} \hat{A}; \hat{B} \end{bmatrix}^T
\]
\[
\phi_k = \begin{bmatrix} x_k \ u_k \end{bmatrix}^T
\] (32)

The unknown parameter is obtained with the recursive LS approach in Equation (24). The estimated \( \hat{\epsilon}_k \) is

\[
\hat{\epsilon}_{k+1} = \hat{\epsilon}_k + \eta^{\text{un}} \delta_k \hat{A}
\] (34)

where parameter update rate \( \eta^{\text{un}} \) is \( 0 < \eta^{\text{un}} < 2/\|\hat{A}\| \). Configuration of the constructed OE estimator is shown in Fig. 3.

4. Simulation results

This section initially presents the parameters of the extended SpID-N model and then analysis the obtained results with the ARX, ARMAX and OE models representing the uncertainties in different forms. For the analysis, we use the data provided by the Health Ministry of Turkey. However, the proposed model can be easily modified for the other countries as well.
The estimated ARX based extended SpID-N model is

\[
\begin{bmatrix}
\hat{S}_{p,k+1} \\
\hat{I}_{k+1} \\
\hat{D}_{k+1}
\end{bmatrix} = \begin{bmatrix}
1.09 & 0.5e - 1 & 0.4e - 2 \\
0.5e - 1 & 0.5e - 1 & 0.9e - 3 \\
0.9e - 3 & 0.2e - 2 & 0.2e - 4
\end{bmatrix} \begin{bmatrix}
\hat{S}_{p,k} \\
\hat{I}_{k} \\
\hat{D}_{k}
\end{bmatrix} + \begin{bmatrix}
0.1e - 3 & 0.1e - 3 & 0.2e - 4 \\
-0.3e - 4 & -0.1e - 4 & -0.5e - 4 \\
-0.8e - 6 & 0.9e - 6 & -0.2e - 5
\end{bmatrix} \begin{bmatrix}
\hat{u}_{k}^{I} \\
\hat{u}_{k}^{wh} \\
\hat{u}_{k}^{su}
\end{bmatrix}
\]

\[
+ \begin{bmatrix}
\hat{e}^{Sp}_{k} \\
\hat{e}^{I}_{k} \\
\hat{e}^{D}_{k}
\end{bmatrix}
\]

The estimated ARMA based extended SpID-N model is

\[
\begin{bmatrix}
\hat{S}_{p,k+1} \\
\hat{I}_{k+1} \\
\hat{D}_{k+1}
\end{bmatrix} = \begin{bmatrix}
0.97 & 0.3e - 2 & 0.1e - 2 \\
0.5e - 1 & 0.1e - 1 & 0.6e - 4 \\
0.1e - 2 & 0.1e - 2 & 0.1e - 4
\end{bmatrix} \begin{bmatrix}
\hat{S}_{p,k} \\
\hat{I}_{k} \\
\hat{D}_{k}
\end{bmatrix} + \begin{bmatrix}
0.7e - 5 & 0.1e - 4 & -0.4e - 4 \\
-0.2e - 4 & 0.1e - 4 & -0.7e - 4 \\
-0.6e - 6 & -0.7e - 6 & -0.3e - 5
\end{bmatrix} \begin{bmatrix}
\hat{u}_{k}^{I} \\
\hat{u}_{k}^{wh} \\
\hat{u}_{k}^{su}
\end{bmatrix}
\]

\[
+ \begin{bmatrix}
\hat{e}^{Sp}_{k} \\
\hat{e}^{I}_{k} \\
\hat{e}^{D}_{k}
\end{bmatrix}
\]

The internal and coupling parameters learned with the ARMAX model in Equation (36) have similar properties with the ARX model parameters discussed in the Section 3.2. The difference between them is that the ARMAX model has a matrix colouring the \(e_k\) uncertainties. Fig. 5 shows the reported and estimated casualties with the ARMAX based extended SpID-N model.

