

An Overflow Loss Network Model for Capacity Planning of a Perinatal Network

Md Asaduzzaman and Thierry J. Chausalet

University of Westminster, London, UK.

Summary. In this paper, a model framework is developed to solve capacity planning problems faced by many perinatal networks in the UK. We propose a loss network model with overflow based on a continuous-time Markov chain for a perinatal network with specific application to a network in London. We derive the steady state expressions for overflow and rejection probabilities for each neonatal unit of the network based on a decomposition approach. Results obtained from the model are very close to observed values. Using the model, decisions on number of cots can be made for specific level of admission acceptance probabilities for each level of care at each neonatal unit of the network and specific levels of overflow to temporary care.

Keywords: Queueing network model; Decomposition; Rejection; Continuous-time Markov chain

1. Introduction

Every year over 80,000 (approximately 10%) neonates are born premature, very sick, or very small and require some form of specialist support in England (DH, 2003; RCPCH, 2007). Neonatal services aim to offer high quality care for these vulnerable babies. Over a six month period in 2006-07, neonatal units were shut to new admissions for an average of 24 days. One in ten units exceeded its capacity for intensive care for more than 50 days during a six month period (Bliss, 2007). The National Audit Office reported that capacity and staffing problems at unit level continue to constrain neonatal service (NAO, 2007). Therefore, capacity planning has become a major concern for all neonatal units in England.

Neonatal care is provided in specialist units that are graded into three levels set by the British Association of Perinatal Medicine. A level 1 unit called special care baby unit provides special care which is the least intensive and most common type of care. In this unit, neonates may need to be fed through a tube, be supplied with extra oxygen or treated with ultraviolet light for jaundice. In a level 2 unit, neonates may receive high dependency care such as breathing via continuous positive airway pressure or intravenous feeding and called high dependency unit. These units may also provide some short-term intensive care (often referred as intensive therapy unit). A level 3 unit provides the whole range of medical neonatal care: special care, high dependency care, and intensive care. In intensive care, called neonatal intensive care unit, neonates will often be on a ventilator and need constant care to be kept alive. It is a national recommendation that neonates with complex needs or requiring long periods of respiratory support have their initial care in a level 3 unit. A level 3 unit, also known as perinatal centre or lead centre of a perinatal network, plays a major role by providing

Address for correspondence: Md Asaduzzaman, Institute of Statistical Research and Training (ISRT), University of Dhaka, Dhaka 1000, Bangladesh.
E-mail: asad@isrt.ac.bd

the highest level of care for the most vulnerable neonates. Level 2 and level 3 units may also have some transitional care cots to tackle overflow and rejection from special care baby unit.

In a neonatal unit, neonates are admitted to neonatal intensive care unit (NICU), high dependency unit (HDU), or special care baby unit (SBU) according to the severity of their condition, which is usually driven by their gestational age and birth weight. As their condition improves, neonates will eventually be transferred to a sub-unit providing a lower level of care (i.e. from neonatal intensive care unit to high dependency unit or high dependency unit to special care baby unit) within the same unit or in another neonatal unit from where they are discharged. Similarly, those admitted to high dependency unit move to special care baby unit before discharge. Neonatal intensive care unit (respectively high dependency unit and special care baby unit) neonates are sometimes initially cared at high dependency unit (respectively neonatal intensive care unit or special care baby unit and high dependency unit or transitional care) when all neonatal intensive care unit (respectively high dependency unit and special care baby unit) cots are occupied. This temporary care is provided by staffing a cot with appropriate staff and equipment resources and referred to as overflow. This only occurs when neonates must be admitted because they are booked at the same hospital or severely ill. Rejection from neonatal intensive care unit (respectively high dependency unit and special care baby unit) occurs only when all cots at the neonatal intensive care unit (respectively high dependency unit and special care baby unit) and high dependency unit (respectively special care baby unit and transitional care) are occupied. Therefore, no neonatal unit allows queue to form when all cots are occupied; and also overflow and rejection are evident in a neonatal unit. In case of rejection, neonates are transferred to another neonatal unit.

‘Managed clinical networks’ (MCN’s) for neonatal care were established in 2004 to achieve an appropriate level of care in the right place, and have to provide all levels of care for 95% of mothers and neonates within the network area where they live. An ‘MCN’, also called a neonatal or perinatal network, is organised in such a way that there is at least one level 3 unit which closely works with other level 2 and level 1 units. A survey on neonatal units reported that the 182 neonatal units established in England are organised in 24 networks (Redshaw and Hamilton, 2005).

In London, for example, there are five perinatal networks set up in south west, south east, north west, north east and north central, to coordinate and manage neonatal care. The rejection of admission requests from NICU and HDU is also increasing at neonatal units in London due to capacity shortage. In this paper we develop a capacity planning model for a perinatal network based on loss queueing network approach. The expressions for steady state overflow and rejection probabilities are derived. The model is then applied to a network in London.

The paper is organised as follows: in the next section we describe a network: the North Central Perinatal Network (NCLPN) which forms our case study, its organisation and guidance for transfers between units. In section 3, we develop a loss network model with overflow for the NCLPN; steady state expression of overflow and rejection probabilities for each neonatal unit are derived. Data and numerical results are presented in Section 4. Discussion and further research direction are given in Section 5.

2. Case study: the north central London perinatal network

The NCLPN is one of 24 perinatal networks across England and coordinates intensive, high dependency and special care provision for newborn infants in the London Boroughs of Barnet, Enfield, Haringey, Camden and Islington. Fig. 1[†] shows the neonatal network, working in north central London, consists of five neonatal units: Barnet, Chase Farm, Whittington, Royal Free and University

[†]Map source: <http://www.neonatal.org.uk/a/427>



Fig. 1. Neonatal units in the NCLPN.

Table 1. The number of funded cots at each hospital of the NCLPN in 2006.

Neonatal Unit	Level of care	NICU/HDU	SCBU	TC
Chase Farm	1	–	10	–
Royal Free	1	2	12	–
Barnet	2	6	14	4
Whittington	2	11	5	5
UCLH	3	12	15	8
Total	–	31	56	17

College London Hospital. There is also a highly specialised centre Great Ormond Street Hospital (GOSH) for treating very sick infants including neonatal surgery and some other specialised services. The network was setup to deliver a single neonatal service across London North Central area. The underlying aim of the network is to ‘Improve standards and achieve capacity so that 95% women and neonates may be cared for within the network.’ Cot distribution in each neonatal unit are presented in Table 1.

