ELEMENTS OF A HYBRID SIMULATION MODEL: A CASE STUDY OF THE BLOOD SUPPLY CHAIN IN LOW- AND MIDDLE-INCOME COUNTRIES

Bhakti Stephan Onggo
Department of Management Science
Lancaster University Management School
Lancaster University
Lancaster, LA1 4YX, UNITED KINGDOM

ABSTRACT

A hybrid simulation model is a simulation model that is formed from at least two different simulation modelling methods (e.g. discrete event, system dynamics, agent-based). The use of different simulation modelling methods in one model requires modellers to specify additional model elements. This paper discusses three elements, namely, the modules, module interfaces and updating rules. Each module may use a different simulation method. The interface between modules defines the information that will be passed between them (including aggregation and disaggregation). The updating rules define how the information sent by one module affects other modules. These three elements are explained using a case study of a blood supply chain simulation model for low- and middle-income countries (LMIC) which has different characteristics and challenges in comparison to the typical blood supply chain in high-income countries (HIC).

1 INTRODUCTION

A hybrid simulation model is a simulation model that is formed from at least two different simulation modelling methods, e.g. discrete-event simulation (DES), system dynamics (SD) and agent-based simulation (ABS). Hybrid simulation has been attracting the interest of researchers for some time. Some researchers argue that one of the key benefits of combining different simulation methods is the potential to combine the benefits and virtues of each method (Brailsford, Desai, and Viana 2010; Zulkepli, Mustafee, and Eldabi 2012). Others use a hybrid simulation model because they want to capture both the continuous and discrete elements present in the real world system that is being modelled (e.g. Lee et al. 2002) or the ability to observe the behaviour of the model at macro and micro levels concurrently when the behaviour at one level affects the behaviour at the other level (Onggo, Kusano, and Sato 2007). While it might be possible to build a model for such a real-world system using a non-hybrid method, Morgan, Howick, and Benton (2011) argue that, in a complex problem, our model may not achieve all of the intended objectives if the model is developed using one method without making additional assumptions. Given the growing body of knowledge about hybrid simulation methods, it would be useful for simulation practitioners (industry and academic) if we could formulate what needs to be done differently when building a hybrid simulation model in comparison to building a simulation model using one method only.

This paper discusses three elements that a modeller needs to define when building a hybrid simulation model. They are the modules (each module may use a different simulation method), the interface between them (including the aggregation and disaggregation of information) and the updating rules that define how information sent by one module affects other modules. Although these three elements are crucial when developing a hybrid simulation model, this does not mean that this work is exhaustive. Indeed, there may be other elements that we have not discussed in this paper. To make the discussion more concrete, a
case study of a typical blood supply chain in low- and middle-income countries (LMICs) will be used. The blood supply chain model is formed from three modules: donors, blood centre and hospitals. Based on the key performance indicators (KPIs) required for each module, a hybrid simulation seems to be a sensible approach. The justification for using hybrid simulation will be explained in the ensuing discussion.

The case study itself is chosen because of the author’s interest in the modelling of blood supply in LMICs. It is disheartening to know that shortages of blood for transfusion in LMIC have resulted in many otherwise avoidable deaths, especially of women of reproductive age and young children (Bates et al. 2008). In contrast, in high-income countries (HICs) the capacity of transfusion services to meet demand has been improved by research in various disciplines (Williamson and Devine 2013), including the use of simulation modelling. Simulation modelling has also frequently been used to improve supply chains in other contexts (Tako and Robinson 2012). Hence, simulation modelling should also be applicable to blood supply chain modelling in LMICs. Unfortunately, research into blood supply chains is rarely conducted in the context of LMICs. Furthermore, blood supply chains in LMICs have different characteristics and challenges in comparison to the typical blood supply chain in HIC. Therefore, best practice in the blood supply chains in HICs may not be applicable to LMICs. We believe that simulation modelling may help us to tackle many of these challenges and provide evidence suggesting effective interventions to improve the blood supply chain in LMICs.

The remainder of this paper is organized as follows. Section 2 provides an overview of the blood supply chain in LMICs. Section 3 presents the conceptual model of hybrid simulation. This section also explains the three elements of hybrid simulation and discusses them using a case study of a typical blood supply chain in LMIC. Finally, Section 4 concludes our paper.

