Annual Reviews in Control 46 (2018) 94–106

Contents lists available at ScienceDirect

Annual Reviews in Control

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

Review article

Dealing with observational data in control

Emma D. Wilson[a], Quentin Clairon[b], Robin Henderson[b], C. James Taylor*,[a]

[a] Engineering Department, Lancaster University, UK
[b] School of Mathematics and Statistics, University of Newcastle, UK

ARTICLE INFO

Keywords:
Interdisciplinarity
Observational data
Causation
System identification
Missing measurements
Missing data
Longitudinal data
Periodic control
Irregular control
Event driven control

ABSTRACT

There is growing interest in the use of control theory for interdisciplinary applications, where data may be sparse or missing, be non-uniformly sampled, have greater uncertainty, and where there is no opportunity to collect repeat measurements. In such applications, problems posed by observational data and the issue of missing or irregular data need to be considered. We present a review on dealing with observational, missing and irregular data for control applications. This considers the following issues: (i) how to identify a system model from observational data subject to missing measurements, (ii) how to determine control inputs when output data includes missing measurements, and (iii) how to ensure stability when future update times may be missed. Dealing with observational data and missing measurements is a key problem within the statistics literature, so we introduce statistical methods for dealing with this type of data. We aim to enable the integration of well-developed statistical methods of dealing with missing data into control theory. An example problem of using anticoagulants to control the blood clotting speed of patients is used throughout the paper.

1. Introduction

Concepts and techniques from control theory are increasingly being used for interdisciplinary applications. Such applications (with selected examples) include:

- Biological systems, where controllers have been designed to manage the uptake of CO_2 during leaf photosynthesis (Paré, Walker, & McGrath, 2017; Taylor & Aerts, 2014); model predictive control (MPC) algorithms have been applied to control animal heart rate (Aerts, Gebruers, Van Camp, & Berckmans, 2008; Hunt & Liu, 2018; Leor-Librach, Bobrovsky, Eliash, & Kaplinsky, 1999) and growth (Aerts et al., 2003); and bio-inspired methods have been used to develop adaptive controllers (Wilson et al., 2015).
- Climate control, on a global scale, where optimal control theory has been used to develop strategies for climate manipulation aimed at reducing global warming (Jarvis, Leedal, Taylor, & Young, 2009; MacMartin, Kravitz, Keith, & Jarvis, 2014; Soldatenko & Yusupov, 2017); and on the micro-scale, where microclimates within greenhouses, ventilation chambers and grow cells have been controlled (Luan, Shi, & Liu, 2012; Tsitsimpelis, Wolfenden, & Taylor, 2016; Underwood, 2002).
- Economics, where constrained optimal control has been used to manage economic performance (Gaimon, 2002; Westcott, 1984).
- Epidemic analysis, where optimal control has been applied with the aim of increasing recovery rates and decreasing infection rates with finite resources (Behncke, 2000; Hansen & Day, 2011; Nowzari, Preciado, & Pappas, 2016).
- Adaptive interventions, which are sequences of treatments that are adapted to individuals to achieve health behaviour change, where MPC has been applied to behavioural research to develop controller-design-based interventions (Bekiroglu, Lagoa, Murphy, & Lanza, 2017; Deshpande, Nandola, Rivera, & Younger, 2014; Nandola & Rivera, 2013; Rivera, Pew, & Collins, 2007).
- Resource management and operations research, where control theory has been applied to planning, ecology and resource management, including MPC to determine policies for oceanic fisheries (Cliff & Vincent, 1973), optimal control for water resource management (Nicklow, 2000), and MPC to control the flow of traffic (Van den Berg, Hegyi, De Schutter, & Hellendoorn, 2003).
- Medical treatment control, for which examples include using Proportional-Integral-Derivative (PID) (Hoekstra, Vogelzang, Verbitskiy, & Nijsten, 2009; Rattan & Nasraway, 2013) and MPC algorithms (Chakrabarty, Zavitsanou, Doyle III, & Dassau, 2017; Hovorka et al., 2004; Plank et al., 2006) to control blood glucose levels by adjusting insulin inputs, and using Proportional-Integral (PI) and PID based algorithms to control anesthesia (Bibian, Ries, Huzmezan, & Dumont, 2005; Dumont, 2012; Gentilini et al., 2001).

Applying control theory is attractive as it provides a systematic way

* Corresponding author.
E-mail address: c.taylor@lancaster.ac.uk (C. . Taylor).

https://doi.org/10.1016/j.arcontrol.2018.05.001
Received 1 March 2018; Received in revised form 6 May 2018; Accepted 21 May 2018
Available online 29 May 2018

of designing interventions to achieve a specified performance in the presence of noise, uncertainty and external disturbances. However, in many of these applications the controller needs to be built from observational data which may be sparse or missing, be non-uniformly sampled and have large uncertainties. Further, often there is no opportunity to collect repeat measurements and there are questions as to causality of effects. This contrasts with the standard control approach where multiple measurements can be made on the system, the experimental protocol can be designed to enable 'ideal' data to be collected, retesting is straightforward and generally a single system is considered over a long time period. There is a risk, therefore, that when attempting control in some interdisciplinary applications there may be lack of consideration of the nature of the data that are available, especially when the data are observational rather than experimental.

If control system techniques are to be adopted for interdisciplinary applications where data collection is problematic, an awareness of the complexities of the data is required to avoid introducing bias. Our aim in this article is to review and discuss some of these issues, with most attention on causation and problems arising through missing and perhaps irregular data. We consider the challenges of: identifying a system model from observational data when measurements are missing, determining control inputs when output data are missing, and ensuring stability when future update times may be missed.

To fix ideas we will root our discussion on a medical example, namely the long term anticoagulation of patients with chronic conditions, using the records of 152 patients from Newcastle upon Tyne, UK. Anticoagulants are prescribed to control the blood clotting speed of patients at risk of thrombosis (blood too easily clots) or severe bleeding (blood does not clot sufficiently quickly). Clotting speed is measured through the International Normalised Ratio (INR), which is a standardised measure with high values indicating long clotting times (Baglin, Keeling, Watson, & the British Committee for Standards in Haematology, 2006). Usually INR is deemed under control if it is in the range 2–3 units. To achieve control the most common anticoagulant is warfarin, which patients can take in various dosages. Raising the dose increases the clotting time and lowering the dose the opposite. Hence, with warfarin dose prescribed at clinic visit k considered as the input $u(k)$, and INR considered as the output $y(k)$, we have a classic control problem. Unfortunately INR is rarely stable, the effect of warfarin differs between patients, and in addition it interacts with many lifestyle factors, including diet, alcohol, exercise and co-medication, meaning that control is not always straightforward.

Fig. 1 exemplifies, using the results from three patients. The upper part of each plot shows the dose of warfarin prescribed, and the lower part shows the measured INR at clinic visits. Patient A is typical of one class of patient. There is considerable noise in the INR values, even when the patient is essentially under control. There are periods where the INR increases, requiring a corresponding decrease in dose, and other periods where the opposite occurs. The need for changing dose frequently to react to changing INR is evident. On the other hand, the second plot shows results for a second class of patient, where there is some instability when the patient is first prescribed warfarin, but then a suitable dose is found and the patient remains stable for a considerable time. Unfortunately, once stability is achieved it is not necessarily maintained, as illustrated in the third plot. Here a patient was well controlled for several years until, for unknown reasons, the INR became highly volatile. Note also that the dose level for stability differs considerably between the three patients: optimum treatment is patient-specific.

This example will be referenced throughout the paper. In Section 2 we will consider briefly the issue of assigning causation in observational data. In Section 3 we will discuss system identification from observational data subject to missingness, explaining some of the ideas and methods that have been developed in the statistical research literature, where there is considerable experience in dealing with these problems. In Section 4 we turn to control in the presence of missing data. We assume that the intention is to sample periodically in time, but that occasional planned sample times are omitted. In Section 5 we turn attention to more irregular and variable sampling rates and concentrate on event-based control. Discussion in Section 6 concludes the paper. A list of abbreviations used throughout the paper is given in Table 1.

2. Causation

Simple causal reasoning about feedback control systems is difficult as the inputs are influenced by the outputs (Åström & Murray, 2010). Therefore, in conventional control system design, the problem is usually split into two steps: i) system identification from open loop input-output experiments and ii) control design and implementation. To identify the system, chosen inputs are applied in open-loop experiments. The term open-loop implies that the response has no influence on the input and so the issue of causation is not considered. By splitting the control design into two steps, the issue of causation in identifying a suitable plant model can effectively be ignored. However, experimental data can sometimes only be obtained in a closed-loop situation, for example when a feedback controller is always required due to safety or stability issues. Hence, methods to identify system models from closed-loop data have been developed within the control systems literature (e.g. Gilson and Van Den Hof, 2005; Van den Hof, 1998; Ljung, 1999; Söderström and Stoica, 1989; Taylor, Young, and Chotai, 2013; Verhaegen, 1993; Young, 1970; 2012, Section 8.7).

A schematic of an open-loop plant used for system identification is

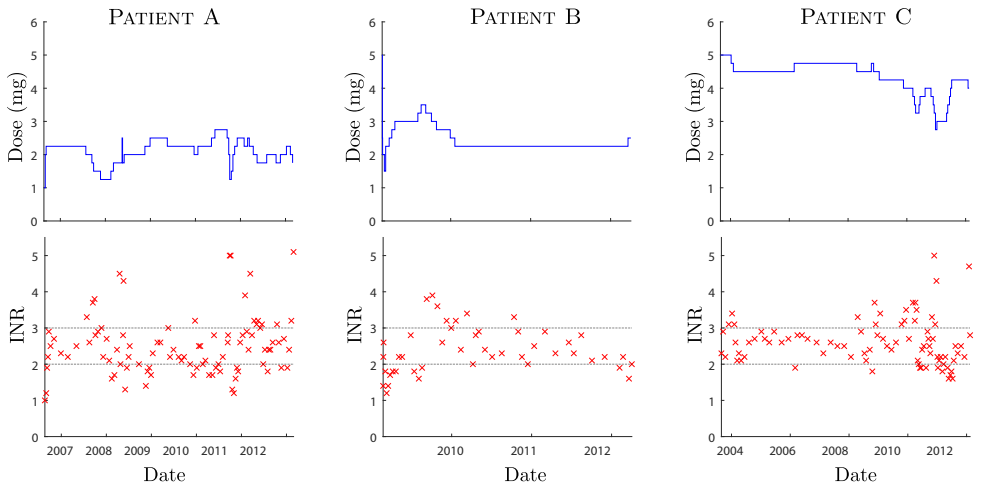

Fig. 1. Anticoagulation of three patients. The three subplots have the same vertical scales but different horizontal scales.

Table 1
Abbreviations used throughout document.

Abbreviation	Definition	Context
AR	Autoregressive	Model structure
ARMA	Autoregressive moving average	Model structure
ARMAX	Autoregressive moving average with exogenous terms	Model structure
ARX	Autoregressive with exogenous terms	Model structure
CC	Complete case	Imputation method
DDE	Delay differential equations	Method to analyse NCSs
EM	Expectation maximisation	Parameter estimation
INR	International normalised ratio	Measure of blood clotting speed
KF	Kalman filter	State estimation
LMI	Linear matrix inequalities	Method to analyse NCSs
LOCF	Last observed carried forward	Imputation method
LQ	Linear quadratic	Optimal control
LQG	Linear quadratic gaussian	Optimal control
MAR	Missing at random	Missingness mechanism
MCAR	Missing completely at random	Missingness mechanism
ML	Maximum likelihood	Parameter estimation
MNAR	Missing not at random	Missingness mechanism
MPC	Model predictive control	Controller
NCS	Network control system	Control system
ODT	Optimal dynamic treatment	Decision rule
PI	Proportional integral	Controller
PID	Proportional integral derivative	Controller
VAR	Visiting at random	Visiting schedule mechanism
VCAR	Visiting completely at random	Visiting schedule mechanism
VNAR	Visiting not at random	Visiting schedule mechanism

presented in Fig. 2a. The experimenter is free to set the input at any of the feasible levels and can identify changes in output caused by changes in input. However, when a plant is to be identified from observational rather than laboratory or experimental data, the problem of interference in input determination can rarely be ignored. The most obvious of these is that in observational data the input is set with the intention of maintaining control, not for the benefit of the observer. Consequently there may be rather little spread in the range of inputs selected, and bad decisions should be rare, which makes system identification more difficult. Further, often the input will be at least partially informed by the previous outputs, as in Fig. 2b.