The underlying guidance for admission, discharge and transfer of the network are discussed below.

- (a) All mothers expecting birth < 27 week gestational age or all neonates with < 27 week gestational age are transferred to UCLH.
- (b) All mothers expecting birth ≥ 27 but < 34 week gestational age or all neonates with the same gestational age are transferred to Barnet or Whittington depending upon the booked place of delivery.
- (c) All neonatal unit accept neonates for special care booked at the same unit.
- (d) Neonates admitted into units other than booked place of delivery are transferred back to the respective neonatal unit after the required level of care.

In 2006, a common information system called SEND (South England Neonatal Database) was established. This information system gives an opportunity to compare the neonatal units over time. A preliminary analysis shows that the UCLH neonatal NICU-HDU had rejected 33% neonates in 2006. Statistics for level 1 and 2 rejection from all units are also high. Table 1 shows the cot

distribution in the NCLPN neonatal units. The NICU capacity at UCLH unit has increased by 4 cots in 2007 and 1 cot in 2008. In Whittington, the capacity has expanded by 1 NICU cot in 2008; and 3 SCBU cots in 2007 and 8 SCBU cots in 2008 to tackle the increasing demand. But recent analysis shows that capacity crisis exists and yet to be resolved.

3. A loss network framework for a perinatal network

Queueing models have been used as capacity planning tool for a long time in many areas such as communication networks, manufacturing, health care etc. The earliest work on queueing networks dates back to J. R. Jackson (Jackson, 1957, 1963). The theory of queueing network models has advanced considerably over the past several decades (e.g. Gordon and Newell, 1967; Baskett et al., 1975; Reiser and Kobayashi, 1975; Kelly, 1979). Loss networks, a special class of queueing networks with no waiting space, provide a framework for studying the blocking behavior of connection-oriented services in circuit-switched networks, ATM (Asynchronous Transfer Mode) networks, optical networks and wireless networks (Kobayashi and Mark, 2002). The Erlang loss model, the simplest case of loss model, is of a special class of queueing models where the system has no waiting space. Systems with overflow have also been studied by many researchers (e.g. Wilkinson, 1956; Kaufman et al., 1981; Jagerman, 1984; Guerin and Lien, 1990). Overflow occurs only when primary servers are in blocking state, i.e., when all regular servers are busy an arriving job is rerouted to a secondary group of overflow servers. A number of overflow servers may be kept available for 'emergency' situations. In health care, systems with loss queueing have been considered by Kortbeek and van Dijk (2007), Bekker and de Bruin (2009), and with overflow by Litvak et al. (2008). Kortbeek and van Dijk (2007) proposed the Erlang loss bounds for Operating Theatre-Intensive Care Unit (OT-ICU) tandem system approximating M/G/c/0. Bekker and de Bruin (2009) used an M(t)/H/c/0 model to determine the impact of the time-dependent arrival pattern on the required number of operational beds and rejection of admission for clinical wards using approximations based on the infinite-server queue. Litvak et al. (2008) approximated the ICU rejection probability through the equivalent random method (ERM) considering a common overflow for a particular class of patients. In a preliminary study we developed an Erlang loss network model for UCLH where we ignored overflows (Asaduzzaman and Chausalet, 2008). Then, we developed a more realistic model with overflow for NICU-HDU and SCBU neonates for the UCLH (Asaduzzaman et al., 2009). The model did not capture overflow from SCBU to NICU-HDU as this was not observed at UCLH in 2006. But at UCLH in 2007 and 2008 and Barnet and Whittington in all three years, this flow has been observed. However, our present goal to develop a model framework for the whole perinatal network consists of level 1, 2 and 3 units with complex transfers and back transfers.

Now we shall develop a model for capacity planning of the NCLPN based on a loss network framework with overflow being motivated by the communication network models. The steady state overflow and rejection probabilities will be derived and based on these probabilities, cot capacities may be planned for the neonatal units in the NCLPN. Fig. 2 shows the topology of the NCLPN. The network consists of a level 3, two level 2, and two level 1 units. The Royal Free level 1 unit has also an ITU. We assume the inter-arrival times i.e., times between admission requests and the length of stay (LoS) at the neonatal units follow the Markov property, that is, follow exponential distribution for model tractability. The Markov property states that the probability distribution for the system at the next step (and in fact at all future steps) only depends on the current state of the system, and not additionally on the state of the system at previous steps. Then, the network can be modelled as a continuous-time Markov chain $\mathbf{X} = (\mathbf{X}(t), t \geq 0)$ that expresses the number of neonates in the

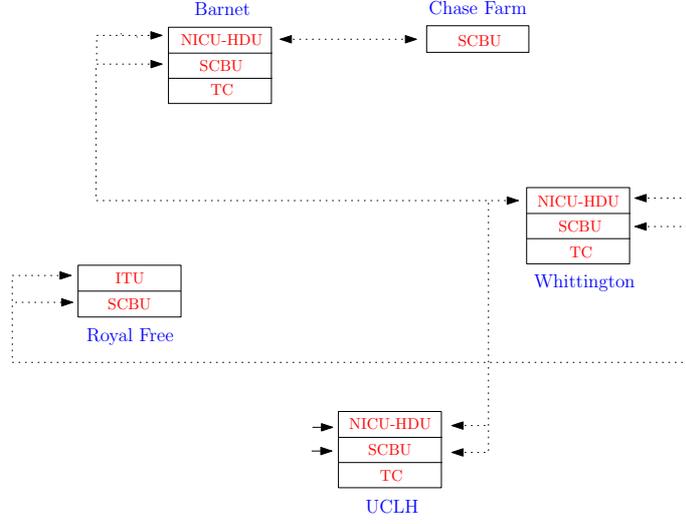


Fig. 2. Topology of the NCLPN.

network.