2 BLOOD SUPPLY CHAIN IN LOW- AND MIDDLE-INCOME COUNTRIES

Following the advocacy of the World Health Organisation (WHO), many LMICs have moved towards co-ordinating their blood services through regional blood centres (and in some cases, a national blood centre), a common practice seen in HIC. However, blood donations in some LMICs are still hospital-based, e.g. in sub-Saharan Africa (Field and Allain 2007), with heavy reliance on family replacement donors instead of voluntary non-remunerated donors. Another donor-related issue in LMICs is that the blood donation rate is significantly lower than the rate in HICs. To make matters worse, some LMICs, especially in Africa, suffer from a high prevalence of HIV and other transfusion-transmitted infections (TTIs) which further reduce the supply of safe blood. This has led to a donation practice that relies heavily on younger donors (age 16 to 25) because they are perceived to have a lower risk of carrying any TTIs. Donors in this age group are mostly students. Hence, the number of donations is seasonal (lower during school holidays).

A blood centre in an LMIC faces challenges that are different from a HIC blood centre. A blood centre in an LMIC rarely separates blood into its components. This is because the main purpose of blood transfusion in LMICs is to cover emergencies such as pregnancy-related complications and severe childhood anaemia. Besides, the cost of separating blood into several components (red blood cells, platelets, plasma etc.) is less affordable for LMICs. There are also concerns over the cost of having more centralised blood centres which are more expensive than a hospital-based system (Field and Allain 2007). Apart from this, there is also the issue of the low quality of blood screening due to a lack of training and standardised procedures.

Hospitals in LMICs also face different challenges than hospitals in HICs. Bates et al. (2008) reported a lack of safe blood supplies in hospitals in LMICs. There is also an issue concerning the quality and availability of data for the demand for blood. Hence, it is hard to know the real demand for blood. The real situation could be worse. Due to reasons such as access to a hospital and affordability, some people may not be able to receive the healthcare they need and in some cases cannot even travel to the nearest hospital. This demand is not reflected in the existing data. As mentioned earlier, the main purpose of
blood transfusions in LMICs is to handle emergencies. Hence, more thorough blood-compatibility testing such as complete cross-matching may not be performed.

In summary, the challenges faced by LMICs are complex and differ from the challenges faced by HICs. These challenges include an inadequate blood supply, a scarcity of data on blood supply and demand, a high dependence on replacement donors (usually family members of the patient), very low rates of repeat donation (i.e. the most reliable blood supply and the lowest rates of TTIs), fragmented and diverse blood providers, scarce resources and a high prevalence of TTIs resulting in a high risk of transmission and high discard rates (Vermeulen and Reddy 2010).

3 HYBRID MODEL DESIGN

This section discusses the three elements needed when building a hybrid simulation model: modules, module interfaces and updating rules. These elements are explained sequentially. In practice, the process can be iterative. For example, after we identify a number of modules, as we are working on the interfaces between modules, we may decide to combine modules or create a new module.

3.1 Modules

A module encapsulates one logical component of the model that can be implemented using any of the simulation modelling methods. The decision regarding the boundary of a module is case-dependent. In principle, a module should be self-contained with predefined interfaces (input and output) to the outside world, including other modules. Hence, it is possible to run a module as an independent model provided that we specify the necessary information for its input interface. A modeller needs to justify the use of a particular simulation method for each module (i.e. to determine if the chosen simulation method is the best one for the job (Brailsford et al. 2013)). A module can produce two types of output: simulation output (to be used for analysis) and input for other modules (to be used in a hybrid simulation run).

![Figure 1: Modules and their interfaces.](image)

The case study used in this paper is a typical blood supply chain in LMICs in which a regional blood centre covers a number of hospitals and pools the supply of blood in the region. At the same time the hospitals may recruit family replacement donors when the required blood is not available. Hence, there are three main components in this supply chain: donors as the suppliers of blood, a blood centre that maintains the inventory and distribution of blood, and hospitals that require blood for transfusion services. As we will explain later, policymakers are interested in different KPIs for each component. Based on the KPIs, the blood supply chain model is divided into three modules: donor, blood centre and
hospital. The structure of the modules and their interactions are shown in Figure 1. A detailed description of each module will be discussed in this section and their interfaces will be discussed in Section 3.2.

3.1.1 Module Donor

This module models the life cycle of blood donors in the system. The main concerns related to blood donation for policymakers in LMICs are how to encourage people to become repeat donors and how to keep the current pool of repeat donors. In other words, the main KPI is the number of repeat donors. Hence, the model needs to show the number of people in various blood donation statuses (such as non-repeat donor, repeat donor and ineligible donor) in the system. The model does not need to track individual donors. Hence, the blood donor system can be seen as a flow model in which people move from one status to another over time. System dynamics is chosen because it has been widely used to model the flow of people in a system.