In the warfarin example a large proportion of high INR values are followed by decreases in dose, and a large proportion of low INR values are followed by an increase in dose. We will model this in the next section and show that, although dose is associated with previous INR, the relationship is not deterministic and there are many occasions where the dose is unchanged despite high or low INR, and a rather smaller number of times where the dose moved in the opposite direction to expectations. Association between output and selected input can be allowed for by careful consideration of direct and indirect effects. Thus an output at time k can influence outputs at time $k+1$ directly, through serial correlation in the response, and indirectly by its effect on the input $u(k)$. Path analysis techniques can be used to disentangle these (Borgan, Fiaccone, Henderson, & Barreto, 2007; Fosen, Ferkingstad, Borgan, & Aalen, 2006).

Exogenous variables may also drive input selection. Continuing with the anticoagulation application, in the hospital notes accompanying the data there are regular explanatory comments on unusual INR possibly being caused by co-medication, increased or decreased alcohol, missed doses, preparation for surgical procedures, and so on. This often meant the change in dose was different from that which might be expected. One example alongside some unhealthily high INR values is the comment “Patient has bad sight and may have been taking 3 mg instead of 2 mg. Request large print”. Another is a comment alongside a sharp decrease then later recovery of INR for one elderly patient: “Nursing home has been giving patient 1–2 glasses of cranberry juice every day!!. Advised of effects of cranberry on warfarin. Home will stop”. The nursing home, like the current authors, was evidently unaware that cranberry can react with warfarin. In both cases the action was to fix the underlying problem and not to change dose to accommodate unusual INR. In these cases, if a reason for a dose change is recorded then it can be allowed for in modelling.

Most often there is no record of exogenous factors or *confounders* that could affect either the output or, importantly, decisions on input level. This is the scenario of Fig. 2c, where both input and output are associated in some way with a disturbance. Since the disturbance is not measured, at least directly, then disentangling the effect of the input from the effect of the disturbance confounder becomes challenging. This leads naturally to issues that have been discussed in various literatures, especially clinical epidemiology and social science, under the name of *causal inference*. Causation in such literature does not mean determining the true physical process, at a molecular level for instance when a drug is administered, but instead reliably attributing an effect to a cause at a more macro level. Thus changing warfarin by this amount will lead, on average, to a change in INR of that amount.

There are many approaches to causal inference, some of which are contentious (e.g. Dawid, 2000; Ding & Li, 2018; Pearl, 2009; Rubin, 2005). The most common is based on the *potential outcomes* or *counterfactual* framework introduced by Rubin (1974), though he acknowledged antecedents as far back as Neyman (1923). Associated with each *potential* input $U(k)$ is the *potential* output $Y(k+1, U(k))$ that would be observed if the actual input $u(k)$ took that value. Causal effects are then defined in terms of comparisons of $Y(k+1, U(k))$ over different $U(k)$. Unfortunately the fundamental problem of causal inference (Holland, 1986) is that only one of the potential outcomes can ever be observed, that corresponding to the selected $u(k)$. Consequently, as Pearl (2009) states, behind every causal conclusion there must lie some causal assumption that is not testable in observational studies.

The most important of many technical assumptions are on the input selection mechanism: how, in the observed data, was the input selected? As stated, if the reasons are known then they can be adjusted for. The problem is when the reasons are not known. For instance, if a clinician prescribes a low dose because they are aware that the patient likes to drink cranberry juice, then an analyst unaware of the additional information will misinterpret the effect of cranberry as the effect of warfarin. A fundamental assumption is then of *no unmeasured confounders*: anything that can influence both the input and the output is recorded in the available data and can be adjusted for in the analysis, at least in principle (e.g. Ding & Li, 2018; Streeter et al., 2017). How realistic such an assumption is will be context-specific. Returning to the

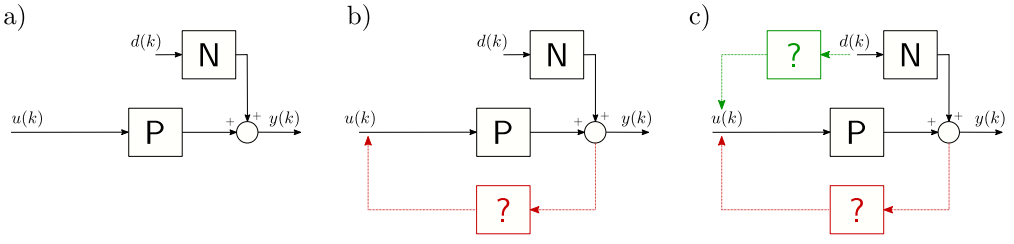

Fig. 2. Measured and unmeasured confounders: (a) open-loop system with disturbance; (b) output as measured confounder; (c) disturbance as unmeasured confounder. Here, $u(k)$ and $y(k)$ represent the input and output respectively, and $d(k)$ is the disturbance; the boxes represent general systems, which may not be known but, in system identification, P is typically the plant and N the noise model.

warfarin example the assumption is certainly questionable, as it relies on all relevant information provided by a patient being recorded in the notes at each clinic visit. Sensitivity methods are recommended to assess the consequences of the assumption failing (e.g. VanderWeele & Onyebuchi, 2011).

3. System identification

Much of the control literature on missing-data, or non-uniform sampling, is concerned with designing stable control systems in the presence of timing uncertainty and missing information (see Section 5). A key concern is the performance of the system in the presence of bounded uncertainties. The approach to this problem tends to be theoretical and system models are either pre-specified (Arzén, 1999; Åström, 2008; Heemels, Sandee, & Van Den Bosch, 2008), or a general state-space system model is considered (Cloosterman et al., 2010; Heemels & Van De Wouw, 2010; Lunze & Lehmann, 2010; Sala, 2005). It is generally assumed that the plant is stable and there are no model uncertainties (Lunze & Lehmann, 2010). The question of how to identify the system model from data with missing measurements is often not considered. It is nonetheless a key issue in designing controllers from observational data.

There are some examples in the control domain where identification from data with missing measurements is considered (see Shi & Fang, 2010 and Ding & Ding, 2010 and the references within). The estimation of autoregressive with exogenous terms (ARX) models (Isaksson, 1993; Wallin, Isaksson, & Ljung, 2000), autoregressive (AR) models (Larsson & Söderström, 2002; Mirsaidi, Fleury, & Oksman, 1997), and autoregressive moving average (ARMA) models (Jones, 1980; Rosen & Porat, 1989) from data with missing, or incomplete measurements has been considered. A general frequency domain approach to system identification with missing data is presented in Pintelon and Schoukens (2000). This approach works for any model structure (e.g. ARX, ARMAX, output-error, etc.) and is based on treating missing input, or output data as unknown parameters. This can potentially lead to a large number of parameters to be estimated. The particular problem of missing output data is considered by Sanchis and Albertos (2002), where the input (control action) is assumed to be updated at a fixed rate, with the output measured synchronously with the input, but with an irregular availability pattern. This is similar to the situation that could be encountered in our anticoagulation example, where the patient continues to take a regular fixed dose (input), but appointments may be missed, so output data are sometimes unavailable.

In the control literature, when modelling missing data, the reason or mechanism for the data being missing is rarely considered. Data are either assumed to be missing periodically (Markovsky, 2013; Rosen & Porat, 1989; Sanchis, Sala, & Albertos, 1997), or a random Bernoulli pattern of missing data is assumed where each measurement has a fixed probability of being missing and misses are independent (Rosen & Porat, 1989; Shi & Fang, 2010).

3.1. Missing data mechanisms

The question of how to develop a model from data that contains missing measurements is a key problem in the statistics literature. Rubin (1976) introduced three important classifications for missing data. These are: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR).

Suppose that a vector of responses Y is intended to be measured. Let R be a vector of observation indicators, with the convention that the elements of R are one when the corresponding elements of Y are observed, and zero if they are missing. Divide Y into sub-vectors of observed and missing components, Y_{obs} and Y_{miss} respectively. Use $f(.)$ as generic notation for probability distributions. Data are said to be:

- *Missing completely at random* (MCAR) if the missingness mechanism is independent of the responses:
  $f(R|Y_{obs}, Y_{miss}) = f(R)$.
- *Missing at random* (MAR) if the missingness mechanism is conditionally independent of the missing data, given the observed data:
  $f(R|Y_{obs}, Y_{miss}) = f(R|Y_{obs})$.
- *Missing not at random* (MNAR) otherwise:
  $f(R|Y_{obs}, Y_{miss}) \neq f(R|Y_{obs})$.

These concepts, and variants thereof (e.g. Mealli & Rubin, 2015; Seaman, Jackson, & J., 2013), underpin much of the statistical research into methods for dealing with missing data. They are not considered in the control literature, where data are generally assumed to be MCAR.

Assuming that the parameters that determine R are distinct from those that determine Y, it is accepted that simply ignoring the missing data reduces efficiency under MCAR (as there is less data to use) but otherwise leads to no difficulty. Some statistical methods, notably maximum likelihood (ML) estimation, remain valid under MAR if the missing data are ignored (Little, Rubin, & Zangeneh, 2017). To see this note that we need to maximise the observed data likelihood. This is obtained by integrating out Y_{miss} from the full data distribution:

$$\begin{aligned} f(Y_{obs}, R) &= \int f(Y_{obs}, Y_{miss}, R) dY_{miss} \\ &= \int f(R|Y_{obs}, Y_{miss}) f(Y_{obs}, Y_{miss}) dY_{miss} \\ &= \int f(R|Y_{obs}) f(Y_{obs}, Y_{miss}) dY_{miss} \quad \text{under MAR} \\ &= f(R|Y_{obs}) \int f(Y_{obs}, Y_{miss}) dY_{miss} \\ &= f(R|Y_{obs}) f(Y_{obs}). \end{aligned}$$

So all parameters in $f(R|Y_{obs})$ and $f(Y_{obs})$ can be estimated consistently. Åström and Torsten (1965) introduced the ML framework to the control community, and Isaksson (1993) implements ML with a Kalman filter to estimate ARX model parameters from missing data.

Data are said to be ignorable in the context of likelihood inference when the parameters describing the measurement process are functionally independent of the parameters describing the missingness process and the data are MCAR or MAR. When these conditions are not met the data are non-ignorable. If data are ignorable, inferences can proceed by analysing the observed data only. Ignoring missing data under MNAR will usually lead to biased and inconsistent estimation. Correct inference needs to take into account the missing data and mechanism, and then depends upon the modelling assumptions all being correct. This is of course a common requirement across all areas of modelling and estimation but a difference in the missing data sphere is that it is *impossible* to verify MNAR assumptions no matter how much data are available: see, for example, the informatively-titled article "Every missingness not at random model has a missingness at random counterpart with equal fit" by Molenberghs, Beunckens, Sotto, and Kenward (2009).

