FIRE DEPARTMENT DEPLOYMENT ANALYSIS – MODELLING IN PUBLIC POLICY

Produced from material originally prepared by Clive Savory,
The Open University

Introduction

Fire is a strange beast for humans. On the one hand we control it to do our bidding, as a heat and power source for example, but on the other we fear it because it can get out of control and wreak havoc on us and our environment.

This case study is predominantly based on the book *Fire Department Deployment Analysis* (Walker *et al*., 1979), which details a series of studies carried out by the New York City–RAND Institute between 1969 and 1975, principally as a partnership between RAND and the Fire Department of the City of New York (FDNY) but also including other projects in Trenton and New Jersey City, New Jersey, Wilmington in Delaware, and Tacoma in Washington.

The RAND Corporation at that time was concerned with investigating the use of the then young discipline of operational research and, more specifically, systems analysis. It had used operation research techniques on many projects in defence and industrial operations, but this project was the first where it was to be applied to urban problem solving and services. Central to the work of the project was the development of models through simulation and empirical validation resulting from

- the complexity of the situation;
- the criticality of fire service provision, and
- the potential costs of large scale studies.

You will come across various pieces of mathematics in this text. It is not necessary for you to understand them, rather just to appreciate the basis of a lot of the work. Indeed, we could have included more than we have, as the book *Fire Department Deployment* contains large sections on the background to their work. If you are interested in exploring this facet in greater detail, I recommend you read the book for yourself.

Background

New York, in common with other large cities, has always suffered from fires, but incidents had been climbing at an alarming rate since 1960. Between 1956 and 1969 the number of fire alarm calls had grown from 69 000 to 240 000 per year (Blum, 1971). Why this was so is complex to explain, but factors included building conditions, construction techniques, population growth, changing social attitudes and rising social tensions in areas of urban decay. The situation eventually reached a point where the number of accidental fires within reasonably maintained buildings was being outstripped by other sources of alarms as the primary source of realized demand for the FDNY. These included false alarms, rubbish fires, fires in abandoned buildings, arson and non-fire emergencies. Much of the increased demand was focused in areas of social deprivation, where the housing stock was also poor. The increase in, and changing composition of, this *realized* demand for their services was causing the FDNY serious deployment problems.

The issues facing the FDNY were not all demand led; although demand had traditionally played the major role in the planning of fire-fighting resources, the service had also to consider the changing nature of the 'latent' demand, fire prevention services and code compliance responsibilities. Latent demand is the potential demand arising from the existence of the infrastructure and processes by which humans choose to organize their everyday lives, and is expressed as hazards. Improving technologies had led to changes in the nature of the risk and the consequences of a fire event. As developers and architects experimented with new designs and materials it became increasingly possible to site large numbers of people in the same enclosed space, for example in high-rise office buildings and apartment blocks. The use of new materials also began to change the nature of fires in the way that they burnt, and the fumes given off.

The increasing strain placed on the fire-fighting services was not just a New York issue; the National Commission on Fire Prevention and Control echoed this as a national problem in its 1973 report *America Burning*.

Historically the fire service had developed to cope with ever-rising demands by increasing the number of fire companies which, at an annual cost of \$250 000–\$750 000 each, depending on manning levels, put severe pressure on budgets. By 1972 the number of fire companies had increased to 375, and during the period 1957–1971 the FDNY budget increased from \$99 million to \$311 million (Blum, 1971). By 1975 this had further increased to \$375 million (Ignall *et al.*, 1975). Additionally, it was becoming clear that the traditional approach did not resolve the problems being faced by the FDNY because other factors had unforeseen effects. For example, in 1967 a new fire company was created, and housed in the same fire house as the existing busiest engine. The aim was to relieve pressure and workload from that engine, but it did not work out as expected. A subsequent analysis one year later revealed that the fire house now had two very busy fire engines. The number of incidents had not dramatically increased during the year; the high call-out rate was actually the result of existing predetermined rules used to deploy resources when an alarm was received.

There was a clear need to do something about the situation, but the lack of relevant data and analytical techniques made the development of a new effective policy difficult, if not impossible. It was against this backdrop that Fire Commissioner Robert Lowery and Chief of Department John O'Hagan requested the resources for analytical assistance from Mayor Lindsay.

The methodology

The framework used by the RAND–New York team is shown in Figure 12. You will note its resemblance to the 'hard systems' approach you came across in Block 1 (Figure 32 from Block 1 is reproduced here as Figure 13).

The material following is structured in terms of the research methodology shown in Figure 12 up to and including step 6. The section titled 'Issues addressed' contains material relevant to steps 1–7. It was not possible to include material for step 8 as this was not covered in the book, probably because the New York City–RAND Institute was closed in late 1975. RAND did not undertake any further large scale fire studies after its closure.