Figure 2 shows a conceptual model of module donor represented as a stock-and-flow diagram. Given that the planning horizon is typically less than five years, we assume that the numbers of people joining and leaving the population of blood donors (henceforth, population) are the same, which leads to a constant size. If we are interested in a non-constant population, we can add incoming rate and leaving rate to the model. The population of blood donors is individuals in the population who are within the age limit(s) for blood donation (some countries may specify a lower age limit only).

The population can be divided into three mutually exclusive groups (see the highlighted rectangles in the figure). They are non-repeat donors (NRDs), repeat donors (RDs) and ineligible donors (NDs). NRDs include people who have never donated blood and those who have donated blood but not regularly. A repeat donor is an individual who donates blood regularly. People in the ND category are not eligible to donate blood because blood test results show that they may have at least one disease that can be transmitted through blood transfusion, such as HIV, hepatitis B (HPB) or hepatitis C (HPC). Given the high dependency on family replacement donors (FRDs), we need to include them in this module. An FRD is usually a friend or family member of the person needing the blood transfusion. Hence, a FRD can come from among either NRDs or RDs. The blood donation is taken at the hospital where the person is treated. In the model, we separate an FRD who is an NRD (FRD1) from an FRD who is an RD (FRD2) so that we can return him/her to his/her original status after the blood donation has been taken at the hospital.

A non-repeat donor who wants to donate blood must fill in a self-assessment questionnaire to identify potential health problems for the donor and potential TTIs that may affect people needing a blood transfusion. The donors who pass this stage will be checked to see if their haemoglobin (Hb) level is acceptable. Those with an acceptable Hb level are allowed to donate blood. Subsequently, the blood will be tested for TTIs at a blood centre. The details of the process at the blood centre will be explained in the next subsection. For simplicity, in the model, those who fail TTI testing are treated as if they are not eligible to donate blood anymore. This should not affect the results significantly, given the relatively short planning horizon and that some people with TTIs may not be completely cured over the planning horizon.

Ideally, we need to encourage those who pass TTI testing to become repeat donors. Repeat donors are not only good for providing a stable blood supply but they are also likely to pass future TTI testing (which will bring the cost down due to a lower discard rate). Repeat donors typically do not need to fill in a self-assessment questionnaire when they want to donate blood. They only need to undergo Hb and TTI tests. Although the chances of repeat donors failing TTI testing are low, the model needs to capture this possibility. Likewise, there is also a possibility that an RD will become an NRD.

The number of NRDs and RDs who have just donated blood in Figure 2 will generate new blood bags to be tested by the module blood centre. The TTI test results from the module blood centre will determine the donor’s next state in the system (the proportions may be different between RDs and NRDs). The interaction between these two modules will be discussed in Section 3.2.
3.1.2 Module Blood Centre

This module models the lifecycle of blood bags in the blood centre’s inventory. A typical blood centre manager is interested in certain KPIs, such as the number of fulfilled requests from hospitals, the number of expired blood bags and the number of new blood bags processed within a predefined duration (throughput). Since blood bags may have different expiry dates, we need to keep track of individual bags in the system. Hence, a DES or ABS model is suitable for this module. DES is chosen over ABS because of the familiarity of policymakers (clients) with a more process-oriented representation of the model. Figure 3 shows a conceptual model of the module represented using BPMN. For a detailed description of the use of BPMN for conceptual model representation please refer to Onggo (2009, 2010).

For each donor, the blood centre will draw a sample from the donor’s blood bag(s) to be tested for TTIs. The blood bags that are deemed cleared of TTIs will be put into different blood groups. A number of donors who pass or fail TTI tests will go through an ABO and rhesus-grouping procedure. The blood bags that have been labelled with their blood group will remain in the inventory until they are requested by a hospital or expire. This process is shown in the top lane in Figure 3 (to save space it only shows the process for RD because the process for RD and NRD is the same except for the start event). The second process in Figure 3 is triggered when a request note from a hospital arrives. The blood centre will decide whether they will allocate blood bags based on blood availability and their allocation policy (for example, a blood centre may want to maintain fairness when dealing with requests from a number of hospitals). As a result, a request may be fulfilled, partially fulfilled or not fulfilled. The allocated blood bags and an accompanying delivery note will be sent to the requesting hospital. The last process in Figure 3 is triggered when a blood bag expires. In this case, the blood bag needs to be removed from the inventory.