3.2. Longitudinal data

Observational data for control design are likely to be longitudinal. In healthcare applications, longitudinal data (often termed time-series data in control) contain measurements from individuals that are taken repeatedly through time. In longitudinal data, partially observed sequences due to missed appointments, or especially dropout (patient leaves study after a time and there are no subsequent measurements taken) are very common (Diggle, 2002). There is a huge literature on missing data methods in longitudinal studies; see, for example:

Diggle (2002), Molenberghs and Kenward (2007), Ibrahim and Molenberghs (2009), Molenberghs et al. (2004), Ibrahim, Chen, Lipsitz, and Herring (2005), Pigott (2001), and Farewell, Huang, and Didelez (2017) and the references therein.

Simple methods for analysing longitudinal data subject to missingness include complete case (CC) and last observation carried forward (LOCF) techniques. In CC analysis only subjects without missing data are included in the estimation procedure, while under LOCF missing values are simply imputed from the last observed value. Both of these methods have been justifiably criticised but seem still to be in use (National Research Council, 2010).

More principled methods rely on models to at least some extent, with early approaches falling into one of three general classes. Recalling that Y represents the responses of interest and R indicates whether values are observed or missing, writing $f(Y, R) = f(R|Y)f(Y)$ leads to the *selection* model class, in which the marginal distribution of Y is modelled, by a multivariate Gaussian distribution for example, and then the observation indicator is allowed to depend upon Y through an appropriate binary data model. On the other hand writing $f(Y, R) = f(Y|R)f(R)$ leads to the *pattern mixture* class of models based on the marginal distribution of R and then separate models for Y given each possible R. Mathematically the two factorisations are equivalent but they lead naturally to different modelling assumptions. A third general class is also in widespread use, the *shared parameter* approach. Here it is assumed that association between Y and R arises because of shared dependence upon an unobserved random effect W, i.e. $f(Y, R, W) = f(Y|W)f(R|W)f(W)$ (Fitzmaurice, Davidian, Verbeke, and Molenberghs, 2009; Henderson, Diggle, and Dobson, 2000, Chapter 19).

Given the assumed model, estimation techniques include direct likelihood, Bayesian analysis, the expectation-maximisation (EM) procedure, single and multiple imputation and estimating equations (Fitzmaurice et al., 2009, Chapters 15, 17–21). Doubly robust procedures have been developed which attempt to reduce the dependence upon untestable assumptions by creating valid inference *either* if assumptions about R are correct *or* if assumptions about Y are correct, but not necessarily both (Bang & Robins, 2005; Carpenter, Kenward, & Vansteelandt, 2006; Ding & Li, 2018).

3.3. Anticoagulation example

A simulated model based on the warfarin data is developed and used here, and later in the article, to demonstrate various issues with missing data.

The output $y(k)$ is the *log*(INR) and the input $u(k)$ the dose (mg) of warfarin. In the data (9851 records from 152 patients), 98% of dose *changes* are from the set $\{-0.5, -0.25, 0, 0.25, 0.5\}$. We fitted a proportional odds model (McCullagh, 1980) to the observed dose changes and used this to generate changes in inputs in the simulations. The most important factors are the most recent INR (if high tend to reduce dose) and the previous change in dose (tendency to reverse). Table 2 compares the actual changes with those predicted by the model (as the most likely).

For the outputs we used the simple model,

$$y(k) = a_1 y(k-1) + b_1 u(k-1) + \varepsilon. \tag{1}$$

Based on a typical patient, we take $a_1 = 0.4$ and $b_1 = 0.25$, together with $\varepsilon \sim \mathrm{N}(0, \sigma^2)$ where $\sigma = 0.25$. Example input-output data generated using this model are given in Fig. 3: the patterns look very similar to those in the real data. To generate different mechanisms for missingness in the observed data, we used a logistic model for the probability $p(k)$ that $y(k)$ is observed:

$$z(k) = y(k) - a_1 y(k-1) - b_1 u(k-1) \tag{2}$$

$$p(k) = \mathrm{expit}(\theta_0 + \theta_1 z(k-1) + \theta_2 u(k-1) + \theta_3 z(k)) \tag{3}$$

where $\mathrm{expit}(x) = \exp(x)/\{1 + \exp(x)\}$.

To illustrate we chose the parameter values in Table 3 to give 50% missingness under each of the three mechanisms. The non-zero coefficients are large as $z(k)$ is usually small. Parameters were estimated by ML with allowance for correlation between successive $y(k)$. When there were missing data, we estimated using all values k such that $y(k)$, $y(k-1)$ and $u(k-1)$ were observed. Table 4 summarises our results. If there are no missing data, the parameters are estimated well. All means are within the simulated noise of the true values. Under MCAR there is no bias but there is more variability – this is simply because there are less data available for estimation. Under MAR there is no bias because we used ML estimation, which remains consistent if MAR missing data are ignored. Under MNAR there is severe bias in the estimated parameters.

Table 2
Frequencies of observed and predicted dose changes.

		Predicted change				
		−0.5	−0.25	0	0.25	0.5
Actual change	−0.5	25	58	49	5	0
	−0.25	78	418	481	6	1
	0	43	492	5537	619	59
	0.25	3	3	622	448	75
	0.5	0	2	46	70	17

4. Periodic control

This section considers control design and implementation with periodic missing data, i.e. data are sampled periodically but for certain samples the output data are missing. For model-based control, the problem can be split into independent issues of model identification and model-based design. The issue of model identification was summarised in the previous Section 3.3, where an ARX model was fitted to simulated data with missing outputs. In this section we study control design and performance. We first evaluate the performance of controllers designed using models developed from missing data, but where all control data are available, before considering the case when not all the control data are available. We end this section with discussion on optimal dynamic treatment (ODT) techniques and explore how these link to a general control problem.

4.1. Control design using incomplete observations

Controllers developed for non-standard applications tend to be based on PID, PI, or MPC designs. PID controllers are one of the most ubiquitous control solutions, and as expected PI and PID controllers have been applied to the control of non-traditional systems. Examples include: in the control of anaesthesia (Dumont, 2012; Dumont, Martinez, & Ansermino, 2009); for the control of blood glucose using insulin (Hoekstra et al., 2009; Rattan & Nasraway, 2013); and for micro-climate control (El Ghoumari, Tantau, & Serrano, 2005). Due to the ability to handle constraints, which can be numerous in real world problems, MPC methods are also popular, probably more so than PID and PI methods. Examples include: again the control of blood glucose using insulin (Chakrabarty et al., 2017; Hovorka et al., 2004; Plank et al., 2006); the design of adaptive interventions (Rivera et al., 2007); for the solution of decision problems in economics (Grüne, Semmler, & Stieler, 2015); and to control the heart rate (Aerts et al., 2008; Leor-Librach et al., 1999) and growth (Aerts et al., 2003) of animals.

In applying control to non-traditional areas it may be necessary to use observational data for control design. As described in the previous section, this can lead to a biased estimated model when data are missing. To illustrate the effect of this model bias on control system performance, we use the warfarin model of the previous section and

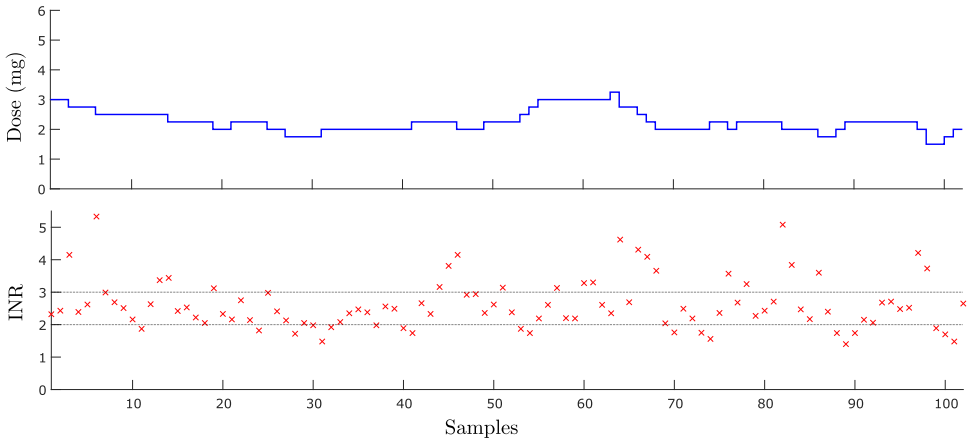

Fig. 3. Example of anticoagulation simulation. The blue lines represent dose in mg of warfarin and the red crosses INR. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Missingness simulation parameter values.

	θ_0	θ_1	θ_2	θ_3
MCAR	0	0	0	0
MAR	0	−10	0	0
MNAR	0	−10	0	10

Table 4
Effect of missingness on parameter estimates. Results of 20,000 simulations, each of sequence length 100. Approximately 50% of the observations are missing under each of MCAR, MAR and MNAR.

	$a_1 = 0.4$		$b_1 = 0.25$		$\sigma = 0.25$	
Missingness	Mean	Std dev	Mean	Std dev	Mean	Std dev
None	0.42	0.07	0.24	0.03	0.25	0.02
MCAR	0.39	0.17	0.26	0.07	0.24	0.04
MAR	0.39	0.23	0.26	0.07	0.24	0.04
MNAR	0.68	0.20	0.22	0.08	0.19	0.04

evaluate the performance of PI and MPC controllers designed using both an accurate system model ($a_1 = 0.4$, $b_1 = 0.25$), and a biased system model ($a_1 = 0.68$, $b_1 = 0.22$), as obtained under MNAR. For the PI control design we used LQ optimisation, as described in Section 5.4 of Taylor et al. (2013), with weights of $W_e = 0.1$, $W_u = 1$, $W_y = 0.5$. For the MPC controller, we used a receding horizon control law with a prediction horizon of three, that was based on minimising the sum of the squares of differences between a reference trajectory and predicted future outputs (based on a discrete state space model). No constraints on the inputs were used, so a simple analytical solution for the optimal inputs could be computed (Camacho & Alba, 2013). In this simple example, offset-free MPC was not used i.e. the MPC control law does not include the equivalent of integral action (Maeder, Borrelli, & Morari, 2009).

Sample sizes of 100 were used and the desired INR output (reference trajectory) defined to be 2.5 for the first 50 samples, then 3.5 for the remaining 50 samples. These are the normal targets for patients with occasional and recurrent deep vein thrombosis respectively (Baglin et al., 2006), with the increase mimicking a change in diagnosis. Performance is evaluated by considering the proportion of INR outputs that are within $\pm$ 0.5 of the desired INR, and performance metrics are the averages from 1000 different trials. Example control responses from one trial are summarised in Fig. 4 and performance metrics given in Table 5. In all cases the noise to signal ratio is high, so none of the methods yield results that are within the desired INR range for more than 50% of samples. Performance is improved when MPC is used with the correct model, however using MPC with an incorrect model results in steady-state errors. This is expected as the performance of MPC is strongly associated with the accuracy of the model. Steady state errors could be removed by, for example, introducing an integral of error state into the MPC rule (Akçakaya & Sümer, 2009; Di Ruscio, 2013; Exadaktylos & Taylor, 2010). MPC techniques that are robust to model mismatch have also been developed (Kothare, Balakrishnan, & Morari, 1996). In practice these methods have rarely been used when MPC has been applied to non-conventional control examples. However, they could increase control performance in cases when the model is estimated from observational data and may be biased. The performance of the PI controller is less affected by model mismatch. Again this is as expected since, although the PI weights were tuned using a specific model, the resultant response is relatively slow and reasonably robust; and inclusion of integral action ensures there are no steady state errors when the model is incorrect.

4.2. Control implementation with missing data

During control implementation there are two distinct approaches to dealing with missing data, either the missing values can be estimated (reconstructed/ imputed) (Phillips & Tomizuka, 1995; Shi & Fang, 2010), or the control decision can be delayed until data are available (Astrom & Bernhardsson, 2002; Heemels et al., 2008). The second of these is an event based control strategy (Åström, 2008) which will be discussed in Section 5, while the first of these is considered here.