It can be clearly seen from Fig. 5 that the infected \(\hat{I}_k\) and death \(\hat{D}_k\) estimated casualties have large variances around the days \(k = 50\) and \(k = 100\). That region corresponds to the transient descent part of the non-pharmacological policies where the curfews and restrictions have been released and this resulted in fluctuations in the estimates. Moreover, starting from the day \(k = 170\), the \(\hat{I}_k\) and the \(\hat{D}_k\) sub-models produced oscillatory outputs where the COVID-19 casualties are about the increase. This increase has continued in the estimated future casualties until the second peak shown in Fig. 9. It is important to note that all the non-pharmacological policies have been lifted from the day \(k = 98\) as given in Table 7. Since the absence of the non-pharmacological policies cause low damping effect on the outputs, the casualties increase due to changes in the internal and/or external parameters.
4.4. Analysis of the OE based SpID-N model

The estimated OE based extended SpID-N model is

\[
\begin{bmatrix}
\hat{s}_{p_{k+1}} \\
\hat{i}_{k+1} \\
\hat{d}_{k+1}
\end{bmatrix}
= \begin{bmatrix}
1.1 & 0.6e - 1 & 0.4e - 2 \\
0.1e - 1 & 0.2e - 1 & -0.4e - 4 \\
0.4e - 3 & 0.2e - 2 & 0.2e - 4
\end{bmatrix}
\begin{bmatrix}
\hat{s}_p_k \\
\hat{i}_k \\
\hat{d}_k
\end{bmatrix}
+ \begin{bmatrix}
0.1e - 3 & 0.1e - 3 & 0.3e - 4 \\
-0.6e - 4 & 0.3e - 4 & -0.1e - 3 \\
-0.1e - 5 & 0.1e - 5 & -0.5e - 5
\end{bmatrix}
\begin{bmatrix}
u^i_k \\
u^{ih}_k \\
u^{iu}_k
\end{bmatrix}
+ \begin{bmatrix}
1.1 & 0.6e - 1 & 0.4e - 2 \\
0.1e - 1 & 0.2e - 1 & -0.4e - 4 \\
0.4e - 3 & 0.2e - 2 & 0.2e - 4
\end{bmatrix}
\begin{bmatrix}
\varepsilon^{ih}_k \\
\varepsilon^{iu}_k \\
\varepsilon^{ihh}_k \\
\varepsilon^{iuu}_k \\
\varepsilon^{i}$^{30}_k
\end{bmatrix}
\]  
\tag{37}

The learned OE model in Equation (37) has similar characters for the internal and coupling dynamics as in the ARX and ARMAX models analyzed in Sections 3.2 and 3.3. The difference is that the OE model has larger parameters for the non-pharmacological policies \(u_k\). This could occur due to the OE uncertainty structure which is scaled with the same internal and coupling parameters. Fig. 6 shows the reported and estimated COVID-19 casualties with the OE based SpID-N model.

The OE model weights the uncertainties with the \(\hat{A}\) matrix, which represents the internal and coupling dynamics of the diseases. Therefore, during the sharp learning process of the internal and coupling dynamics (around \(k = 10\) to \(k = 100\) days), the uncertainty fluctuates as can be seen from the estimated infected (\(\hat{i}_k\)) and death (\(\hat{d}_k\)) casualties in Fig. 6. However, this is not noticeable with the estimated suspicious (\(\hat{s}_p\)) casualties since it does not have a distinctive peak up to the \(k = 100\) day.

4.5. Comparison of the SpID-models

Fig. 7 demonstrates the effects of the uncertainties on the estimates. As can be seen from Fig. 7 there is an offset among the real casualties and the estimated casualties with the SpID, SpID-N, and extended SpID-N models. While the extended SpID-N model can closely follow the real casualties, the SpID model yields the largest error since it does not consider the non-pharmacological policies and uncertainties. The character of the offset can vary for different model structures, for instance for the OE model structure it changes with respect to the internal dynamics \(A\) as in Equation (31) whereas for the ARMAX model structure the uncertainty varies with respect to the \(C\) parameter as in Equation (26). However, with respect to the ARX, its coefficient is unity as in Equation (19); henceforth, contribution of the corresponding uncertainty is almost proportional (due to the modelling error) to the output of the ARX SpID-N model. The next section provides analysis of the ARX, ARMAX, and OE model structures.

4.6. Mean errors in the estimated models

Fig. 8 shows mean errors in the estimated casualties. As can be seen from Fig. 8, the ARX based extended SpID-N model has the smallest errors in its infected (\(\hat{i}\)) and death (\(\hat{d}\)) estimates. This implies that the disturbances in these estimates have more random characters. On the other hand, the ARMAX based has the smallest error in its suspicious (\(\hat{s}_p\)) estimates. This signifies that the corresponding uncertainty is correlated, so it can be predicted with a characterized function.

