A Systems Dynamics Representation of Resilience

Robert L Wears[1,2] and Shawna J Perry[3]
[1]Clinical Safety Research Unit, Imperial College London, UK
wears@ufl.edu; r.wears@imperial.ac.uk
[2]University of Florida, Jacksonville, FL, USA
[3]Virginia Commonwealth University, Richmond, VA, USA
sperry4@mcvh-vcu.edu

Abstract. Resilience can be thought of as a property of a system that permits it to survive and achieve its goals in the face of expected threats and challenges to its operations. Systems dynamics modelling is a technique useful in exploring the behaviour of complex systems, especially the nonlinear interactions and feedback delays are present. In this paper, we use systems dynamics modelling to explore the nature of resilience, using small, highly abstract modules built for incorporation into a larger model of the crowding problem in emergency departments.

1 INTRODUCTION

Emergency departments (EDs) are dynamic, open, high risk, continuously operating systems that demonstrate considerable resilient capacity (Wears & Perry, 2006; Wears, Perry, & McFauls, 2007), but occasionally perform in less resilient, more brittle ways (Anders, Woods, Wears, Perry, & Patterson, 2006). Systems dynamics is a family of techniques for representing and exploring the behaviour of complex systems and their response to change over time (Sterman, 2000a). The objective of this paper is to use systems dynamics modelling of the problem of ED overcrowding to explore the nature of the transitions from resilience to brittleness and back again.

Most US EDs have experienced severe and increasing over-crowding problems over the past decade (Derlet, Richards, & Kravitz, 2001; Goldberg, 2000; Richardson, Asplin, & Lowe, 2002). This is thought to be due primarily to a decrease in the total number of inpatient beds *via* hospital closures, mergers and acquisitions (although there are many other causal influences), leading to the 'boarding' of large numbers of admitted patients in the ED. EDs have adapted to this problem in a variety of ways, such as dedicating entire units to inpatients, adapting previously unused space such as hallways to use as treatment spaces, and dynamically changing the manner in which work is performed (Wears & Perry, 2006). These adaptations create a series of reverberations throughout the organisation that eventually feed back to affect the ED, although subject to various time delays. As the over-crowding problem has increased in severity, this adaptive capacity has become increasingly strained, and a highly respected study of the problem has concluded that EDs as a whole are near a point of complete breakdown (Committee on the Future of Emergency Care in the US, 2006).

Because of time delays in feedback and complex interactions with the rest of the hospital, the effect of current or proposed future strategies to maintain safe ED operations is difficult to determine; in fact, some of the proposed solutions may even be making the problem worse in the long run. The problem has largely been viewed as intractable, and has resisted many attempts at solution or mitigation.

This paper stems from a larger project to model the overcrowding problem and potential approaches to it in an attempt to provide policy guidance to organizations and managers. In this paper, we examine small, highly abstract modules that will be linked together as components of the larger ED model. The objective here is to characterize the sorts of model behaviours that might represent resilience or its converse, brittleness. Identifying these behaviours in very simple abstract models will be an important aid to assessing them in the more complex, ‘full ED’ that is currently under construction.

2 METHODS

In this section we describe the work system under consideration, then discuss the philosphy guiding model development, and finally describe the components of the model used for this analysis.

2.1 Work System

An urban 653 bed US teaching hospital that is part of an 8 hospital network served as the data source. The ED has roughly 100,000 visits per year, and is a Level 1 trauma center. It is divided into 5 major treatment areas totaling 79 beds; 2 treatment areas are dedicated to severe trauma patients and to pediatric cases. One of the non-dedicated treatment areas (comprising 22 beds) is reserved for ‘boarders.’ Two large hallways are routinely used as additional treatment space.

EDs are staffed by three distinct groups – physicians, nurses, and technicians – who have a strong sense of professional identity and a distinct sense of a gradient in authority. (Other groups also work in the ED but typically do not self-identify as ED practitioners, do not work there exclusively). These groups must coordinate their work, but act in highly independent manners, at a ‘cooperative distance;’ coordination among workers is largely implicit, mediated in part by external artefacts such as the status board (Wears, Perry, Wilson, Galliers, & Fone, 2007), synchronous and asynchronous communication, and cross-monitoring

2.2 Guiding Principles for Model Development

Many different modeling disciplines for this problem are available, and even within a single discipline there are a large number of choices to be made.

Balancing Scope and Detail. In evaluating this tradeoff, we decided to favor broad model boundaries over more extensive detail. A broader scope helps to avoid the problem of tacitly mistaking endogenous factors for external causes. We therefore bounded the model at the organizational level (the hospital), rather than at the departmental, or departmental unit level, because we wanted to explore the possible feedback relationships between the ED and the hospital.