Let n_{i1} denotes the number of neonates admitted at NICU-HDU, and n_{i12} denotes the number of neonates in the SCBU overflowing from the NICU-HDU in the i th neonatal unit (level 3/2). Let n_2 be the number of neonates in the SCBU, n_{i21} be the number of neonates overflowing from the SCBU to the NICU-HDU, and n_{i23} be the number of neonates overflowing from SCBU to the TC in the i th neonatal unit (level 3/2). Then, for the i th neonatal unit (level 3/2), we can define

$$\mathbf{n}_i = (n_{i1}, n_{i12}, n_{i21}, n_{i2}, n_{i23}), \quad i = 1, 2, 3.$$

Then, a state of the Markov chain for the NCLPN is a vector

$$\mathbf{n} = (n_{11}, n_{112}, n_{121}, n_{12}, n_{123}, n_{21}, n_{212}, n_{221}, n_{22}, n_{223}, n_{31}, n_{312}, n_{321}, n_{32}, n_{323}, n_{41}, n_{42}, n_5).$$

We assume that the Markov chain is homogeneous, aperiodic and irreducible on its finite state space

$$S = \{\mathbf{n} : n_{i1} + n_{i21} \leq c_{i1}, n_{i21} + n_{i2} + n_{i23} \leq c_{i2}, n_{i3} \leq c_{i3}, n_{41} \leq c_{41}, n_{42} \leq c_{42}, n_5 \leq c_5, i = 1, 2, 3\},$$

where c_{i1} , c_{i2} and c_{i3} are the number of cots in the NICU-HDU, SCBU and TC, respectively, for the i th level 3/2 neonatal unit (UCLH, Barnet and Whittington), c_{41} and c_{42} are the cots for the level 1 unit with ITU (Royal Free), and c_5 is the number of cots for the level 1 unit (Chase Farm). Thus, \mathbf{X} is uniquely determined by its transition rates $\mathbf{Q} = (q(\mathbf{n}, \mathbf{n}') : \mathbf{n}, \mathbf{n}' \in S)$. The equilibrium distribution of \mathbf{X} is the unique distribution $\pi = (\pi(\mathbf{n}), \mathbf{n} \in S)$ that satisfies the *global balance equation*

$$\sum_{\mathbf{n}' \in S, \mathbf{n}' \neq \mathbf{n}} [\pi(\mathbf{n})q(\mathbf{n}, \mathbf{n}') - \pi(\mathbf{n}')q(\mathbf{n}', \mathbf{n})] = 0, \quad \mathbf{n} \in S.$$

The above stochastic process is a multidimensional Markov process which can be analysed through overflow loss network framework. We are particularly interested in the overflow and rejection probabilities in equilibrium state.

The equilibrium distribution of a multidimensional Markov process is often computed by numerically solving a set of *balance equations*. In overflow loss networks, equilibrium distribution generally does not admit a product form solution. The computational complexity increases radically with the number of dimensions. In such case explicit solutions for the equilibrium distribution π are usually obtained through a decomposition of the *global balance equation* into an appropriate set of *partial balance equations* (van Dijk, 1993). Although it is possible to construct an explicit expression of the equilibrium distribution, it cannot usually be computed for higher dimensional Markov chains. In our case, the NCLPN with 5 neonatal units each having NICU-HDU/SCBU/TC including overflow has 18 dimensions and have complex transfers and back transfers for which the equilibrium distribution cannot be computed practically. Even for a network with moderate capacity, one is forced to use alternative methods. Decomposition or other approximations therefore play a crucial role for computing steady state behaviour or estimating blocking probabilities. Among approximation techniques, *Erlang Fixed Point Approximation* (EFPA) is the most popular one (Kelly, 1986). The basic idea of EFPA is to decompose the whole loss system into a number of server-group subsystems and treat each subsystem as if it were an independent Erlang B sub-system. With a network of reasonable dimensions and capacity, the occupancy of neighbouring links may be highly dependent and the equilibrium distribution will no longer have a product form solution. Modelling dependencies in this context is therefore critical and EFPA may not perform well. Bebbington et al. (1998) proposed an improvement of EFPA to calculate blocking probabilities more accurately by specifically accounting for the dependencies between adjacent links. We adopt a similar approach to find the steady state behaviour for each unit of the network via decomposition. We first decompose the whole network into a set of subnetworks (i.e., ‘node-wise’ decomposition), namely, neonatal units. Then, deriving *station balance equation* (see van Dijk (1993) for details) for each node we obtain the steady state behaviour and expression for overflow and rejection probabilities for the subnetworks as if each subnetwork itself a full network. While deriving the steady state behaviour as well as estimating overflow and rejection probabilities, transfer of neonates between subnetworks are incorporated implying we specifically account the dependencies between adjacent units. In brief, our analysis principle by decomposition can be described by the following steps

- Decomposition of the whole network into subsystems, i.e., neonatal units which can be referred to as single queuing stations or subnetworks.
- Superposition or merging of arrivals at a neonatal unit.
- Analysis of each neonatal unit in isolation. The units are related to their network surroundings by input (arrival) and output (departure) processes.
- Performance measure for each neonatal unit as partial performance of the network.

Modelling assumptions

- (a) Admission requests follow a Poisson process and length of stay at neonatal units follows an exponentially distribution.

These assumptions make the model more tractable and are generally supported by the data. Although the sojourn times in perinatal units could be far from exponential as appeared in various medical references (e.g., Griffiths et al, 2006), it should be noticed from the insensitivity property in queueing models that a product form result (if any) obtained with Poisson arrival assumptions, will be valid for arbitrary non-exponential sojourn times (see van Dijk (1993) for more details and proof of the so-called *insensitivity* property). The Poisson arrival assumption is valid for the neonatal units considered for the NCLPN.

- (b) Arrivals of the aggregate product i.e., merging or superposition of arrivals at a neonatal unit follow a Poisson process. This assumption is inspired by the results of Palm and Khintchine (see Heyman and Sobel (1982)) and relies on the notion that the superposition of a large number of independent renewal processes can be approximated by a Poisson process.
- (c) A neonate in the overflow will join, i.e., switch-back to the original unit (known as *call packing* principle in the telecommunication literature (van Dijk and van der Sluis, 2009)) when a cot becomes available again in the original unit.
- (d) Mean sojourn times depends only on the type of patient and not on the unit, which can be referred to as ‘server independence’.

Capacity of the neonatal units may be planned based on the overflow and rejection probabilities for each of the neonatal units.

3.1. Model for UCLH/Barnet/Whittington neonatal unit

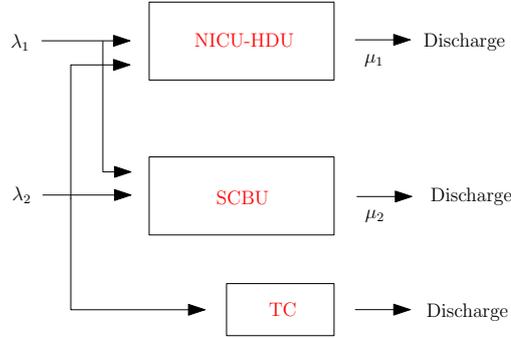


Fig. 3. Sub-network model for UCLH/Barnet/Whittington.