The theory for time-driven, periodic systems is well developed. Therefore, when data are missing it is attractive to provide estimates of the outputs at a constant sampling rate, so as to have access to the systems theory for time driven systems (Heemels et al., 2008). A Kalman filter can be used to estimate outputs with uniform sampling period when measurements are missing (Heemels et al., 2008; López-Orozco, de la Cruz, Besada, & Ruipérez, 2000). Similarly, a Leuenberger-type observer has also been used to estimate the state of the plant at uniform times (Heemels et al., 2008; Phillips & Tomizuka, 1995).

Several Kalman filter based methods can be used to estimate the state of a system when data are missing (Khan & Gu, 2009; Sinopoli et al., 2004a). Of these, the simplest method is zero-correction (KF zero), where if no output data are available (e.g. data are missing) the predicted state estimate is used as the true state without updating (Khan & Gu, 2009). This method yields simple, fast estimation without any need to store observations. However, drawbacks include reliance on the system model and spike phenomena in the estimation error. An alternative approach is to use past observations (KF past), where the

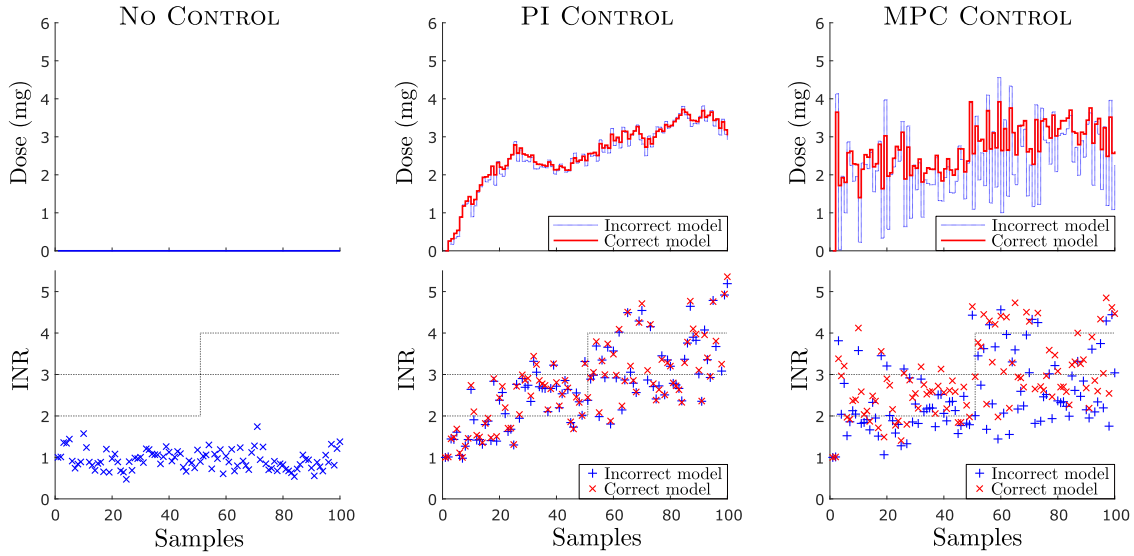

Fig. 4. Example control results. The desired INR is defined as the region between the dashed lines.

Table 5
Control performance averaged across 1000 trials for different control designs.

Control method	Estimated model	In range (%)
None	N/A	0.39
PI	$a_1 = 0.4$, $b_1 = 0.25$	40.39
PI	$a_1 = 0.68$, $b_1 = 0.22$	40.00
MPC	$a_1 = 0.4$, $b_1 = 0.25$	47.35
MPC	$a_1 = 0.68$, $b_1 = 0.22$	38.62

previous observations are used to correct the predicted state (Khan & Gu, 2009). This reduces some of the shortcomings of the zero-correction method, but it requires the most recent observation to be stored. Khan and Gu (2009) have also considered the development of more complex Kalman filter based optimisation algorithms.

Missing data can be even more simply estimated by replacing the missing measurement with the last measured value. This removes the need for Kalman filtering, but requires the previous observation to be stored. This last observed carried forward (LOCF) technique is an *ad hoc* method for inference that assumes data are MAR (Molenberghs et al., 2004). More principled statistical approaches to imputation include direct likelihood, Markov chain Monte Carlo, the EM algorithm and multiple imputation (Ibrahim & Molenberghs, 2009; Molenberghs & Kenward, 2007).

To illustrate, PI control of the warfarin model system is evaluated when approximately 50% of output data are missing and the missingness can be described as MCAR, MAR or MNAR using the method described in Section 3.3. Different methods are used to impute the missing data: KF zero, KF past and LOCF (also, an event based method, as discussed in Section 5.1.2). Results averaged over 1000 trails are given in Table 6. Values for the percentage in range are only given for the measured data, as in practice unmeasured values would be unavailable. A key issue during control is ensuring that the unmeasured values also have reasonable values. Results show that the KF zero imputation method, which is most reliant on the model, provides best performance when the model is correct and worst when it is biased.

4.3. Optimal dynamic treatment

There is a link, so far perhaps not fully exploited, between control methodology and research into optimal dynamic treatment (ODT) techniques developed for personalised medicine. In ODT, also called adaptive strategy or adaptive intervention, medical interventions adjust

Table 6
Control performance when data are missing, averaged across 1000 trails for different control designs.

Estimated model	Method	In range (%)		
		MCAR	MAR	MNAR
$a_1 = 0.4$, $b_1 = 0.25$ (correct)	KF zero	42.32	44.13	41.79
	KF past	41.04	42.83	42.65
	LOCF	40.56	41.78	41.98
$a_1 = 0.68$, $b_1 = 0.22$ (biased)	Event	38.99	40.01	39.03
	KF zero	33.45	35.83	30.07
	KF past	38.42	42.10	40.50
	LOCF	39.54	42.47	41.50
	Event	38.75	39.23	37.19

to changing individual circumstances. For overviews and a range of techniques see, for example: Chakraborty and Moodie (2013), Luedtke and van der Laan (2016), Orellana, Rotnitzky, and Robins (2010), Schulte, Tsiatis, Laber, and Davidian (2014), Zhao, Zeng, Rush, and Kosorok (2012), Zhang, Tsiatis, Laber, and Davidian (2012), and Zhou, Mayer-Hamblett, Khan, and Kosorok (2017). An archetypal situation is the anticoagulation scenario of Section 1, where dosage of warfarin is adjusted to reflect change in INR as a result of new diets, co-medication and so on. Thinking of treatment as input $u(k)$ and an appropriate health measure such as INR as outcome $y(k)$, the generic control and ODT scenarios are one and the same. Despite this, there seems to have been little input from the control community into developments in ODT. This may be because many early applications of ODT, starting with Murphy (2003), concentrated on cases where there were few potential treatment options available, classically simply a choice between two or three possible treatments, and few (again, just two or three) repeat visits. There has been more recent interest in applications in which the treatment can take a large number of values, such as the dose of a drug, and can be considered as effectively continuous, and for chronic conditions in which follow-up is long, again as in the warfarin application (Barrett, Henderson, & Rosthøj, 2014; Henderson, Ansell, & Alsibani, 2010; 2011; Rich, Moodie, & Stephens, 2014; Rosthøj, Keiding, & Schmiegelow, 2012).

The ODT literature has concentrated primarily on modelling, estimation and inference, i.e. system identification as discussed in Section 3. There has been less attention given to subsequent use of the

modelled systems for control, which means large and well established areas of control expertise could and should be brought to bear in personalised medicine; robust control for instance. In the other direction some of the techniques developed in ODT could be taken into control applications, including careful handling of sparse and noisy observational data, and perhaps the use of A-learning techniques borrowed from machine learning and adapted to allow causal inference (Chakraborty & Moodie, 2013; Murphy, 2003; Robins, 2004).

Missing data causes two distinct issues in ODTs. These are introduced by Rosthøj, Henderson, and Barrett (2014). The first of these is the issue of missing data in the past. This refers to the situation where the data set used for estimation of the optimal regime contains missing values. This issue is comparable to the identification of models from missing data (see Section 3) and standard techniques of parameter estimation in the presence of missing covariates or responses in longitudinal data can be used. The second issue is missing data in the future; when calculating the optimal treatment it is assumed that the patient will be seen again at all future time points, so the treatment can be changed if necessary. This second issue is potentially more of a problem and has been given much less consideration within the literature.

To our knowledge the issue of missing data in the future has been considered in the ODT literature only by Rosthøj et al. (2014), who proposed the use of a fixed-dose allocation rule. In their optimal dynamic fixed-dose strategy it is assumed that each patient has visits scheduled for regular times $k = 1, 2, \ldots$ but the possibility that some visits may be missed is allowed for, at least under the assumption that missingness is completely at random. At each time point k a fixed dose strategy assigns the current dose $u(k)$ to the patient for all times within a future observation window. The optimal fixed dose is the dose the patient should stay on to optimise the potential outcome if it is assumed that the dose will not change throughout the window, hence providing protection against the patient missing future visits. It is proposed to recalculate the optimal fixed dose at each visit that does take place. Rosthøj et al. (2014) showed that when there are no missed visits the optimal dynamic fixed dose strategy does not differ considerably from the optimal dynamic strategy. When there are missed visits the optimal dynamic fixed dose strategy outperforms the optimal dynamic strategy. A similar strategy, without missing data, is proposed by van der Laan, Petersen, and Joffe (2005) and Petersen, Deeks, Martin, and van der Laan (2007). At each visit (or time point) a future treatment regime is fully specified, this is updated at subsequent visits. Such a strategy is a form of MPC, something that we return to in the next section.

5. Irregular control

The usual approach to digital control is to sample periodically in time, which is known as periodic, Reimann (Astrom & Bernhardsson, 2002), or single-rate (Hu & Michel, 2000) sampling. The previous sections have considered periodic data with missing outputs. However, in many applications such as medical treatment control, or resource management, scheduling may be irregular and not synchronised with any underlying sampling period. Irregular data are not uniformly sampled and occur when scheduled visits (or measurements) are not uniformly distributed. Data could be constrained to treat the problem as one of periodic missing data, but if a fast underlying time period is used then data may be very sparse; alternatively using a slower time period may give poor resolution of when the actual event occurred. In such cases it may be better to treat the data as irregular.

In our anticoagulation example, Patient A (Fig. 1) has relatively regular observations, with 68% of intervals being within 10% of 7, 14, 21 or 28 days. There are nonetheless some 13% of intervals in excess of one month. Table 7 shows the distribution of gaps between visits. Many patients have more irregular sampling times.

As in the case of missing data, irregular visiting schedules can be categorised into: visiting completely at random (VCAR), visiting at random (VAR) and visiting not at random (VNAR). The term VCAR is used when the visit time is independent of the outcomes, VAR when the visit time is independent of the outcome given previously observed data, and VNAR when the visit time is not independent of the outcome given previously observed data (Pullenayegum & Lim, 2016).

Table 7
Sampling intervals for Patient A. We take intervals to be $\simeq 7$ etc. if within 10%. The maximum interval was 84 days. The balance of intervals are within one month but not close to multiples of 7 days.

	Sampling interval (days)						
	≤ 5	$\simeq 7$	$\simeq 14$	$\simeq 21$	$\simeq 28$	31–60	61–90
Percentage	2%	6%	30%	22%	10%	9%	4%

5.1. Event based control

In an event-based system, the sampling is event-triggered, as opposed to time-triggered. It is the occurrence of an event, rather than the passing of time, that determines when samples should be taken (Arzén, 1999; Åström, 2008; Heemels et al., 2008). Within the literature different terminologies have been used to describe event-based control (Lunze & Lehmann, 2010). These include event-based sampling (Åström, 2008), event-driven sampling (Heemels et al., 2008), Lebesgue sampling (Astrom & Bernhardsson, 2002), aperiodic control (Arzén, 1999), and asynchronous control (Losada, Rubio, & Bencomo, 2015). All describe sampling that is not time triggered, but driven by the occurrence of an event.