With a broad scope, attempting to model fine-grained detail would become unmanageable, and in addition would limit the generalizability of the results.

Because two important goals of the project are to illuminate some aspects of resilience in many, not just this, ED; and even more broadly, to tease out some aspects of resilience in complex work systems in general, we purposely abstracted much of the underlying detail into simpler and hopefully more general model structures. For example, patients differ greatly in the amount of resources and effort they require of the ED, but we treat them as uniform; ED work is pulsatile, responding to at least 4 temporal cycles (daily, weekly, and seasonal cycles of visits and acuity, and weekly cycles of hospital bed availability), but we focus here on the averages. All of these assumptions will require empirical inspection and / or sensitivity analysis.

The temporal scope of the model is more limited than is typical for most systems dynamics models. We limit the temporal scope to the span of control typically wielded by ED operations and hospital middle managers, *ie*, to dealing with problems of flow and crowding in a span of days to weeks, not months to years; this essentially limits responses to reallocations of existing resources and priorities. Thus, increasing capacity by building new space, or hiring additional staff, are bounded out of this specific model because the time course for these actions is too long. Strategies at this level will be explored in related models but not dealt with further here.

Model Choice. The foregoing considerations led us to a generic modeling type, the compartment aggregation model, in which various states of the process are considered separate but communicating compartments; flows and levels are modeled, but not individual agents or entities. Two particular types of compartment aggregation models, aging chain models and supply chain models (Sterman, 2000b) are both well understood and particularly suitable for this problem. In an aging chain, material (or information) flows through a series of compartments in which it typically is delayed; the output of one compartment is the input into the next. Thus an aging chain model lends itself nicely to the logical progress of patients through various stages of the ED and on into the hospital. However, patients can also flow backward, or be recycled through the chain, so supply chain models, which allow for 're-work', are also attractive. One salient difference between the ED and supply chain models is that services cannot be provided in advance or excess and banked in inventories in the same way that 'widgets' can. Thus the ED model will be a hybrid of the classic aging chain and supply chain models (Orcun, Uzsoy, & Kempf, 2006).

Data. The model development process is continuing, and involves direct observations of ED operations, with special attention to 'limit cases' – situations of extreme congestion or overload. Quantitative data on numbers of patients, triage acuity, times of presentation and disposition are obtained from the hospitals computerized information system. Information on system performance was gathered from interviews with nursing, physician, ancillary and management staff in various departments of the organization (ie, not just the ED).

2.3 The Model and the Abstract Modules

A example model for studying the overcrowding problem is shown in Figure 1. It is a hypbrid of classical aging and supply chain models. ED patients present for care, undergo some processing (with delays), and eventually are admitted to the hospital or discharged. If admitted, they may be physically boarded in the ED or transported to an inpatient ward, depending on available inpatient beds. Poor discharge decisions may lead to ‘rework’; patients returning to a previous stage (by presenting again to the ED).

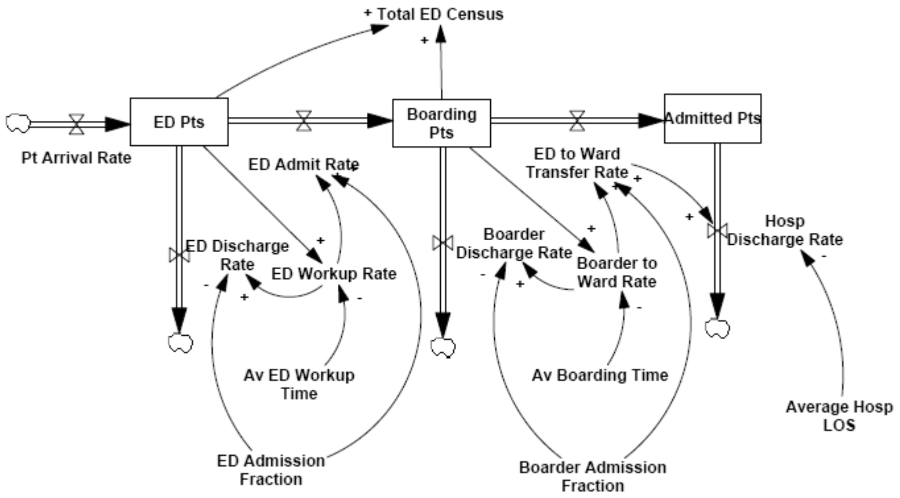

Fig. 1. Schematic model of patient flow through the ED and processing delays

Each of the ‘processing units’ shown in Figure 1 has a substructure (shown in Figure 2) and it is the general substructure and behaviour of those units that is the focus of this paper. These modules are highly simplified, capacitated, input-output units. Patients arrive at some rate, are processed, and leave at a rate that depends not only on the number of patients to be processed, but also on the relative proportions of patients to available resources, and on adaptations to workload by actors in the system.