UCLH is the only level 3 unit of the NCLPN which consists of NICU-HDU, SCBU and TC. Assume neonates arrive according to a Poisson process with rate λ_1 (respectively λ_2) and have an exponentially distributed LoS with mean μ_1^{-1} (respectively μ_2^{-1}) at NICU-HDU (respectively SCBU and TC). The number of cots in NICU-HDU, SCBU and TC are c_1 , c_2 and c_3 respectively. If all c_1 cots are occupied at NICU-HDU, neonates are admitted to SCBU initially (overflow) if there is an empty cot, and rejected otherwise. Similarly, neonates arrive at SCBU; if all c_2 cots are occupied they are moved (overflow) to the NICU-HDU if there is an empty cot, otherwise move to TC (overflow) and rejected if all cots are occupied. We assume the LoS distribution in overflow is also exponential with mean μ_{12}^{-1} for NICU neonates overflow to SCBU and μ_{21}^{-1} and μ_{23}^{-1} SCBU neonates overflow to NICU-HDU and TC respectively. NICU-HDU neonates who are in SCBU because of overflow eventually join NICU-HDU at rate γ_1 . Similarly, neonates who are in NICU-HDU eventually join SCBU at rate γ_2 , who are in TC join SCBU at rate γ_3 .

Let $X_i(t)$ be the number of neonates at unit i and $X_{ij}(t)$ be the number of neonates overflowing from unit i to unit j , $i, j \in 1, 2, 3$, at time t . Then,

$$\mathbf{X} = (X_1(t), X_{12}(t), X_2(t), X_{21}(t), X_{23}(t), t \geq 0)$$

is a continuous-time Markov chain with state space given by

$$\mathbf{S} = \{ \mathbf{n} = (n_1, n_{12}, n_2, n_{21}, n_{23}) : n_1 + n_{21} \leq c_1, n_{12} + n_2 \leq c_2, n_{23} \leq c_3 \}.$$

Now we assume LoS distribution is the same for NICU-HDU neonates at NICU-HDU and residing at SCBU for temporary care and also SCBU neonates at SCBU and TC, i.e. $\mu_{ij} = \mu_i$, $i, j \in 1, 2, 3$; and $\gamma_i = \infty$, $i = 1, 2, 3$, implies that as soon as a cot becomes available a neonate in temporary care is brought back to the main unit. Then, the steady state solution can be given as

$$\pi(\mathbf{n}) = G^{-1} \frac{\left(\frac{\lambda_1}{\mu_1}\right)^{(n_1+n_{21})} \left(\frac{\lambda_2}{\mu_2}\right)^{(n_{12}+n_2+n_{23})}}{(n_1+n_{21})!(n_{12}+n_2+n_{23})!}, \quad \mathbf{n} \in \mathbf{S}, \quad (1)$$

where G is the normalising constant. It is interesting to note that equation (1) is similar to the steady state distribution obtained by van Dijk and van der Sluis (2009).

The rejection probability for i th unit, $i = 1, 2$, is then

$$R_i = \sum_{\mathbf{n} \in T_i} \pi(\mathbf{n}),$$

where

$$T_1 = \{\mathbf{n} \in \mathbf{S} | (n_1 + n_{21} = c_1 \text{ and } n_{12} + n_2 = c_2)\},$$

and

$$T_2 = \{\mathbf{n} \in \mathbf{S} | (n_{12} + n_2 = c_2 \text{ and } n_{23} = c_3)\}.$$

The overflow probability O_i for i th unit, $i = 1, 2$, can be computed by the following formula

$$O_i = \sum_{\mathbf{n} \in \{T_i^* \setminus T_i\}} \pi(\mathbf{n}),$$

where

$$T_1^* = \{\mathbf{n} \in \mathbf{S} | (n_1 = c_1 \text{ and } n_{12} + n_2 < c_2)\},$$

and

$$T_2^* = \{\mathbf{n} \in \mathbf{S} | (n_1 < c_1 \text{ and } n_2 = c_2) \text{ or } (n_{12} + n_2 = c_2 \text{ and } n_{23} < c_3)\}.$$

Barnet and Whittington are level 2 units in the NCLPN; and have similar structure. Each unit consists of NICU-HDU, SCBU and a TC and has major similarities with UCLH (level 3). The overflow and rejection occur the same way in these units as observed for the UCLH (level 3). Therefore, models for Barnet and Whittington are similar to the UCLH except from the fact that often discharges at UCLH NICU-HDU join at Barnet or Whittington NICU-HDU after specialists support at UCLH depending upon the booked place of delivery. Hence, the back transfers from the UCLH are merged with the arrivals in Barnet and Whittington. Rejection and overflow probabilities can be computed similar way as computed for the UCLH.

3.2. Model for Royal Free/Chase Farm neonatal unit

Royal free neonatal unit consists of two units: NICU-HDU and SCBU. Unlike UCLH or Barnet/Whittington, overflow does not occur in this unit due to the admission policy. Therefore, the simplest Erlang loss system is used to calculate the blocking behaviour for this unit. Assume neonates arrive according to a Poisson process with rate λ_1 to the NICU-HDU and λ_2 to the SCBU. The LoS is exponentially distributed with mean μ_1^{-1} at NICU-HDU and μ_2^{-1} at SCBU. The number of cots in NICU-HDU and SCBU are c_1 and c_2 respectively. Let δ_2 be the rate of transfer of neonates to Royal Free SCBU from Whittington NICU-HDU which is assumed to form a Poisson process with λ_2 .

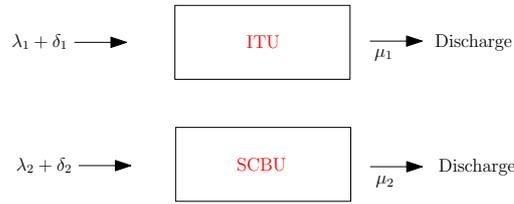


Fig. 4. Sub-network model for Royal Free.



Fig. 5. Sub-network model for Chase Farm.