3.1.3 Module Hospital

The module hospital represents the transfusion services and the inventory system at a hospital. A hybrid simulation model may be run with multiple instances of module hospital. Each instance represents one hospital in the region served by the same blood centre. One of the most important KPIs in this module is
the number of stock-outs at each hospital, i.e. when the required blood is not available. In some cases, this has caused fatalities (Bates et al. 2008). To measure this KPI, we need to keep track of individual blood bags because, if the blood group needed is not available, the hospital will try to find a compatible blood group. The hospital also maintains a blood inventory and so we are also interested in reducing the number of expired blood bags. These requirements make DES or ABS suitable for this module. DES is chosen over ABS for the same reason as in the module blood centre.

Figure 3: Module blood centre.

Figure 4 shows a conceptual model for module hospital. The top lane shows the blood transfusion process followed at a typical LMIC hospital. In general, the process is relatively straightforward. When a patient requires a blood transfusion and compatible blood is available, the patient will receive a transfusion. This forms the happy path in the process “transfusion service” in Figure 4. When compatible blood is not available, friends or family members (i.e. FRDs) with a compatible blood group may be asked to donate. If this fails, there might be other options, such as requesting blood bags from another hospital. But for the performance analysis needed in this modelling work, we do not need to model the details. All we need is the number of stock-outs. To save space, Figure 4 does not show the complete cross-matching activity. Furthermore, most blood transfusions in LMICs are used for emergency patients. Hence, complete cross-matching activity may not be carried out. However, if cross-matching is done in the hospital, this activity can easily be added to the model.

The second lane shows that the hospital conducts a regular check on their blood supply (in this example, on each working day). If the stock of a certain blood group is low, they will order it from the blood centre. For simplicity, this model does not include any ad hoc requests made to the blood centre.
When the hospital receives blood bags from the blood centre, they will add them to the inventory as shown in the third lane. The final lane shows that when a blood bag expires, it has to be discarded.

**Figure 4: Module hospital.**

### 3.2 Module Interfaces

One of the factors to identify a module is the presence of a clear and logical interface between the module and other module(s). An interface between two modules should decouple them. If the interface is messy and intertwined, it may be more practical to combine them into one integrated module. An interface defines the information that will be passed between one module to the other module(s), the module that generates the information, and the receiving module(s). Ideally, the information should have an actual meaning in the real-world system being modelled (i.e. not modelling artefacts). This will enhance the credibility of the hybrid model to the model users (or clients).

Two modules in a hybrid model may have different levels of detail. For example, one module may operate at the population level and the other at an individual level. In this case, the information flow from the module with the more detailed level to that with a less detailed level has to be aggregated. The flow in
the opposite direction is more challenging because the information needs to be disaggregated. The
disaggregation may require one or more distribution functions to transform a given population size to a
set of finer units that form the population with the necessary heterogeneity.

In the case study, donors provide the blood supply to a blood centre. This is represented as an
interface between module donor and blood centre. Module donor operates at the population level. Hence,
it does not track individual donors. In this interface, module donor will send the number of donors who
have donated within a predefined unit of time (in this case one working day) to module blood centre.
Since module blood centre operates at the individual level, the number of blood donors will be
disaggregated to create a set of individual donors and their blood bag(s). Individual donors will be
generated based on a set of distribution functions to represent the heterogeneous profile of the donors
based on existing data.

Subsequently, module blood centre determines individual donors who pass or fail TTI testing. The
blood bags from donors who pass will be added to the inventory. The TTI results will be sent to module
donor. The TTI results will be aggregated into the number of passes and fails grouped by status (RD or
NRD). As we can see from the conceptual model of module donor (Figure 2), these numbers will be used
to control the flow of donors in the module.

Module hospital generates the demand for blood by regularly sending requests for blood bags to
module blood centre to replenish its inventory. Subsequently, module hospital will receive the blood bags
allocated by module blood centre. Both modules operate at the same level of detail. Hence, the
information passed between them does not need to be aggregated or disaggregated.

In cases where compatible blood groups are not available, a hospital may ask family replacement
donors (FRDs) with a compatible blood group to donate. Module hospital recruits a FRD from the pool of
NRDs and RDs from module donor. After they have donated blood at the hospital, they will return to
their original status (as NRD or RD). Because the two modules operate at different levels of detail,
aggregation and disaggregation processes are needed. Module hospital generates the number of FRDs and
sends the aggregated number to module donor. Module donor will use that number to move some NRDs
and RDs to FRD. The FRD will move back to their original status (NRD or RD) after some delay (i.e.
donors have to wait for a few months before they may next donate blood). This is why FRDs are split into
FRD1 and FRD2 in Figure 2, i.e. to allow them to return to their original status.