In time-driven controllers the focus is on performance, while in event-driven controllers the aim is to balance performance with practical aspects (Heemels et al., 2008) such as the availability of data, and resource utilization. There are several reasons for using event-triggered sampling. These include: (i) the nature of the measurement e.g. the system may use event driven sensors such as encoders (Åström, 2008); (ii) the difficulty in sticking to a time-triggered paradigm, such as in modern distributed control systems (Zhang, Branicky, & Phillips, 2001), or applications like medical treatments; and (iii) resource utilisation, for which event-triggered sampling may reduce the number of control updates necessary (Liu, Wang, He, & Zhou, 2014). Also, interestingly in the context of developing optimal treatment rules in healthcare, event driven control is the dominating control principle in biological systems, including humans (Åström, 2008; Gawthrop, Loram, Lakie, & Gollee, 2011; Loram, Van De Kamp, Gollee, & Gawthrop, 2012).

In event-driven control, the nature of the event can vary, and examples of events could include a signal passing a threshold, or the arrival of a data packet to a node. The arrival of new data, or lack of data, is an example of unintentional event driven control and the passing of a threshold an example of intentional event driven control. Most examples of intentional event driven control consider the event as the signal, or a state, passing a threshold. Research into unintentional event driven control has focused on the area of Networked Control Systems (NCS).

A block diagram of a simplified event based control loop is shown in Fig. 5. The controller operates in open-loop between the events, with feedback actions occurring at events. Periodic systems also operate in open loop between sampling instants. In periodically sampled systems, the standard procedure is to keep the control constant between the sampling instants using a zero-order-hold scheme. Sensor signals are periodically sampled, and a zero order hold used on actuators. In event based systems, update times are not scheduled or known. The way the open-loop signal is generated is key and properties of the system depend on how this signal is generated (Åström, 2008; Cervin & Åström, 2007; Lunze & Lehmann, 2010). In event based control the sampler is replaced by an event generator, and the hold by the control input generator.

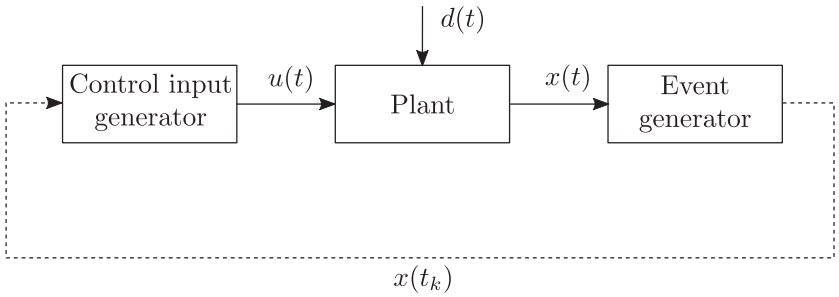

Fig. 5. Block diagram of an event-based control loop (Lunze & Lehmann, 2010). Solid lines denote continuous signal transmission and the dashed lines event based signal transmission. The control input generator may also make use of an observer.

The concept of a control input generator (Fig. 5) was introduced by Åström (2008). It could represent a generalised hold, but is introduced due to the importance of the control signal generator in determining behaviour in event based control. Åström (2008) shows that it is desirable to have holds that give control signals that are initially large, and then decay fast to ensure the system is robust to changes in when the event occurs.

Gawthrop and colleagues (Gawthrop et al., 2011; Gawthrop & Wang, 2007; 2009) use basis functions to specify the controller evolution in the open-loop between events, where the open-loop control signal is constrained to be the linear sum of pre-specified basis functions. The use of basis functions to generate the control signal between events is equivalent to using a generalised hold and, therefore, similar to using the control input generator described by Åström (Åström, 2008; Cervin & Åström, 2007).

5.1.1. Intentional event driven control

In event driven control, the nature of the event can vary. The most commonly studied type of intentional event is the case when a signal, or state, deviates by a set amount. This can be cast, for example, as the measurement signal crossing a specified level, or as the absolute error exceeding a certain amount. The logic for event detection is discretionary, however, having a very complex event detector may needlessly complicate the control algorithm (Arzén, 1999). When the behaviour is event driven, it is not until the event detection criteria has been met (e.g. until the measurement signal has deviated sufficiently from the desired set point) that a new control action is taken.

The difference between an event-based, and a sampled-data controller is considered by Åström and Bernhardsson (Åström, 2008; Åström & Bernhardsson, 1999; Astrom & Bernhardsson, 2002) for a simple system. In this key work, the authors derive analytical results for a simple single-order system, to compare the performance of event-based and sampled-data control loops, where an event occurs when the measurement is out of range. In certain situations the event based strategy performs better in terms of minimising the output variance for similar mean sample times. In the event based strategy communications are increased in time intervals with large disturbances, and decreased when disturbances are small.

Another key early paper on intentional event based control considers an event based PID controller (Arzén, 1999). The event detection logic is kept simple i.e. an event occurs if the absolute value of the difference between the current value of the error $e(t_k)$ and the value of the error the last time a control signal was calculated $e(t_s)$ exceeds a limit, or if a set amount of time has elapsed since the last sample. It is shown through simulations that large reductions in resource utilisation, with only minor control performance degradation, are possible using an event based PID controller. The difficulty is guaranteeing stability and the lack of systems theory for event based control.

Event-based MPC algorithms have also been developed (Bernardini & Bemporad, 2012; Eqtami, Dimarogonas, & Kyriakopoulos, 2011; Henriksson, Quevedo, Peters, Sandberg, & Johansson, 2015). MPC is an optimal control strategy. At each sampling instant (k) a control strategy that minimises an objective function (cost) is computed over a given time horizon ($k + p$). It is standard practice to use only the first computed control input and then resample the plant state and repeat the optimisation starting from this next state. In event triggered MPC, rather than just use the first computed control inputs, the control inputs until the next event are applied. The computed input sequence is applied to the system from time k until time instant $k + i$. At this time instant a new sample is taken and the optimisation is repeated from this state. Chakrabarty et al. (2017) apply event driven MPC to control blood glucose using insulin. Control actions in the optimal MPC input sequence are applied until the event condition is met. Event conditions are: the norm of the output estimation error exceeds a given threshold, or a given number of control actions in the optimal input sequence have been implemented, or safety conditions have been violated. This requires the system to be continually monitored and when the event condition is met, control is triggered (Heemels, Johansson, & Tabuada, 2012). The nature of the event is dependent on the system and method of data collection: for many applications if signals cannot be continually monitored and measured, an event may just be the arrival of new data.

5.1.2. Unintentional event driven control

Aperiodic sampling is not necessarily implemented through choice, instead it can be imposed due to the nature of the system being controlled. We term the case where the limitation is the information exchange and availability of data 'unintentional' event driven control. In our anticoagulation control example, the outputs are only measured when the patient attends an appointment, so it is not possible to continually monitor the system until an event defined by the signals occurs. Instead, we could treat the arrival of new information in itself as an event. This means that the system will not be able to react to increases in the error (e.g. due to set point changes, or disturbances). Decisions on changes to medications can take place only when the patient attends an appointment.

To illustrate, we modified the PI controller described in the simulation study in Section 4 into an event based PI controller using the approach described by Arzén (1999). In divergence from continually monitoring the error, as in Arzén (1999), the event condition is the arrival of data, e.g. $Y_{obs}(k)$ is available. The integral control gain is recalculated at each event by using the elapsed time since the last event to determine the sampling period, $k_I T_s \rightarrow h k_I T_s$, where k_I is the integral control gain, T_s the sample period and set to one in simulations, and h is the number of elapsed samples since the last update. When data are missing the inputs are held. Control simulation results with 50% missing data are given in Table 6, Section 4. The method is not quite as accurate as imputing data, but requires approximately half as many updates, as decisions are only made when data are available.

Event based control shows promise for applications where data are not available at every periodic sampling instant. The key issues are the analysis of the control system and guarantees of worst-case performance. The absence of a systems theory for event driven controllers is a major reason why time driven control still dominates (Heemels et al.,

2008). However, due to the huge increased interest in network control systems, the event driven control literature is rapidly expanding and staring to address these issues, and this includes steps towards developing an event-driven systems theory (Heemels et al., 2012; Heemels et al., 2008; Hetel et al., 2017; Lunze & Lehmann, 2010).

Network control systems provide an example of a system where aperiodic sampling and timing issues are imposed, and a vast amount of literature on aperiodic sampling is concerned with the NCS control problem. An overview of some of the main research themes is provided here, however, it is by no means exhaustive and it should be noted that several reviews of research within the area of NCSs exist (Gupta & Chow, 2010; Heemels & Van De Wouw, 2010; Hespanha, Naghshtabrizi, & Xu, 2007; Hetel et al., 2017; Liu et al., 2014; Tipsuwan & Chow, 2003; Yan, Yan, Zhang, & Zhao, 2014; Yang, 2006; Zhang et al., 2001). The main limitations of research into NCS are that most of the literature only considers linear systems, and most results do not consider plant uncertainty (Yang, 2006).

Unlike sampled data systems, there is no generalised framework for analysing and designing NCSs. Research into NCSs can be grouped according to the type of network-induced uncertainty (e.g. time varying sampling, time varying delay, packet loss), according to the approach used to model and analyse the system under these network induced uncertainties, and how the network induced uncertainty is modelled (Hespanha et al., 2007). Hetel et al. (2017) provide a thorough recent overview on the stability of systems with aperiodic sampling.

Network timing uncertainties are generally modelled in one of two ways (Cloosterman et al., 2010), either by imposing bounds on the delays, sampling intervals, and maximum number of packet dropouts (Cloosterman et al., 2010; Fujioka, 2009b; García-Rivera & Barreiro, 2007), or by using a stochastic modelling approach (Nilsson, Bernhardsson, & Wittenmark, 1996; Nilsson et al., 1998; Sinopoli, Schenato, Franceschetti, Poolla, & Sastry, 2004b). In the first of these two approaches the time variations resulting from network imperfections can be considered as disturbances. Many studies into NCS consider only one or two network induced imperfections, and several approaches have been proposed for modelling and control of the system. These approaches can be broadly categorised as: (i) an input delay approach (Fridman & Shaked, 2005; Gao, Chen, & Lam, 2008; Hespanha et al., 2007; Mirkin, 2007); (ii) a discrete-time approach (Cloosterman, van de Wouw, Heemels, & Nijmeijer, 2006; Cloosterman et al., 2010; Dritsas & Tzes, 2009; Fujioka, 2009a; 2009b; García-Rivera & Barreiro, 2007; Heemels et al., 2008; Sala, 2005); (iii) a model-based approach (Estrada, Lin, & Antsaklis, 2006; Montestruque & Antsaklis, 2004; 2002; 2003); (iv) an emulation approach (Dačić & Nešić, 2007; Nesic & Teel, 2004); and (v) a stochastic optimal control approach (Nilsson et al., 1996; Nilsson et al., 1998; Sinopoli et al., 2004b). Further details are given by Heemels and Van De Wouw (2010) and the previous references, however, the general methodology behind each approach is briefly summarised here.