The steady state distribution is given by

$$\pi(\mathbf{n}) = G^{-1} \frac{\left(\frac{\lambda_1 + \delta_1}{\mu_1}\right)^{n_1} \left(\frac{\lambda_2 + \delta_2}{\mu_2}\right)^{n_2}}{n_1! n_2!},$$

where G is the normalising constant.

The rejection probability for NICU and SCBU can then be calculated by following formula

$$R_i = \frac{1}{c_i!} \left(\frac{\lambda_i + \delta_i}{\mu_i}\right)^{c_i} \left[\sum_{n=0}^{c_i} \frac{1}{n!} \left(\frac{\lambda_i + \delta_i}{\mu_i}\right)^n \right]^{-1}, \quad i = 1, 2.$$

The Chase Farm unit has only a SCBU. The steady state probability distribution and the rejection probability at SCBU will be calculated by the same formula as given for Royal Free SCBU.

4. Application of the model

4.1. Data

The study uses data collected through SEND by each neonatal unit in the NCLPN. When a neonate is admitted to a neonatal unit booked place of delivery, place of birth, birth time, gestation, birth weight, episode, reason for admission, admission time, admission temperature, admission from hospital, discharge destination, discharge destination hospital name, discharge time, level 1 days, level 2 days and level 3 days are recorded in SEND. Therefore, SEND contains comprehensive and accurate information about each admitted neonate's arrival and LoS in neonatal units.

The basic parameters required for the analysis of the model are actual arrival rate and LoS at each level of care of all neonatal unit in the NCLPN. Model parameters were constructed in three different time periods to assess the consistency of results over time: period 1 (2006), period 2 (2007) and period 3 (2008). SEND does not have any record of requests for admission, which are required to estimate the arrival rate to the units. Therefore, we use a computer package SIMUL8[®] (Simul8, 2000) to obtain these arrival rates. These estimated arrival rates are validated by the clinicians and staff at each of neonatal units. Table 2 shows mean inter-arrival times and mean LoS for each of the units.

Table 2. Mean inter-arrival times (in days) and mean LoS (in days) of neonates at each neonatal unit in the NCLPN during 2006-08

Unit	2006		2007		2008	
	Inter-arrival	LoS	Inter-arrival	LoS	Inter-arrival	LoS
UCLH						
NICU-HDU	0.82	14.26	0.79	13.20	0.58	11.51
SCBU	0.40	8.23	0.33	7.26	0.24	5.83
Barnet						
NICU-HDU	2.20	13.31	1.45	9.52	1.12	6.78
SCBU	1.15	14.17	0.83	9.87	0.83	9.71
Chase Farm						
SCBU	–	–	1.30	7.60	1.05	8.03
Whittington						
NICU-HDU	1.76	8.37	0.95	6.19	1.21	5.16
SCBU	0.92	16.24	0.88	16.26	0.98	14.61
Royal Free						
NICU-HDU	5.11	2.91	2.43	1.17	2.77	2.21
SCBU	0.69	12.53	0.83	9.67	0.91	9.99

4.2. Results

The model analyses have been performed to neonatal units in the NCLPN. The calculated rejection probability by the overflow model with existing arrival and LoS pattern at each level of care for the neonatal units are presented in Table 3. The calculated rejection probability by the Erlang model and the observed probability are also given for comparison. The calculated rejected probabilities by the overflow model are close to the observed probabilities implying that the models are accurate enough to produce the rejection probability at each level of care in the neonatal units and are an improvement over the standard Erlang's loss model.

Table 3 shows that the UCLH unit has been performing better than past years since its NICU-HDU capacity has increased. Consequently, a substantial decrease of rejection probability has been achieved from 0.3250 to 0.1895. The model confirms the same level of decrease of rejection probability. With 22 NICU-HDU, 19 SCBU and 8 TC cots, the rejection probability at NICU-HDU would be dropped to 0.1059 and 0.0483 at SCBU assuming arrival and LoS pattern remains the same as 2008.

The rejection probabilities for the Barnet unit are also shown in Table 3. With 14 SCBU cots, the SCBU has kept rejection probability at an acceptable level. But the NICU-HDU requires 4 extra cots to get the rejection level 0.0597 with arrival and LoS patterns in 2008.

The Whittington neonatal unit has been performing better than any other unit in the NCLPN. With the current 12 NICU-HDU cots, the unit has less than 3% rejection probability at NICU-HDU. Due to the increase of 11 SCBU cots, rejection probability at SCBU has been decreased significantly.

We notice that the calculated probabilities for SCBU-TC at Barnet, NICU-HDU at Whittington and for SCBU-TC at UCLH in 2006 are not very close to the observed ones. However, since these values are < 0.05 or close to 0.05, they do not have any impact on management decisions regarding the number of cots. On the other hand, when rejection probability is high (say, 0.10), then observed and calculated values are close to each other.

Table 3 also show the calculated overflow probabilities for UCLH, Barnet and Whittington neonatal units. At UCLH, the calculated overflow probabilities are low. However, at Barnet NICU-SCBU, more than 10% NICU-HDU neonates are cared at SCBU. At the Whittington unit, the over-

Table 3. Rejection probability at each level of care of the neonatal units in the NCLPN during 2006-08

Year	Unit	No. of cots	Obs. rej. Prob.	Cal. rej. prob. (Erlang model*)	Cal. rej. prob. (Overflow model)	Cal. over. prob. (Overflow model)
2006	UCLH					
	NICU-HDU	12	0.3250	0.3770	0.3328**	0.0442
	SCBU-TC	15(8)	0.0380	0.0700	0.0565**	0.0135
2007	UCLH					
	NICU-HDU	16	0.1662	0.1970	0.1743	0.0227
	SCBU-TC	15(8)	0.0889	0.1272	0.0828	0.0444
2008	UCLH					
	NICU-HDU	17	0.1895	0.2515	0.1962	0.0553
	SCBU-TC	15(8)	0.1319	0.1781	0.1271	0.0510
2006	Barnet					
	NICU-HDU	6	0.1787	0.2684	0.1633	0.1051
	SCBU-TC	14(4)	0.0207	0.0313	0.0114	0.0199
2007	Barnet					
	NICU-HDU	6	0.1893	0.3034	0.1760	0.1274
	SCBU-TC	14(4)	0.0164	0.0191	0.0060	0.0131
2008	Barnet					
	NICU-HDU	6	0.1644	0.2687	0.1508	0.1179
	SCBU-TC	14(4)	0.0142	0.0225	0.0076	0.0149
2006	Whittington					
	NICU-HDU	11	0.0033	0.0061	0.0011	0.0050
	SCBU-TC	5(5)	0.2239	0.4849	0.2090	0.2759
2007	Whittington					
	NICU-HDU	11	0.0141	0.0345	0.0127	0.0218
	SCBU-TC	8(5)	0.2385	0.3705	0.1930	0.1775
2008	Whittington					
	NICU-HDU	12	0.0216	0.0011	0.0007	0.0004
	SCBU-TC	16(5)	0.0138	0.0303	0.0018	0.0285
2006	Royal Free					
	ITU	2	0.0917	0.0936	—	—
	SCBU	12	0.4072	0.4056	—	—
2007	Royal Free					
	ITU	2	0.0732	0.0726	—	—
	SCBU	12	0.1849	0.1847	—	—
2008	Royal Free					
	ITU	2	0.1468	0.1504	—	—
	SCBU	12	0.1558	0.1580	—	—
2007	Chase Farm					
	SCBU	10	0.0368	0.0386	—	—
2008	Chase Farm					
	SCBU	10	0.1078	0.1060	—	—