2.4 Conjectures

If the model is to be useful, we postulate (hypothesize) that is should show several behaviours that seem characteristic of resilience.

Conjecture 1: Non-Adaptive Systems Show Brittleness. The simple modules developed as components for the larger system model should show some sensitivity to a set of inputs that produces sudden and dramatic changes in state.

Conjecture 2: Adding Adaptive Components Mitigate Brittleness. The addition of a capacity that adjusts to exogenous shocks should mitigate this brittleness.

Conjecture 3: Repeated Shocks Plus Adaptive Components Lead to ‘Anticipatory’ Compensation. If exogenous shocks are recurrent, and if the memory of adaptations is sufficiently strong, systems will migrate towards the adapted state and can respond more quickly to exogenous shocks.

Conjecture 4: Brittleness transmits, but resilience contains, exogenous shocks. When systems are arranged in chains, transmission of exogenous shocks along chains suggests brittleness, while their isolation in a small number of modules (ideally one) suggests resilience.

It is important to note here that overall performance of the ED model is likely to be emergent; *ie*, that resilient (or brittle) behaviour of the model as a whole need not necessarily be found in any of its components, and conversely that brittleness (or resilience) at the component level may not necessarily imply that it exists for the entire model. The purpose here is to gain a better understanding of how those behaviors arise and how they might appear in the outputs by using the simplest possible modules.

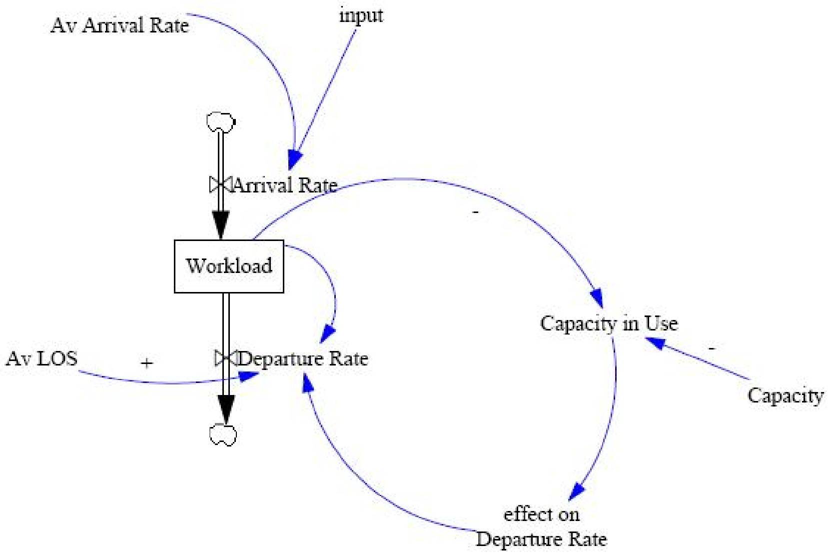

Fig. 2. Simple input-output component of ED model. This is a capacitated processing model; patients present at some rate (which may be modified by external shocks), undergo processing, and depart at a rate which is dependent on the number of patients and the capacity of the unit

3 RESULTS

The highly abstract modules shown in Figure 2 were simulated using historical data for on of the units of the ED, to evaluate the first three conjectures. Using historical data, the module exhibits steady-state performance (not shown) in that the number of patients in the unit stabilizes around 32, and inputs and outputs are balanced.

We first evaluated conjecture 1, that the system could show brittleness under certain conditions when adaptation was limited. Figure 3 shows this behaviour. When shocked by a single input pulse, system response increases to compensate, but eventually the system becomes overloaded and crashes catastrophically, even though the initial pulse was limited in time and input had returned to normal. It is interesting to note that this sudden degradation occurred at some time removed from the initial insult.

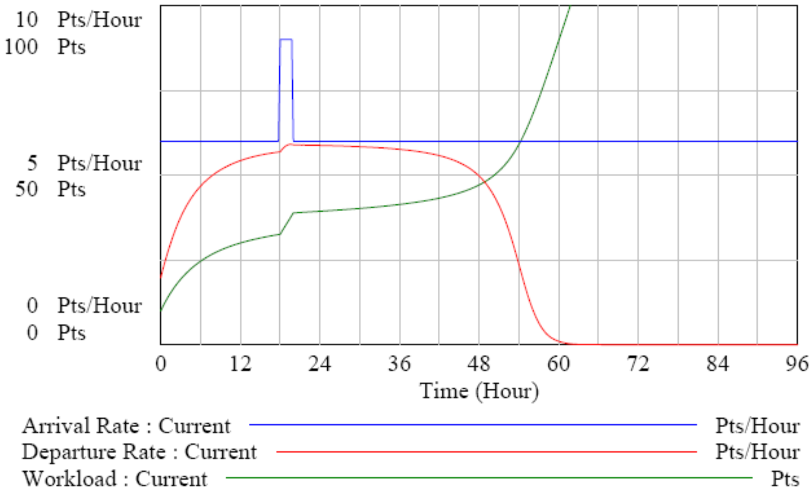

Fig. 3. Brittle response of the simple module to a single pulse load. Performance initially compensates, but eventually deteriorates catastrophically, some time after the initial shock