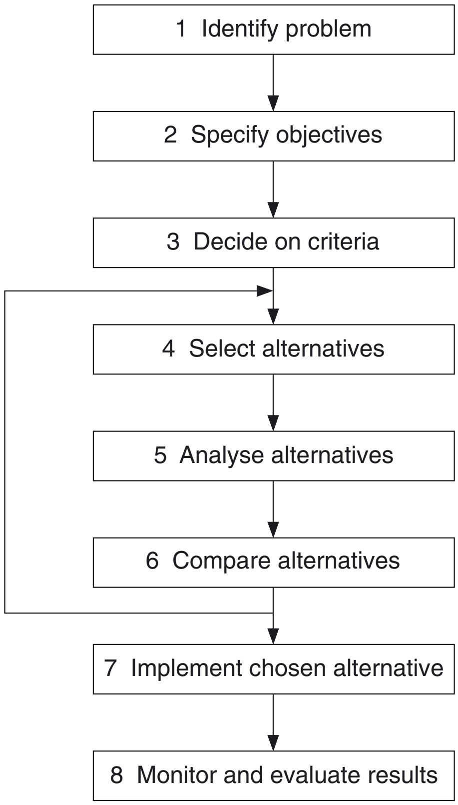

Figure 12 Steps in a systems analysis study. (Source: Walker *et al.*, p. 70)

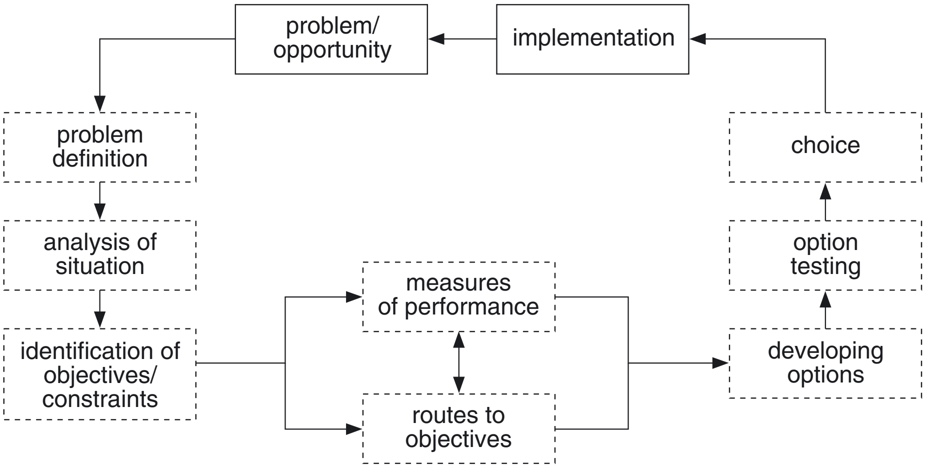

Figure 13 Stages of the hard systems approach

Step 1: Identify problem

The primary problem for any fire service is to ensure that fire-fighters and their equipment are in the right place at the right time. This can be further decomposed into strategic and tactical sub-problems.

Strategic problems relate to the longer-term issues of resource allocation and location, both of which involve long-term planning. Three sub-problems require addressing:

- How many fire companies should be on duty at any one time? This is affected by budgetary considerations as well as the temporal nature of the requirement; there is a daily and seasonal variation in demand.
- How should the resources be allocated geographically? Issues to consider would include, for example, notions of equitableness as well as the nature of realized and latent demand. The allocation of a fire-fighting resource clearly has political and social dimensions along with the more obvious requirement to fight fires.
- Where should the resources be located within a region? Consideration should be given to issues such as land availability, existing road configurations and traffic conditions.

Unlike strategic problems, tactical problems relate to the daily operation of the fire service. For this study on deployment the related sub-problems are:

- What is the most effective dispatch policy? That is, what types of resource should be dispatched on receipt of an alarm and which fire companies should be employed for the fire event?
- How should ‘gaps’ in cover be managed? This relates to relocation and redeployment policies, the aims of which are to ensure adequate coverage throughout the entire region at all times.

Step 2: Specify objectives

The primary objective for a fire system is to protect lives and property from fire. However, there were political and financial constraints placed upon the FDNY, reflected in the appointment of senior personnel, the non fire-fighting duties imposed and budgets available. In addition, objectives might change during a study period, and this is more likely the greater the study duration. For example, towards the end of 1971 the City of New York faced a budget crisis and the FDNY was asked to present plans to cut its budgets. Other objectives were developed for specific projects – determining the most effective dispatch policy for example.

Step 3: Decide on criteria

Given the primary objective in step 2, any changes in policy by the City of New York should result in either no effect or a change in the number of fire-related fatalities and the value of property lost due to fire (Walker *et al.*, page 80). This is a somewhat restricted view in that it does not address injuries caused by fire, but it at least suggests two metrics for evaluation purposes.