*Standard Erlang loss model **Overflow loss model as in Asaduzzaman et al. (2009)

flow probability at SCBU decreased from 0.2759 in 2006 to 0.0285 in 2008 mainly due to the increase of 11 SCBU cots. Increasing the number of cots at transitional care can only reduce rejection of admissions at SCBU at the expense of increasing overflow. Similarly, extra SCBU cots can reduce SCBU rejection probability, and possibly NICU-HDU rejection probability because of increased overflow to SCBU. However, these overflows may affect the quality of care. Therefore, while planning capacity allocation for the units one should aim at keeping the expected overflow to SCBU from NICU-HDU and from SCBU to TC to a minimum.

Table 3 shows the rejection probabilities for the level 1 units Chase Farm, and Royal Free. The Royal Free unit has a 2-cot ITU and 12 SCBU cots. In the unit, 1 extra cot at ITU and 3 extra cots at SCBU would decrease the rejection probabilities to 0.0385 and 0.0583 at ITU and SCBU respectively. The Chase Farm unit has been funded for 10 SCBU cots since 2006, and included in the SEND system since mid 2006. The unit would need 2 cots more to keep the rejection level at 0.0419.

5. Conclusion

In this paper, we developed an overflow loss network model for capacity planning in a perinatal network. Using the continuous time Markov chain framework, we derived the steady state distribution, expressions for rejection and overflow probabilities. By modelling the system of a perinatal network as a whole, we first decomposed the network into sub-networks, namely neonatal units, then we captured the movement of neonates between neonatal units each having different levels within the whole network. We calculated rejection and overflow probabilities at each level of care. The number of cots required to keep the rejection probability at a specific level was also derived. The overflow probability measures the overflow to temporary care which should be kept in mind while planning capacity for a unit.

The NCLPN has the basic aim to provide all levels of care for 95% of mothers and neonates within the network area where they live. UCLH is the only level 3 unit (perinatal network centre) in the NCLPN. At UCLH, the NICU-HDU rejection probability has decreased over the 3 years. This was achieved by both increasing the number of cots and decreasing LoS. At SCBU-TC, the number of cots required the same while demand increased significantly, thus increasing the rejection probability despite the slight decrease in LoS. With current arrival and LoS pattern (2008), the unit requires 5 extra cots in NICU-HDU to reduce rejection to less than 10%, and 4 extra cots in SCBU to reduce rejection to less than 5%. The rejection probabilities at NICU-HDU and SCBU show that the Barnet unit (level 2) requires 4 more cots to reduce rejection level to 5%. In the Royal Free unit, 1 extra cot at ITU and 3 cots at SCBU, and at Chase Farm, 2 extra cots would decrease rejection level to 5%.

We believe the model developed in this paper would help for solving capacity problem currently faced by many perinatal networks in the UK. Although the model was formulated for a specific network, it can be extended for any perinatal network since all perinatal networks in the UK consist of level 1, 2 and 3 units with NICU/HDU/SCBU-TC, and have similar structure. However, networks may vary in their number of units and also transfers and back transfers between units (known as “routing” and “alternative routing” in the telecommunication literature). Since we have used a decomposition approach which takes into account the dependency between units through arrivals, the number of units and complex transfers within a network will not affect the derivation of steady state distribution, rejection probability etc. Therefore, the model framework can be applied to any network. Moreover, the model formulation can be applied to plan capacity for other areas such as computer, teletraffic and other communication networks. If covariate data such as expected length

of stay in the units, number of staff, population density, proportion of women in reproductive age groups in area, fertility rate in area, proportion of these women with certain health conditions etc. are available alternative statistical methods such as multilevel modelling approach can be applied to predict overflow and rejection probabilities. The proposed model assumes that arrival and LoS pattern are Markovian which was satisfactory for the network considered here. However, future work is required where alternative non-Markovian arrival and LoS patterns will be considered. Given the importance of capacity planning for perinatal networks in the UK to reduce rejection level in the neonatal hospitals which results in high risk of neonatal mortality and cost, the findings of this paper should be of interest to the Department of Health (DH), health care researchers, and perinatal network managers.

Acknowledgements

We would like to thank to the clinical lead at each neonatal unit in the NCLPN, especially Dr. Jane Hawdon at UCLH, for their help to get honorary contract with each unit, data gathering, and useful discussions.

A. Derivation of steady state behaviour for UCLH/Barnet/Whittington

For $\mathbf{n}, \mathbf{n}' \in \mathbf{S}$, the transition rates $\mathbf{Q} = q(\mathbf{n}, \mathbf{n}')$ are given by

$$q(\mathbf{n}, \mathbf{n}') = \begin{cases} \lambda_1 & \mathbf{n}' = \mathbf{n} + \mathbf{I}_1 \text{ or } \mathbf{n}' = \mathbf{n} + \mathbf{I}_{12} \text{ if } n_1 + n_{21} = c_1 \\ \lambda_2 & \begin{cases} \mathbf{n}' = \mathbf{n} + \mathbf{I}_2 \text{ or } \mathbf{n}' = \mathbf{n} + \mathbf{I}_{21} \text{ if } n_{12} + n_2 = c_2 \\ \mathbf{n}' = \mathbf{n} + \mathbf{I}_{23} \text{ if } n_{12} + n_2 = c_2 \text{ and } n_1 + n_{21} = c_1 \end{cases} \\ n_i \mu_i & \mathbf{n}' = \mathbf{n} - \mathbf{I}_i \\ n_{ij} \mu_{ij} & \mathbf{n}' = \mathbf{n} - \mathbf{I}_{ij} \\ n_{ij} \gamma_i & \mathbf{n}' = \mathbf{n} - \mathbf{I}_{ij} + \mathbf{I}_i \\ 0 & \text{otherwise,} \end{cases}$$

where \mathbf{I}_i and \mathbf{I}_{ij} denote unit vectors.