The *input delay* approach models the NCS as a continuous-time delayed differential equation (DDE). The stability is then studied using Lyaponov-Krasovskii methods (Hespanha et al., 2007), and linear matrix inequalities (LMIs). Mirkin (2007) showed that this approach is conservative as it does not take into account the piecewise constant nature of the control signal that occurs due to the zero-order hold. Alternative DDE approaches (Naghshtabrizi & Hespanha, 2005; Naghshtabrizi, Hespanha, & Teel, 2006; 2008) have been proposed that are less conservative, and do take the piecewise constant nature of the control signal into account. In the *discrete-time* approach, discrete-time representations of the sampled data system are constructed based on the exact discretisation of the continuous-time plant over a sample interval. These models are then used in a robust stability analysis based on Lyaponov-functions (Cloosterman et al., 2010; Fujioka, 2009a; 2009b) or LMIs (Cloosterman et al., 2006; Dritsas & Tzes, 2009; García-Rivera & Barreiro, 2007). In the constrained case, where the delay is smaller than the sampling period, lifted state vectors (Cloosterman et al., 2006; García-Rivera & Barreiro, 2007; Yamamoto, 1996), or Lyaponov–Krasonoskii functions (Xie & Wang, 2004) have been used to address the analysis and design problem.

Model based control of NCSs was introduced by Montestruque and Antsaklis (2004, 2002, 2003). It uses an explicit model of the plant to estimate the plant state between transmission times and provide approximate control signals when sensor data are not available. The model is updated using the true values of the plant state when available. The *emulation* approach considers continuous time controllers using a continuous-time (sampled-data) NCS model in the form of a hybrid system. This is used to quantify allowable levels (to ensure stability) of network uncertainties (maximum transmission intervals, maximum delay). The problem is considered as a linear-quadratic-Gaussian (LQG) problem by the *stochastic optimal control* approach, where the LQG matrix is based on the network delay statistics. This method generally assumes that time delays are less than the sampling period.

6. Discussion

Control theory is increasingly being applied to nonstandard applications. The use of control theory is appealing as it provides a systematic way of determining control inputs in the presence of noise and uncertainty. In such applications, data may be observational, sparse, missing or irregular, with no opportunity to collect repeat measurements. This poses the following challenges: i) system identification from observational data with missing measurements; ii) determining control inputs when output data are missing; and iii) ensuring stability when future update times are unknown. In this paper we reviewed methods in the literature that deal with these issues. The example of anticoagulation control using warfarin was used throughout to illustrate some of the challenges and potential solutions.

Identifying models from observational data is a fundamental issue within the statistics community. To analyse such data, the assumption of no unmeasured confounders is key, which requires that there is knowledge of all predictors of treatment decisions (for medical applications) that are also independent risk factors for the outcome of interest. A significant amount of literature exists on estimating model parameters from longitudinal, observational data with missing covariates; key to this are the assumptions that are made regarding the nature of the data and scheduling, or missing, mechanisms. Traditionally, data collection and analysis assumptions are rarely considered within the context of control. Generally, multiple measurements can be made on the system and the experimental protocol can be designed to enable 'ideal' data to be collected. Where data are missing, it is assumed that the data are either missing periodically, or MCAR. If control systems techniques are to be adopted for nonstandard applications, an awareness of the complexity of the data and any missingness mechanisms is required to avoid introducing bias. Here techniques from statistics could be used to ensure careful handling of sparse and noisy observational data. When estimated models are biased, robust control techniques provide a way of ensuring controllers behave well in the presence of modelling uncertainty, external disturbances and sensor noise.

There are two distinct approaches to handling missing data during control. The first of these is to impute missing data. Several methods for imputing missing data exist. If the imputation method is more reliant on a model of the system, then control performance is more reliant on the accuracy of this model. The second general approach is to use an event based control scheme. Using an 'unintentional' event based scheme would mean control inputs were only updated when data were available, resulting in a system that is unable to react to increases in error. Such a scheme still gives reasonable performance when applied to our anticoagulation simulation example, and required $\approx 50\%$ fewer updates than the imputation method. Alternatively, if the system could be continually monitored, an 'intentional' event based control scheme could be used, where control updates are implemented when a signal,

or state, deviates by a set amount. Event based methods lead to asychronous control, and the rapidly expanding control literature on NCS means a systems theory for asynchronous control is starting to develop. Such theory could usefully be applied to ensure stability when applying control to nonstandard applications.

The dominating control principal within biological systems is event based control. Event based control may, therefore, provide a natural solution for controlling the health of biological systems using medical interventions (as well as in the control of other nonstandard systems). Event-driven control requires the system to be continually monitored and when the event condition is met, control is triggered. The development of sensor technology means certain health variables, such as blood glucose using a continuous glucose monitor, can now be measured online. Where online monitoring is possible, event based control strategies may improve resource utilisation and reduce the number of dose (or treatment) changes necessary. However, in certain contexts, continual health monitoring may be infeasible or impractical. In this case the event triggering condition may just be the arrival of data.

If the system cannot be continually monitored, self driven control, an alternative asynchronous, proactive control strategy could be implemented. In self driven control the time of triggering the next control update is precomputed at the current control update time, based on predictions using the history of the data and knowledge of the plant dynamics (Heemels et al., 2012). A model of the plant dynamics is required, which relates back to the system identification problem. In general, self-triggering results in asynchronous sampling that is dependent on the predicted state of the system. The self driven control problem is closely related to event driven control in terms of ensuring stability given bounds on variations in the sample rate. In the context of medical treatment and ODT, self driven control would provide a method of scheduling future appointments (or measurements), but alone would not provide protection against missed visits.

In protecting against future missing measurements, within the control literature it is generally assumed that there are some bounds on the timing, and methods of designing control systems that are able to deal with given variations have been developed. This contrasts with the statistics approach, which considers that a visit may never occur again (Rosthøj et al., 2014). Again, these contrasting approaches are likely to be due to the nature of the system being controlled. Much of the work on missing and non-uniform sampling in control is within the context of NCS, where it may be reasonable to assume bounds on packet dropouts, or system delays. In the statistics literature, within the context of ODTs, patient scheduling is likely to be less predicable and so within this context it is harder to give upper bounds on the timing between visits. However, in certain contexts, assuming some bounds on the missingness (e.g. patients miss no more than two consecutive appointments, or patient must make at least one visit a month) may help to simplify the control problem and could help to develop stricter rules on the scheduling of the next appointments, or measurements, to ensure constraints are not violated.

An awareness of the context and assumptions imposed by applying control to non-traditional areas is essential. Many methods exist for dealing with missing, or non-uniformly sampled data and the suitability of these methods depends on the actual study being considered. For example, in the case where a patient's health is continually monitored (e.g. blood glucose), event-based control theory may be key for determining treatments. However, it is of no use if we are unable to monitor the event triggering condition. Similarly, an awareness of the limitations of the data is essential in order to understand the limitations and bias that can occur in the estimated models.

Acknowledgement

This work is supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/M015637/1.

References

Aerts, J.-M., Gebruers, F., Van Camp, E., & Berckmans, D. (2008). Controlling horse heart rate as a basis for training improvement. *Computers and Electronics in Agriculture, 64*(1), 78–84.
Aerts, J.-M., Van Buggenhout, S., Vranken, E., Lippens, M., Buyse, J., Decuypere, E., & Berckmans, D. (2003). Active control of the growth trajectory of broiler chickens based on online animal responses. *Poultry Science, 82*(12), 1853–1862.
Akçakaya, H., & Sümer, L. G. (2009). An application of robust model predictive control with integral action. *Instrumentation Science and Technology, 37*(4), 410–430.
Arzén, K.-E. (1999). *A simple event-based PID controller. Proceedings 14th IFAC world congress18. Proceedings 14th IFAC world congress* 423–428.
Åström, K. J. (2008). Event based control. *Analysis and Design of Nonlinear Control Systems, 3*, 127–147.
Åström, K. J., & Bernhardsson, B. (1999). Comparison of periodic and event based sampling for first-order stochastic systems. *IFAC Proceedings Volumes, 32*(2), 5006–5011.
Åstrom, K. J., & Bernhardsson, B. M. (2002). *Comparison of Riemann and Lebesgue sampling for first order stochastic systems. 41st ieee conference on decision and controlvol. 2. 41st ieee conference on decision and control* IEEE2011–2016.
Åström, K. J., & Murray, R. M. (2010). *Feedback systems: An introduction for scientists and engineers.* Princeton university press.
Åström, K.-J., & Torsten, B. (1965). Numerical identification of linear dynamic systems from normal operating records. *IFAC Proceedings Volumes, 2*(2), 96–111.
Baglin, T. P., Keeling, D. M., Watson, H. G., & the British Committee for Standards in Haematology (2006). Guidelines on oral anticoagulation (warfarin): Third edition - 2005 update. *British Journal of Haematology, 132*, 277–285.
Bang, H., & Robins, J. (2005). Doubly robust estimation in missing data and causal inference models. *Biometrics, 61*, 962–973.
Barrett, J., Henderson, R., & Rosthøj, S. (2014). Doubly robust estimation of optimal dynamic treatment regimes. *Statistics in Biosciences, 6*, 244–260.
Behncke, H. (2000). Optimal control of deterministic epidemics. *Optimal control applications and methods, 21*(6), 269–285.
Bekiroglu, K., Lagoa, C., Murphy, S. A., & Lanza, S. T. (2017). Control engineering methods for the design of robust behavioral treatments. *IEEE Transactions on Control Systems Technology, 25*(3), 979–990.
Van den Berg, M., Hegyi, A., De Schutter, B., & Hellendoorn, J. (2003). *A macroscopic traffic flow model for integrated control of freeway and urban traffic networks. 42nd ieee conference on decision and control3. 42nd ieee conference on decision and control* IEEE2774–2779.
Bernardini, D., & Bemporad, A. (2012). Energy-aware robust model predictive control based on noisy wireless sensors. *Automatica, 48*(1), 36–44.
Bibian, S., Ries, C. R., Huzmezan, M., & Dumont, G. (2005). Introduction to automated drug delivery in clinical anesthesia. *European Journal of Control, 11*(6), 535–557.
Borgan, Ø., Fiaccone, R., Henderson, R., & Barreto, M. (2007). Dynamic analysis of recurrent event data with missing observations, with application to infant diarrhoea in Brazil. *Scandinavian Journal of Statistics, 34*, 53–69.
Camacho, E. F., & Alba, C. B. (2013). *Model predictive control.* Springer Science & Business Media.
Carpenter, J., Kenward, M., & Vansteelandt, S. (2006). A comparison of multiple imputation and doubly robust estimation for analyses with missing data. *Journal of the Royal Statistical Society Series A, 169*, 571–584.
Cervin, A., & Åström, K. J. (2007). *On limit cycles in event-based control systems. 46th ieee conference on decision and control*3190–3195.
Chakrabarty, A., Zavitsanou, S., Doyle III, F., & Dassau, E. (2017). Event-triggered model predictive control for embedded artificial pancreas systems. *IEEE Transactions on Biomedical Engineering*.
Chakraborty, B., & Moodie, E. (2013). *Statistical methods for dynamic treatment regimes.* New York: Springer.
Cliff, E., & Vincent, T. (1973). An optimal policy for a fish harvest. *Journal of Optimization Theory and Applications, 12*(5), 485–496.
Cloosterman, M., van de Wouw, N., Heemels, M., & Nijmeijer, H. (2006). *Robust stability of networked control systems with time-varying network-induced delays. 45th ieee conference on decision and control.* IEEE4980–4985.
Cloosterman, M. B., Hetel, L., Van de Wouw, N., Heemels, W., Daafouz, J., & Nijmeijer, H. (2010). Controller synthesis for networked control systems. *Automatica, 46*(10), 1584–1594.
Dačić, D. B., & Nešić, D. (2007). Quadratic stabilization of linear networked control systems via simultaneous protocol and controller design. *Automatica, 43*(7), 1145–1155.
Dawid, A. (2000). Causal inference without counterfactuals. *Journal of the American Statistical Association, 95*, 407–424.
Deshpande, S., Nandola, N. N., Rivera, D. E., & Younger, J. W. (2014). Optimized treatment of fibromyalgia using system identification and hybrid model predictive control. *Control Engineering Practice, 33*, 161–173.
Di Ruscio, D. (2013). Model predictive control with integral action: A simple MPC algorithm. *Modeling, Identification and Control, 34*(3), 119–129.
Diggle, P. (2002). *Analysis of longitudinal data.* Oxford University Press.
Ding, F., & Ding, J. (2010). Least-squares parameter estimation for systems with irregularly missing data. *International Journal of Adaptive Control and Signal Processing, 24*(7), 540–553.
Ding, P., & Li, F. (2018). Causal inference: A missing data perspective. arXiv:1712.06170v2.
Dritsas, L., & Tzes, A. (2009). Robust stability analysis of networked systems with varying delays. *International Journal of Control, 82*(12), 2347–2355.
Dumont, G. A. (2012). Closed-loop control of anesthesia-a review. *IFAC Proceedings*