The problem the RAND team had was the lack of reliable data. This was not a problem peculiar to New York; the same issue was raised in *America Burning*. Although the FDNY had been logging and keypunching data on every fire alarm since 1962 the data set was not sufficient in itself to enable reliable evaluation of the effect on the primary objective of changes in policy. The RAND team, therefore, decided to use numerical proxy measures whilst acknowledging the paucity of the data available and its effect on the depth of analysis achievable.

> So the analysts wishing to do practical, applied fire research, while forced to accept the present lack of direct performance measures, may nevertheless do important work using proxies, such as response time.
>
> (p. 81)*

The proxies chosen for the project were:

- response time, including dispatching, turnout, travel and setup;
- coverage, a measure of how often a specific location has nearby units;
- availability, a measure of the time when nearby units may be unavailable to respond;
- initial response adequacy, a measure of how appropriate the response was in terms of the types and number of units required;
- fire company workloads;

---

* This and subsequent page references are from Walker *et al.* (1979).

- cost – as the capital cost of providing a fire house is small compared with staffing costs this is a measure of the number of personnel required on the payroll.

Clearly, collecting the data for the whole of New York would have been both time consuming and expensive had the data not already existed. Where there were gaps in the data the most common methods employed were to conduct trials and experiments within constrained locations. For example, rather than assess the travel time to an incident for all the fire companies, an activity that would have created a large administrative burden, a trial was set up using 15 units. The busiest units were excluded from the trial as it was felt to be too burdensome for them. Participating units logged times, distances and locations for responses from the fire house only; other responses, such as call outs whilst returning from an incident, were not included.

Step 4: Select alternatives

In the early stages of the project, alternative policies for consideration were generated by senior FDNY personnel. This was only to be expected as they, and not the analysts, had the in-depth knowledge of the fire service and its political context. Because of this, there is little in the study which throws light on this complex process.

As the project progressed, however, the project analysts had more of an input as their knowledge grew and the FDNY and City officials had growing confidence in the project outcomes.

Steps 5 and 6: Analyse and compare alternatives

Once alternatives for consideration are selected, the consequences require analysis and comparison before final decisions and rollout can be made. One possible method is to make a change and then observe the results over a given period. When policy changes are expensive, time consuming and relate to the loss of life and property such an approach is infeasible, both economically and politically. Clearly, the criticality of the service provided by a Fire Department makes such an approach problematic. To overcome this, the project used modelling and simulation.

Static and dynamic models were developed to describe the current state and predict into the future, and were then tested against empirical data. The models created were:

- for estimating travel distance and time;
- for analysing demand;
- for availability and dispatch.

It is important to realise that the models were used to inform the decision-making process; they were not the sole source of information. Underpinning the process of modelling was the creation of a simulation program representing the Fire Service operations in their totality. The simulation was used for developing and testing models and is described later in this case study.

Models for estimating travel distance and time

The time it takes for resources to arrive at a fire was seen as a crucial element in the reduction of loss of life and property damage. From observation it was found that, typically, the greatest element of response time was that taken to travel from the fire house to the incident. It was seen at an early stage, therefore, that estimates of fire service travel distance were fundamental to understanding many of the problems in engineering the fire protection system, because travel time is related to distance. It was reasoned that the quicker an engine attended an incident, the greater the chance of reducing losses to life and property. The project considered times and distances between any two locations, and also averages within an area.

Location-to-location times and distances

As part of the study, several cities were observed, and it was found that the relationship between travel time and travel distance was that shown by the curve in Figure 14.

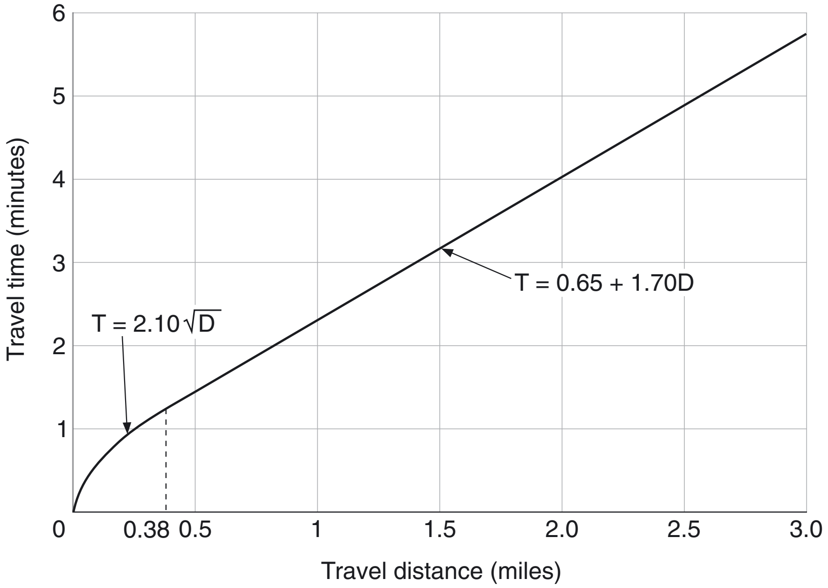

Figure 14 The relationship between travel time and travel distance – average of empirical results from Denver, Trenton, Wilmington and Yonkers. (Source: p. 166)