3.3 Updating Rules

The interfaces between modules discussed in the previous section do not define: (1) when information
will be sent from one module to another and (2) how information will affect receiving modules. When
developing a hybrid simulation model, modellers need to define the times when information will be sent
and how receiving modules handle new information (i.e. updating rules). Updating rules are necessary to
maintain logical consistency among all the modules when information is sent from one module to
another. One of the complications caused by hybrid simulation is that each module in a hybrid simulation
model may reach different points in simulation time when they are run concurrently. There are three
reasons why this might happen during a simulation run. First, each module in a hybrid simulation model,
by definition, may use a different simulation method, and each simulation method may use a different
time advancement method. For example, SD uses fixed-time increments, ABS typically uses fixed-time
increments and DES typically uses variable-time increments (i.e. next event time). Secondly, even when
two modules use the same time advancement method, they may use different time units (e.g. minutes,
days). Finally, each module may be run using different simulation software with its own internal time
management.

Figure 5 shows how the updating points can be defined between two modules. The top part of the
figure (5a and 5b) shows that both modules use fixed-time increments but different units of time. The
updates can be done asynchronously, i.e. every time a module advances its simulation time which may
change its state, the module will send new information to all recipients as defined in its interface(s) (note
that an interface defines the information to be sent and the recipient(s)). The updates can also be done synchronously, i.e. at predefined simulation points. For example, in Figure 5b, we set the updating points to be module 2’s time step (we can of course define it to be module 1’s time step if this is more suitable). In this case, all modules will send their information to other modules as defined in their interfaces. The asynchronous and synchronous updating points can also be applied to other combinations of time-advancement methods. The middle part of the figure (5c and 5d) shows the case when one module uses fixed-time increments and the other module uses variable-time increments. The bottom part (5e and 5f) shows the case when the two modules use variable-time increments. After defining the updating points, modellers need to define how the receiving modules handle each flow of incoming information (i.e. the rules). For example, in Figure 5a, modellers need to define how module 1 will use the information sent by module 2 at times $t_{2,2}$ and $t_{2,3}$. One alternative is that modellers assume that the information will not affect module 1 significantly so that module 1 will use the information in the following time step.

The updating rules discussed in this section refer to the modelling aspect of simulation, i.e. the logical behaviour of modules when they receive information from another module. Readers who are familiar with parallel and distributed simulation will understand that, underneath this, a synchronisation algorithm is needed to ensure that messages that carry information are received correctly by all modules. The synchronisation algorithm is relatively straightforward if the modules are run using the same simulation
Onggo

software on a sequential processor (because they share the same global simulation clock). If the modules are written using different software, we will need a hybrid simulation tool that implements one of the synchronisation algorithms from parallel simulation or uses HLA-RTI (High Level Architecture – Run Time Infrastructure) software to link the modules. A good hybrid simulation tool should free the modeller from this low-level synchronisation algorithm so that they can focus more on model development (including the three elements discussed in this paper).

4 CONCLUSION

This paper has discussed three elements that a modeller needs to define when building a hybrid simulation model. The three elements are modules, module interfaces and updating rules. The three elements are consistent with the two structures identified in Chahal and Eldabi (2008): hierarchical and process environments. This is because the modules in these two structures are clearly separated. The boundary between modules for the other structure, i.e. integrated, is not obvious because the interfaces between modules are intertwined. In my view, even if a hybrid model is highly integrated, it should still be possible to dissect the model into smaller components in such a way that each component (e.g. function, activity, agent) accepts a set inputs and transform those into a set of outputs. It is a design decision to consider whether a small component is worth implementing as a module in a hybrid simulation model or not. Hence, in theory, the three elements are also consistent with the integrated structure. Whether or not it is practical to define the three elements for a highly integrated model is another matter. A conceptual model of the blood supply chain in low- and middle-income countries has been used to explain the three elements. The work in implementing the conceptual model presented in this paper is ongoing.

REFERENCES


Onggo


**AUTHOR BIOGRAPHY**

BHAKTI STEPHAN ONGGO is Lecturer (Assistant Professor) in the Department of Management Science at Lancaster University Management School (LUMS), Lancaster, United Kingdom. He completed his PhD in Computer Science at the National University of Singapore and his MSc in Management Science at Lancaster University. His research interests lie in the areas of simulation methodology (conceptual modeling, discrete-event simulation, system dynamics and agent-based simulation), simulation technology and simulation modelling in business and management. His email address is s.onggo@lancaster.ac.uk.