The equilibrium distribution of \mathbf{X} is the unique distribution $\pi(\mathbf{n}), \mathbf{n} \in \mathbf{S}$ that satisfies the *station*

balance equation,

$$\left. \begin{aligned}
& \left[\lambda_1 (\mathbf{1}_{\{n_1+n_{21}<c_1\}}(\mathbf{n}) + \mathbf{1}_{\{n_1+n_{21}=c_1, n_1+n_{12}<c_2\}}(\mathbf{n})) + \lambda_2 (\mathbf{1}_{\{n_{12}+n_2<c_2\}}(\mathbf{n}) \right. \\
& \quad + \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}<c_1\}}(\mathbf{n}) \\
& \quad + \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}=c_1, n_{23}<c_3\}}(\mathbf{n})) \\
& \quad + \sum_{i=1}^2 n_i \mu_i + \sum_{i,j \in \{1,2,3\}} n_{ij} \mu_{ij} \\
& \quad \left. + \sum_{i,j \in \{1,2,3\}} n_{ij} \gamma_i \mathbf{1}_{\{(n_i+n_{ij})|n_{ji}<c_i\}}(\mathbf{n}) \right] \cdot \pi(\mathbf{n}) \\
& = \lambda_1 [\pi(\mathbf{n} - \mathbf{I}_1) + \pi(\mathbf{n} - \mathbf{I}_{12}) \mathbf{1}_{\{n_1+n_{21}=c_1, n_{12}+n_2<c_2\}}(\mathbf{n})] \\
& \quad + \lambda_2 [\pi(\mathbf{n} - \mathbf{I}_2) + \pi(\mathbf{n} - \mathbf{I}_{21}) \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}<c_1\}}(\mathbf{n}) \\
& \quad + \pi(\mathbf{n} - \mathbf{I}_{23}) \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}=c_1, n_{23}<c_3\}}(\mathbf{n})] \\
& \quad + \sum_{i=1}^2 (n_i + 1) \mu_i \pi(\mathbf{n} + \mathbf{I}_i) + \sum_{i,j \in \{1,2,3\}} (n_{ij} + 1) \mu_{ij} \pi(\mathbf{n} + \mathbf{I}_{ij}) \\
& \quad + \sum_{i,j \in \{1,2,3\}} (n_{ij} + 1) \gamma_i \pi(\mathbf{n} - \mathbf{I}_i + \mathbf{I}_{ij}),
\end{aligned} \right\} \quad (2)$$

normalised such that $\sum_{\mathbf{n} \in \mathbf{S}} \pi(\mathbf{n}) = 1$ and where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function of the event or set $\{\cdot\}$, i.e., $\mathbf{1}_{\{\cdot\}}(\mathbf{n})$ equals to 1 if $\mathbf{n} \in \{\cdot\}$; 0 otherwise, for $\mathbf{n} \in \mathbf{S}$.

Assuming $\mu_{ij} = \mu_i$, $i, j \in 1, 2, 3$ and $\gamma_i = \infty$, $i = 1, 2, 3$, station balance equation becomes

$$\left\{ \begin{aligned}
& \pi(\mathbf{n}) n_1 \mu_1 \mathbf{1}_{\{n_1+n_{21}<c_1\}}(\mathbf{n}) + \\
& \pi(\mathbf{n}) n_{12} \mu_1 \mathbf{1}_{\{n_1+n_{21}=c_1, n_1+n_{12}<c_2\}}(\mathbf{n}) + \\
& \pi(\mathbf{n}) n_2 \mu_2 \mathbf{1}_{\{n_{12}+n_2<c_2\}}(\mathbf{n}) + \\
& \pi(\mathbf{n}) n_{21} \mu_2 \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}<c_1\}}(\mathbf{n}) + \\
& \pi(\mathbf{n}) n_{23} \mu_2 \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}=c_1, n_{23}<c_3\}}(\mathbf{n}) + \\
& \pi(\mathbf{n}) \lambda_1 \mathbf{1}_{\{n_1+n_{21}<c_1\}}(\mathbf{n}) + \\
& \pi(\mathbf{n}) \lambda_1 \mathbf{1}_{\{n_1+n_{21}=c_1, n_1+n_{12}<c_2\}}(\mathbf{n}) + \\
& \pi(\mathbf{n}) \lambda_2 \mathbf{1}_{\{n_{12}+n_2<c_2\}}(\mathbf{n}) + \\
& \pi(\mathbf{n}) \lambda_2 \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}<c_1\}}(\mathbf{n}) + \\
& \pi(\mathbf{n}) \lambda_2 \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}=c_1, n_{23}<c_3\}}(\mathbf{n})
\end{aligned} \right\} = \quad (3)$$

$$\left\{ \begin{aligned}
& \pi(\mathbf{n} - \mathbf{I}_1) \lambda_1 \mathbf{1}_{\{n_1+n_{21}<c_1\}}(\mathbf{n}) + \\
& \pi(\mathbf{n} - \mathbf{I}_{12}) \lambda_1 \mathbf{1}_{\{n_1+n_{21}=c_1, n_1+n_{12}<c_2\}}(\mathbf{n}) + \\
& \pi(\mathbf{n} - \mathbf{I}_2) \lambda_2 \mathbf{1}_{\{n_{12}+n_2<c_2\}}(\mathbf{n}) + \\
& \pi(\mathbf{n} - \mathbf{I}_{21}) \lambda_2 \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}<c_1\}}(\mathbf{n}) + \\
& \pi(\mathbf{n} - \mathbf{I}_{23}) \lambda_2 \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}=c_1, n_{23}<c_3\}}(\mathbf{n}) \\
& \pi(\mathbf{n} + \mathbf{I}_1) n_1 \mu_1 \mathbf{1}_{\{n_1+n_{21}<c_1\}}(\mathbf{n}) + \\
& \pi(\mathbf{n} + \mathbf{I}_{12}) n_{12} \mu_1 \mathbf{1}_{\{n_1+n_{21}=c_1, n_1+n_{12}<c_2\}}(\mathbf{n}) + \\
& \pi(\mathbf{n} + \mathbf{I}_2) n_2 \mu_2 \mathbf{1}_{\{n_{12}+n_2<c_2\}}(\mathbf{n}) + \\
& \pi(\mathbf{n} + \mathbf{I}_{21}) n_{21} \mu_2 \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}<c_1\}}(\mathbf{n}) + \\
& \pi(\mathbf{n} + \mathbf{I}_{23}) n_{23} \mu_2 \mathbf{1}_{\{n_{12}+n_2=c_2, n_1+n_{21}=c_1, n_{23}<c_3\}}(\mathbf{n}) +
\end{aligned} \right\}. \quad (4)$$