The addition of adaptive capacity to the model mitigates this response (Nathanael & Marmaras, 2008). Adaptive capacity is not further specified here – in a real world model it might take the form of work-arounds or short cuts that are employed when workers recognize overload and try to compensate for it. Figure 4 illustrates an adaptive, resilient response to a single pulse shock.

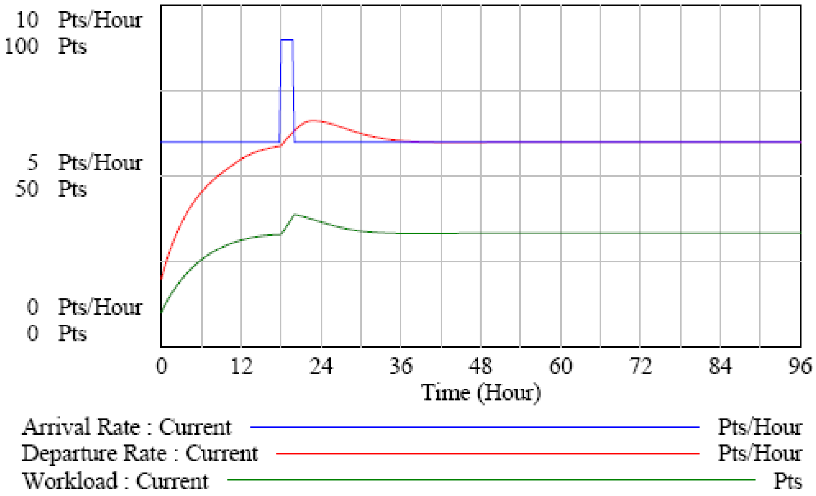

Fig. 4. Resilient response of the simple module to a single pulse load. Performance rises to compensates and gradually returns to the steady state.

Finally, we evaluated Conjecture 3, the ability of the system to permanently adapt to repeated shocks by varying the rate of decay of adaptations. Figure 5 illustrates this effect for two rates of decay. After 3 pulse challenges, the two systems start to diverge, and the system with rapid decay (red line) undergoes a phase transition into catastrophic collapse, while the system with a longer 'memory' is able to shift to a new steady state at a higher load.

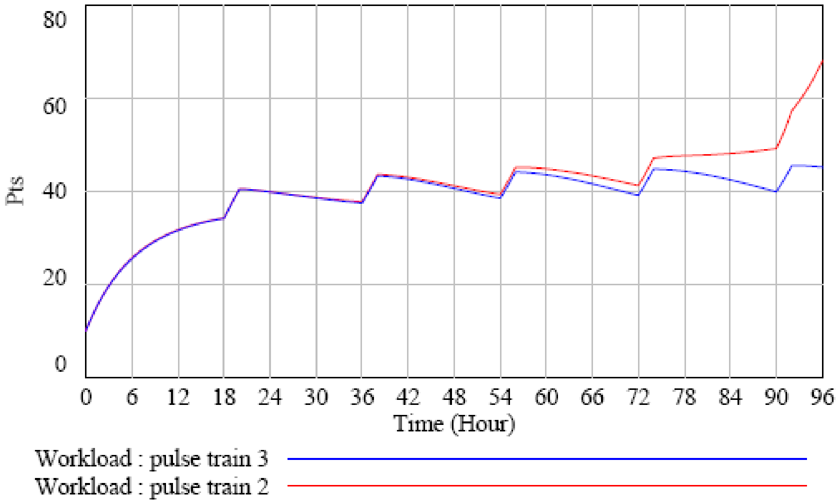

Fig. 5. Comparison of rapid (red) and slow (blue) decay of adaptations when challenged with repeated pulse shocks.

4 DISCUSSION

"All models are wrong, but some models are useful" (Box, 1976). The models presented here are limited and highly abstract, but illustrate the possibility of obtaining more complex behaviours, such as resilience and brittleness from simple components. However, the models currently omit many important variables, such as fatigue, burnout, or the effect on quality (particularly the effect of rework as a result of diminished quality). Once refined and linked into a larger representation of the ED within the hospital, they may provide insight into a public policy problem that has so far resisted solution.

REFERENCES