4.7. Predicted future COVID-19 casualties

Fig. 9 presents the predicted future COVID-19 casualties with the ARX based extended SpID-N model.

The ARX based extended SpID-N model predicts that the death (\(\hat{d}_k\)) casualties will converge to zero around 240 days, the infected (\(\hat{i}_k\)) casualties will disappear around 290 days and the suspicious (\(\hat{s}_p\)) casualties will require more than 300 days to be insignifi-
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.

Appendix A

Table 7
Parameters of the extended SpID-N model (Health Ministry of Turkey, 2020).

| Parameter | Value |
|-----------|-------|
| \( n^c \) | 26,567,000 People with chronic diseases |
| \( n^{65} \) | 1,517,000 People age over 65 without chronic diseases |
| \( n^{20} \) | 25,573,000 People age under 20 |
| \( n^{wh} \) | 29,497,000 Remaining people under curfews |
| \( n^{whu} \) | 26,048,000 Number of the school and university students |
| \( k_i \) | 23/02/2020 First COVID-19 casualty has seen |
| \( k_i = 1 \) | |
| \( k_\text{k}^{\text{I}^c} \) and \( k_\text{k}^{\text{I}^c} \) | 21/03/2020-01/06/2020 Start and end dates of the curfews for the people with chronic diseases and their corresponding day numbers |
| \( k_\text{k}^{65} \) and \( k_\text{k}^{65} \) | 21/03/2020-01/06/2020 Start and end dates of the curfews for the people age over 65 and their corresponding day numbers |
| \( k_\text{k}^{20} \) and \( k_\text{k}^{20} \) | 03/04/2020-01/06/2020 Start and end dates of the curfews for the people age under 20 and their corresponding day numbers |
| \( k_\text{k}^{wh} \) and \( k_\text{k}^{wh} \) | 16/03/2020 – cont. Start and end dates of the schools and universities closures and their corresponding day numbers |

This paper introduced the extended SpID-N model to estimate the casualties of the pandemic diseases. To reveal the character of the uncertainties, ARX, ARMAX and OE uncertainty structures developed for the extended SpID-N model which confirmed that the infected and death casualties have ARX type uncorrelated uncertainty. Moreover, the suspicious casualties have ARMAX type correlated uncertainties. The ARX based SpID-N model predicted convergent behaviour after sharp peaks in casualties in Turkey.

5. Conclusion

This paper introduced the extended SpID-N model to estimate the casualties of the pandemic diseases. To reveal the character of the uncertainties, ARX, ARMAX and OE uncertainty structures developed for the extended SpID-N model which confirmed that the infected and death casualties have ARX type uncorrelated uncertainty. Moreover, the suspicious casualties have ARMAX type correlated uncertainties. The ARX based SpID-N model predicted convergent behaviour after sharp peaks in casualties in Turkey.

Table 4
Mean errors of the extended SpID-N models.

| Parameter | ARX | ARMAX | OE |
|-----------|-----|-------|----|
| \( \hat{S}_p \) | 1204 | 1045 | 1221 |
| \( \hat{l} \) | 102  | 177  | 103 |
| \( \hat{A} \) | 2    | 4.2  | 2.7 |

Table 5
Mean errors of the extended SpID-N models.

| Parameter | ARX | ARMAX | OE |
|-----------|-----|-------|----|
| \( \hat{S}_p \) | 1222 | 1048 | 1232 |
| \( \hat{l} \) | 94   | 207  | 116 |
| \( \hat{A} \) | 2    | 4.6  | 3   |

Table 6
Confidence intervals of the extended SpID-N models.

| Parameter | ARX | ARMAX | OE |
|-----------|-----|-------|----|
| Lower \( \hat{S}_p \) | 1022 | 890  | 1038 |
| Upper \( \hat{S}_p \) | 1385 | 1201 | 1404 |
| Lower \( \hat{l} \) | 88   | 147  | 86  |
| Upper \( \hat{l} \) | 116  | 208  | 121 |
| Lower \( \hat{A} \) | 1.7  | 3.53 | 2.2 |
| Upper \( \hat{A} \) | 2.3  | 4.59 | 3.1 |

cant under the current conditions (Fig. 9). In addition, Fig. 9 clearly shows that the casualties will continue to increase until the second peaks. However, unexpected internal and/or external uncertainties possibly to affect the future character of the casualties.

Fig. 9. Predicted future COVID-19 casualties with the ARX based extended SpID-N model.