The product form expression can be verified by the above equation, or each of the more *detailed balance equations* (3)(k) = (4)(k), $k = 1, 2, \dots, 10$, separately, and with the normalizing constant.

References

- Asaduzzaman, M. and Chausalet, T. J. (2008) Modelling and performance measure of a perinatal network centre in the United Kingdom. In *Proceedings of the 21st IEEE International Symposium on Computer-Based Medical Systems*, pp. 506–511.
- Asaduzzaman, M., Chausalet, T. J. and Robertson, N. J. (2009) A loss network model with overflow for capacity planning of a neonatal unit. *Ann. Oper. Res.*, Forthcoming (published online), DOI 10.1007/s10479-009-0548-x.
- Baskett, F., Chandy, K. M., Muntz, R. R. and Palacios, F. G. (1975) Open, closed, and mixed networks of queues with different classes of customers. *J. ACM*, **22**, 248–260.
- Bebbington, M., Pollett, P. and Ziedins, I. (1998) Two-link approximation schemes for linear loss networks without controls. *J. Korean Math. Soc.*, **35**, 539–557.
- Bekker, R. and de Bruin, A. M. (2009) Time-dependent analysis for refused admissions in clinical wards. *Ann. Oper. Res.*, Forthcoming, DOI 10.1007/s10479-009-0570-z.
- Bliss (2007) Too little, too late? BLISS - The Premature Baby Charity. (Available from <http://www.bliss.org.uk/page.asp?section=677§ionTitle=Too+little%2C+too+late%3F>).
- DH (2003) Report of the Neonatal Intensive Care Services Review Group. Department of Health. (Available from http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4018744.pdf).
- Gordon, W. J. and Newell, G. F. (1967) Closed queuing systems with exponential servers. *Oper. Res.*, **15**, 254–265.
- Griffiths, J. D., Price-Lloyds, N., Smithies, M. and Williams, J. (2006) A queueing model of activities in an intensive care unit. *IMA J. Mangt. Math.*, **17**, 277288.
- Guerin, R. and Lien, L. (1990) Overflow analysis for finite waiting room systems. *IEEE Tran. Comm.*, **38**, 1569-1577.
- Heyman, D. P. and Sobel, M. J. (1982) *Stochastic Models in Operations Research*. Vol. I, McGraw Hill, New York.
- Jackson, J. R. (1957) Networks of waiting lines. *Oper. Res.*, **5**, 518–521.
- Jackson, J. R. (1963) Jobshop-like queueing systems. *Mang. Sc.*, **10**, 131–142.
- Jagerman, D. L. (1984) Methods in traffic calculations. *AT&T Bell Lab. Tech. J.*, **63**, 1291-1309.
- Kaufman, L., Morrison, J. A. and Seery, J. B. (1981) Overflow models for dimension PBX feature packages. *Bell Sys. Tech. J.*, **60**, 661-676.
- Kelly, F. P. (1975) Network of queues with customers of different types. *Adv. App. Prob.*, **12**, 542–554.
- Kelly, F. P. (1979) *Reversibility and stochastic networks*. Chichester: Wiley

- Kelly, F. P. (1986) Blocking probabilities in large circuit-switched networks. *Adv. App. Prob.*, **18**, 473–505.
- Kobayashi, H. and Mark, B. L. (2002) Generalized loss models and queueing-loss networks. *Intl. Trans. Oper. Res.*, **9**, 97–112.
- Marlow, N. and Gill, A. B. (2007) Establishing neonatal networks: the reality. *Arch. Dis. Child.*, **92**, F137-F142.
- Kortbeek, N. and van Dijk, N. M. (2007) On dimensioning intensive care units. *AENORM*, **57**, 22–26.
- Litvak, N., van Rijsbergen, M., Boucherie, R. J. and van Houdenhoven, M. (2008) Managing the over flow of intensive care patients. *European J. Oper. Res.*, **185**, 998-1010.
- NAO (2007) Caring for Vulnerable Babies: The reorganisation of neonatal services in England. National Audit Office. (Available from http://www.nao.org.uk/publications/0708/caring_for_vulnerable_babies.aspx).
- RCPCH (2007) Modelling the future: A consultation paper on the future of children’s health services. Royal College of Paediatrics and Child Health. (Available from <http://www.rcpch.ac.uk/Policy/ServiceReconfiguration/Modelling-the-Future>).
- Redshaw, M. and Hamilton, K. (2005) A Survey of Current Neonatal Unit Organisation and Policy. BLISS - The Premature Baby Charity & National Perinatal Epidemiology Unit, University of Oxford. (Available from <http://www.npeu.ox.ac.uk/downloads/reports/bliss-final-report.pdf>).
- Reiser, M. and Kobayashi, H. (1975) Queueing network with multiple closed chains: Theory and computational algorithms. *IBM J Res. Dev.*, **19**, 283–294
- Ross, K. (1995) Multiservice Loss Models for Broadband Telecommunication Networks. Springer-Verlag.
- Simul8 (2000) Simul8 Manual and Simulation Guide. Simul8 Corporation.
- van Dijk, N. M. (1993) Queueing networks and product forms: A systems approach. Chichester: Wiley.
- van Dijk, N. M. and van der Sluis, E. (2009) Call packing bound for overflow loss systems. *Perform. Eval.*, **66**, 1–20.
- Wilkinson, R. (1956) Theories for toll traffic engineering in the USA. *Bell Sys. Tech. J.*, **35**, 421-514.