Volumes, 45(18), 373–378.
Dumont, G. A., Martinez, A., & Ansermino, J. M. (2009). Robust control of depth of anesthesia. *International Journal of Adaptive Control and Signal Processing, 23*(5), 435–454.
El Ghoumari, M., Tantau, H.-J., & Serrano, J. (2005). Non-linear constrained MPC: Real-time implementation of greenhouse air temperature control. *Computers and Electronics in Agriculture, 49*(3), 345–356.
Eqtami, A., Dimarogonas, D. V., & Kyriakopoulos, K. J. (2011). *Novel event-triggered strategies for model predictive controllers. 50th ieee conference on decision and control and european control conference*3392–3397.
Estrada, T., Lin, H., & Antsaklis, P. J. (2006). *Model-based control with intermittent feedback. 14th mediterranean conference on control and automation*. IEEE1–6.
Exadaktylos, V., & Taylor, C. J. (2010). Multi–objective performance optimisation for model predictive control by goal attainment. *International Journal of Control, 83*(7), 1374–1386.
Farewell, D., Huang, C., & Didelez, V. (2017). Ignorability for general longitudinal data. *Biometrika, 104*, 317–326.
Fitzmaurice, G., Davidian, M., Verbeke, G., & Molenberghs, G. (2009). *Handbooks of modern statistical methods: Longitudinal data analysis*. Boca Raton: Chapman and Hall.
Fosen, J., Ferkingstad, E., Borgan, Ø., & Aalen, O. (2006). Dynamic path analysis: A new approach to analyzing time-dependent covariate. *Lifetime Data Analysis, 12*, 143–167.
Fridman, E., & Shaked, U. (2005). Stability and guaranteed cost control of uncertain discrete delay systems. *International Journal of Control, 78*(4), 235–246.
Fujioka, H. (2009a). A discrete-time approach to stability analysis of systems with aperiodic sample-and-hold devices. *IEEE Transactions on Automatic control, 54*(10), 2440–2445.
Fujioka, H. (2009b). Stability analysis of systems with aperiodic sample-and-hold devices. *Automatica, 45*(3), 771–775.
Gaimon, C. (2002). *Optimal control theory: Applications to management science and economics*. JSTOR.
Gao, H., Chen, T., & Lam, J. (2008). A new delay system approach to network-based control. *Automatica, 44*(1), 39–52.
García-Rivera, M., & Barreiro, A. (2007). Analysis of networked control systems with drops and variable delays. *Automatica, 43*(12), 2054–2059.
Gawthrop, P., Loram, I., Lakie, M., & Gollee, H. (2011). Intermittent control: A computational theory of human control. *Biological Cybernetics, 104*(1), 31–51.
Gawthrop, P. J., & Wang, L. (2007). Intermittent model predictive control. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 221*(7), 1007–1018.
Gawthrop, P. J., & Wang, L. (2009). Event-driven intermittent control. *International Journal of Control, 82*(12), 2235–2248.
Gentilini, A., Frei, C. W., Glattfedler, A. H., Morari, M., Sieber, T. J., Wymann, R., et al. (2001). Multitasked closed-loop control in anesthesia. *IEEE Engineering in Medicine and Biology Magazine, 20*(1), 39–53.
Gilson, M., & Van Den Hof, P. (2005). Instrumental variable methods for closed-loop system identification. *Automatica, 41*(2), 241–249.
Grüne, L., Semmler, W., & Stieler, M. (2015). Using nonlinear model predictive control for dynamic decision problems in economics. *Journal of Economic Dynamics and Control, 60*, 112–133.
Gupta, R. A., & Chow, M.-Y. (2010). Networked control system: Overview and research trends. *IEEE Transactions on Industrial Electronics, 57*(7), 2527–2535.
Hansen, E., & Day, T. (2011). Optimal control of epidemics with limited resources. *Journal of Mathematical Biology, 62*(3), 423–451.
Heemels, W., Johansson, K. H., & Tabuada, P. (2012). *An introduction to event-triggered and self-triggered control. Ieee 51st annual conference on decision and control*. IEEE3270–3285.
Heemels, W., Sandee, J., & Van Den Bosch, P. (2008). Analysis of event-driven controllers for linear systems. *International Journal of Control, 81*(4), 571–590.
Heemels, W., & Van De Wouw, N. (2010). *Stability and stabilization of networked control systems. Networked control systems*. Springer203–253.
Henderson, R., Ansell, P., & Alsibani, D. (2010). Regret-regression for optimal dynamic treatment regimes. *Biometrics, 66*, 1192–1201.
Henderson, R., Ansell, P., & Alsibani, D. (2011). Optimal dynamic treatment methods. *Revstat Statistical Journal, 9*, 19–36.
Henderson, R., Diggle, P., & Dobson, A. (2000). Joint modelling of longitudinal measurements and event time data. *Biostatistics, 1*, 465–480.
Henriksson, E., Quevedo, D. E., Peters, E. G., Sandberg, H., & Johansson, K. H. (2015). Multiple-loop self-triggered model predictive control for network scheduling and control. *IEEE Transactions on Control Systems Technology, 23*(6), 2167–2181.
Hespanha, J. P., Naghshtabrizi, P., & Xu, Y. (2007). A survey of recent results in networked control systems. *Proceedings of the IEEE, 95*(1), 138–162.
Hetel, L., Fiter, C., Omran, H., Seuret, A., Fridman, E., Richard, J.-P., & Niculescu, S. I. (2017). Recent developments on the stability of systems with aperiodic sampling: An overview. *Automatica, 76*, 309–335.
Hoekstra, M., Vogelzang, M., Verbitskiy, E., & Nijsten, M. W. (2009). Health technology assessment review: Computerized glucose regulation in the intensive care unit-how to create artificial control. *Critical Care, 13*(5), 223.
Van den Hof, P. (1998). Closed-loop issues in system identification. *Annual reviews in control, 22*, 173–186.
Holland, P. W. (1986). Statistics and causal inference (with discussion). *Journal of the American Statistical Association*, 945–970.
Hovorka, R., Canonico, V., Chassin, L. J., Haueter, U., Massi-Benedetti, M., Federici, M. O., et al. (2004). Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiological Measurement, 25*(4), 905.
Hu, B., & Michel, A. N. (2000). Stability analysis of digital feedback control systems with time-varying sampling periods. *Automatica, 36*(6), 897–905.
Hunt, K. J., & Liu, M. (2018). Optimal control of heart rate during treadmill exercise. *Optimal Control Applications and Methods, 39*(2), 503–518.
Ibrahim, J. G., Chen, M.-H., Lipsitz, S. R., & Herring, A. H. (2005). Missing-data methods for generalized linear models: A comparative review. *Journal of the American Statistical Association, 100*(469), 332–346.
Ibrahim, J. G., & Molenberghs, G. (2009). Missing data methods in longitudinal studies: A review. *Test, 18*(1), 1–43.
Isaksson, A. J. (1993). Identification of ARX-models subject to missing data. *IEEE Transactions on Automatic Control, 38*(5), 813–819.
Jarvis, A., Leedal, D., Taylor, C. J., & Young, P. (2009). Stabilizing global mean surface temperature: A feedback control perspective. *Environmental Modelling & Software, 24*(5), 665–674.
Jones, R. H. (1980). Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics, 22*(3), 389–395.
Khan, N., & Gu, D.-W. (2009). *State estimation in the case of loss of observations. Iccas-sice*. IEEE1840–1845.
Kothare, M. V., Balakrishnan, V., & Morari, M. (1996). Robust constrained model predictive control using linear matrix inequalities. *Automatica, 32*(10), 1361–1379.
van der Laan, M., Petersen, M., & Joffe, M. (2005). History-adjusted marginal structural models and statically-optimal dynamic treatment regimens. *International Journal of Biostatistics, 1*.
Larsson, E. K., & Söderström, T. (2002). Identification of continuous-time AR processes from unevenly sampled data. *Automatica, 38*(4), 709–718.
Leor-Librach, R. J., Bobrovsky, B.-Z., Eliash, S., & Kaplinsky, E. (1999). Computer-controlled heart rate increase by isoproterenol infusion: Mathematical modeling of the system. *American Journal of Physiology-Heart and Circulatory Physiology, 277*(4), H1478–H1483.
Little, R., Rubin, D., & Zangeneh, S. (2017). Conditions for ignoring the missing-data mechanism in likelihood inferences for parameter subsets. *Journal of the American Statistical Association, 112*, 314–320.
Liu, Q., Wang, Z., He, X., & Zhou, D. (2014). A survey of event-based strategies on control and estimation. *Systems Science & Control Engineering: An Open Access Journal, 2*(1), 90–97.
Ljung, L. (1999). *System identification: Theory for the user* (2nd). Prentice Hall, Upper Saddle River, NJ.
López-Orozco, J., de la Cruz, J. M., Besada, E., & Ruipérez, P. (2000). An asynchronous, robust, and distributed multisensor fusion system for mobile robots. *The International Journal of Robotics Research, 19*(10), 914–932.
Loram, I. D., Van De Kamp, C., Gollee, H., & Gawthrop, P. J. (2012). Identification of intermittent control in man and machine. *Journal of The Royal Society Interface, 9*(74), 2070–2084.
Losada, M. G., Rubio, F. R., & Bencomo, S. D. (2015). *Asynchronous control for networked systems*. Springer.
Luan, X., Shi, Y., & Liu, F. (2012). Unscented kalman filtering for greenhouse climate control systems with missing measurement. *International Journal of Innovative Computing, Information and Control, 8*(3), 2173–2180.
Luedtke, A., & van der Laan, M. (2016). Super-learning of an optimal dynamic treatment rule. *International Journal of Biostatistics, 12*.
Lunze, J., & Lehmann, D. (2010). A state-feedback approach to event-based control. *Automatica, 46*(1), 211–215.
MacMartin, D. G., Kravitz, B., Keith, D. W., & Jarvis, A. (2014). Dynamics of the coupled human–climate system resulting from closed-loop control of solar geoengineering. *Climate dynamics, 43*(1–2), 243–258.
Maeder, U., Borrelli, F., & Morari, M. (2009). Linear offset-free model predictive control. *Automatica, 45*(10), 2214–2222.
Markovsky, I. (2013). *Exact system identification with missing data. 52nd ieee annual conference on decision and control*. IEEE151–155.
McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society Series B, 42*, 109–142.
Mealli, F., & Rubin, D. (2015). Clarifying missing at random and related definitions, and implications when coupled with exchangeability. *Biometrika, 102*, 995–1000.
Mirkin, L. (2007). Some remarks on the use of time-varying delay to model sample-and-hold circuits. *IEEE Transactions on Automatic Control, 52*(6), 1109–1112.
Mirsaidi, S., Fleury, G. A., & Oksman, J. (1997). LMS-Like AR modeling in the case of missing observations. *IEEE Transactions on Signal Processing, 45*(6), 1574–1583.
Molenberghs, G., Beunckens, C., Sotto, C., & Kenward, M. (2009). Every missingness not at random model has a missingness at random counterpart with equal fit. *Journal of the Royal Statistical Society Series B, 70*, 371–388.
Molenberghs, G., & Kenward, M. (2007). *Missing data in clinical studies*. Chichester, UK: John Wiley & Sons.
Molenberghs, G., Thijs, H., Jansen, I., Beunckens, C., Kenward, M. G., Mallinckrodt, C., & Carroll, R. J. (2004). Analyzing incomplete longitudinal clinical trial data. *Biostatistics, 5*(3), 445–464.
Montestruque, L. A., & Antsaklis, P. (2004). Stability of model-based networked control systems with time-varying transmission times. *IEEE Transactions on Automatic Control, 49*(9), 1562–1572.
Montestruque, L. A., & Antsaklis, P. J. (2002). *Model-based networked control systems: Necessary and sufficient conditions for stability. 10th mediterranean conference on control and automation24*.
Montestruque, L. A., & Antsaklis, P. J. (2003). On the model-based control of networked systems. *Automatica, 39*(10), 1837–1843.
Murphy, S. A. (2003). Optimal dynamic treatment regimes. *Journal of the Royal Statistical Society Series B, 65*, 331–355.
Naghshtabrizi, P., & Hespanha, J. P. (2005). *Designing an observer-based controller for a network control system. 44th ieee conference on decision and control, and european control conference*. IEEE848–853.

Naghshtabrizi, P., Hespanha, J. P., & Teel, A. R. (2006). *On the robust stability and stabilization of sampled-data systems: A hybrid system approach. 45th ieee conference on decision and control*. IEEE4873–4878.

Naghshtabrizi, P., Hespanha, J. P., & Teel, A. R. (2008). Exponential stability of impulsive systems with application to uncertain sampled-data systems. *Systems & Control Letters, 57*(5), 378–385.

Nandola, N. N., & Rivera, D. E. (2013). An improved formulation of hybrid model predictive control with application to production-inventory systems. *IEEE Transactions on Control Systems Technology, 21*(1), 121–135.

National Research Council (2010). *The prevention and treatment of missing data in clinical trials.* Washington DC: National Academies.

Nesic, D., & Teel, A. R. (2004). Input-output stability properties of networked control systems. *IEEE Transactions on Automatic Control, 49*(10), 1650–1667.

Neyman, J. (1923). On the application of probability theory to agricultural experiments. essay on principles. section 9. (in polish. translation published in 1990). *Annals of Agricultural Sciences (translation in Statistical Science),* 1–51.

Nicklow, J. W. (2000). Discrete-time optimal control for water resources engineering and management. *Water International, 25*(1), 89–95.

Nilsson, J., Bernhardsson, B., & Wittenmark, B. (1996). Stochastic analysis and control of real-time systems with random time delays. *IFAC Proceedings Volumes, 29*(1), 7106–7111.

Nilsson, J., et al. (1998). *Real-time control systems with delays.*. Lund institute of Technology, Sweden Ph.D. thesis.

Nowzari, C., Preciado, V. M., & Pappas, G. J. (2016). Analysis and control of epidemics: A survey of spreading processes on complex networks. *IEEE Control Systems, 36*(1), 26–46.

Orellana, L., Rotnitzky, A., & Robins, J. (2010). Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes. part I: Main content. *International Journal of Biostatistics, https://doi.org/10.2202/1557-4679.1200*.

Paré, P. E., Walker, B. J., & McGrath, J. (2017). *Feedback linearization control methods for accurate leaf photosynthesis measurements. American control conference*. IEEE801–806.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys, 3*, 96–146.

Petersen, M., Deeks, S., Martin, J., & van der Laan, M. (2007). History-adjusted marginal structural models for estimating time-varying effect modification. *American Journal of Epidemiology, 166*, 985–993.

Phillips, A. M., & Tomizuka, M. (1995). *Multirate estimation and control under time-varying data sampling with applications to information storage devices. American control conference6. American control conference* IEEE4151–4155.

Pigott, T. D. (2001). A review of methods for missing data. *Educational Research and Evaluation, 7*(4), 353–383.

Pintelon, R., & Schoukens, J. (2000). Frequency domain system identification with missing data. *IEEE Transactions on Automatic Control, 45*(2), 364–369.

Plank, J., Blaha, J., Cordingley, J., Wilinska, M. E., Chassin, L. J., Morgan, C., et al. (2006). Multicentric, randomized, controlled trial to evaluate blood glucose control by the model predictive control algorithm versus routine glucose management protocols in intensive care unit patients. *Diabetes Care, 29*(2), 271–276.

Pullenayegum, E. M., & Lim, L. S. (2016). Longitudinal data subject to irregular observation: A review of methods with a focus on visit processes, assumptions, and study design. *Statistical Methods in Medical Research, 25*(6), 2992–3014.

Rattan, R., & Nasraway, S. A. (2013). The future is now: Software-guided intensive insulin therapy in the critically ill. *Journal of Diabetes Science and Technology, 7*(2), 548–554.

Rich, B., Moodie, E., & Stephens, D. (2014). Simulating sequential multiple assignment randomized trials to generate optimal personalized warfarin dosing strategies. *Clinical Trials, 11*, 435–444.

Rivera, D. E., Pew, M. D., & Collins, L. M. (2007). Using engineering control principles to inform the design of adaptive interventions: A conceptual introduction. *Drug & Alcohol Dependence, 88*, S31–S40.

Robins, J. M. (2004). Optimal structural nested models for optimal sequential decisions. In D. Y. Lin, & P. Heagerty (Eds.). *Proceedings of the second symposium on biostatistics* (pp. 189–326). New York: Springer.

Rosen, Y., & Porat, B. (1989). Optimal ARMA parameter estimation based on the sample covariances for data with missing observations. *IEEE Transactions on Information Theory, 35*(2), 342–349.

Rosthøj, S., Henderson, R., & Barrett, J. K. (2014). Optimal dynamic treatment strategies with protection against missed decision points. *Statistics in Biosciences, 6*(2), 261–289.

Rosthøj, S., Keiding, N., & Schmiegelow, K. (2012). Estimation of dynamic treatment strategies for maintenance therapy of children with acute lymphoblastic leukemia: An application of history-adjusted marginal structural models. *Statistics in Medicine, 31*, 470–488.

Rubin, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66*, 688–701.

Rubin, D. (2005). Causal inference using potential outcomes. *Journal of the American Statistical Society, 100*, 322–331.

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581–592.

Sala, A. (2005). Computer control under time-varying sampling period: An LMI gridding approach. *Automatica, 41*(12), 2077–2082.

Sanchis, R., & Albertos, P. (2002). Recursive identification under scarce measurements – convergence analysis. *Automatica, 38*(3), 535–544.

Sanchis, R., Sala, A., & Albertos, P. (1997). Scarce data operating conditions: Process model identification. *IFAC Proceedings Volumes, 30*(11), 453–458.

Schulte, P., Tsiatis, A., Laber, E., & Davidian, M. (2014). Q- and A-learning methods for estimating optimal dynamic treatment regimes. *Statistical Science, 69*, 640–661.

Seaman, S., Galati, J., Jackson, D., & Carlin, J. (2013). What is meant by "missing at random"? *Statistical Science, 28*, 257–268.

Shi, Y., & Fang, H. (2010). Kalman filter-based identification for systems with randomly missing measurements in a network environment. *International Journal of Control, 83*(3), 538–551.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M. I., & Sastry, S. S. (Schenato, Franceschetti, Poolla, Jordan, Sastry, 2004a). Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control, 49*(9), 1453–1464.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., & Sastry, S. S. (Schenato, Franceschetti, Poolla, Sastry, 2004b). *Time varying optimal control with packet losses. 43rd ieee conference on decision and control2. 43rd ieee conference on decision and control* IEEE1938–1943.

Söderström, T., & Stoica, P. (1989). *System identification.* Prentice-Hall.

Soldatenko, S., & Yusupov, R. (2017). On the application of optimal control theory to climate engineering. arXiv:1709.05597.

Streeter, A., Lin, N., Crathorne, L., Haasova, M., Hyde, C., Melzer, D., & Henley, W. (2017). Adjusting for unmeasured confounding in nonrandomized longitudinal studies: A methodological review. *Journal of Clinical Epidemiology, 87*, 23–34.

Taylor, C. J., & Aerts, J.-M. (2014). Control of nonlinear biological systems by non-minimal state variable feedback. *Statistics in Biosciences, 6*(2), 290–313.

Taylor, C. J., Young, P. C., & Chotai, A. (2013). *True digital control: Statistical modelling and non-minimal state space design.* John Wiley & Sons.

Tipsuwan, Y., & Chow, M.-Y. (2003). Control methodologies in networked control systems. *Control Engineering Practice, 11*(10), 1099–1111.

Tsitsimpelis, I., Wolfenden, I., & Taylor, C. J. (2016). Development of a grow–cell test facility for research into sustainable controlled–environment agriculture. *Biosystems Engineering, 150*, 40–53.

Underwood, C. P. (2002). *HVAC Control systems: Modelling, analysis and design.* Routledge.

VanderWeele, T., & Onyebuchi, A. (2011). Unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology, 22*, 42–52.

Verhaegen, M. (1993). Application of a subspace model identification technique to identify lti systems operating in closed-loop. *Automatica, 29*(4), 1027–1040.

Wallin, R., Isaksson, A. J., & Ljung, L. (2000). *An iterative method for identification of arx models from incomplete data. 39th ieee conference on decision and controlvol. 1. 39th ieee conference on decision and control* IEEE203–208.

Westcott, J. (1984). Control engineering and economic modelling: A collaboration aimed at improving control of the national economy. *Electronics & Power, 30*(2), 129–134.

Wilson, E. D., Assaf, T., Pearson, M. J., Rossiter, J. M., Dean, P., Anderson, S. R., & Porrill, J. (2015). Biohybrid control of general linear systems using the adaptive filter model of cerebellum. *Frontiers in Neurorobotics, 9*, 5.

Xie, G., & Wang, L. (2004). *Stabilization of networked control systems with time-varying network-induced delay. 43rd ieee conference on decision and control4. 43rd ieee conference on decision and control* IEEE3551–3556.

Yamamoto, Y. (1996). A retrospective view on sampled-data control systems. *CWI Quarterly, 9*(3), 261–276.

Yan, H., Yan, S., Zhang, H., & Zhao, X. (2014). An overview of networked control of complex dynamic systems. *Mathematical Problems in Engineering*.

Yang, T. C. (2006). Networked control system: A brief survey. *IEE Proceedings: Control Theory and Applications, 153*(4), 403–412.

Young, P. C. (1970). An instrumental variable method for real-time identification of a noisy process. *Automatica, 6*(2), 271–287.

Young, P. C. (2012). *Recursive estimation and time[HYPHEN]series analysis: An introduction to the student and practitioner.* Springer.

Zhang, B., Tsiatis, A., Laber, E., & Davidian, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics, 68*, 1010–1018.

Zhang, W., Branicky, M. S., & Phillips, S. M. (2001). Stability of networked control systems. *IEEE Control Systems, 21*(1), 84–99.

Zhao, Y., Zeng, D., Rush, A., & Kosorok, M. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association, 107*, 1106–1118.

Zhou, X., Mayer-Hamblett, N., Khan, U., & Kosorok, M. (2017). Residual weighted learning for estimating individualized treatment rules. *Journal of the American Statistical Association, 112*, 169–187.