This can be:

$$E\left(T_{ij}\right)=\begin{cases} c\sqrt{D_{ij}} & D_{ij}\le d \\ a+bD_{ij} & D_{ij}\ge d \end{cases} \qquad \text{(1) p. 165}$$

where

$E(T_{ij})$ is the time required to travel the distance between points i and j, D_{ij}
a, b, c and d are parameters which can be estimated empirically.

Based on the assumption that the area under study has a regular road network, as found in American cities, the curve in Figure 14 shows the relationship in two parts: at distances below 0.38 miles it is a square-root relationship while above 0.38 miles it takes on a linear form. The distance D was measured in miles and the time T in minutes. The parameters a, b, c and d when substituted gave:

$$E\left(T_{ij}\right)=\begin{cases} 2.10\sqrt{D_{ij}} & D_{ij}\le 0.38 \text{ miles} \\ 0.65+1.70D_{ij} & D_{ij}\ge 0.38 \text{ miles} \end{cases} \qquad \text{(2) p. 166}$$

It should be noted that the values of the parameters a, b, c and d were found to differ between cities, and so the project set experiments to monitor travel times for specific cities. The New York City experiment yielded 1772 responses and gave similar but slightly different results, as shown in Figure 15.

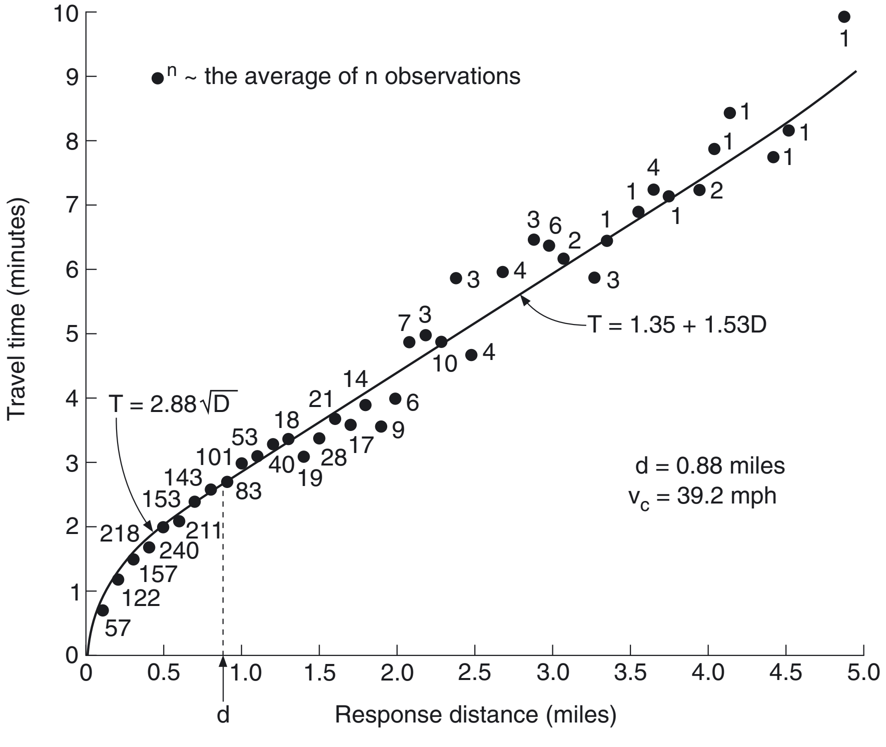

Figure 15 Travel time against distance in New York City. (Source: p. 174)

The values for *a*, *b*, *c* and *d* were calculated to produce the following expression:

$$E(T) = \begin{cases} 2.88\sqrt{D} & D \le 0.88 \text{ miles} \\ 1.35 + 1.53D & D \ge 0.88 \text{ miles} \end{cases} \qquad \text{(3) p. 173}$$

A further level of sophistication was introduced by studying the variation in response speed within a 24-hour period, necessary if a policy considers variations in coverage and response during the day. An intuitive hypothesis would be that the morning and evening rush hours would reduce response time, i.e. increase the time taken to travel. However, the New York data indicated that, whilst there were differences between four time periods – morning rush hour, evening rush hour, daylight (excluding the rush hour) and dark – these were not marked (Table 2). Similar results were found in the four other cities studied.

Table 2 Response speeds in New York City by time of day

Time period	Average speed (miles h^{-1})	Standard deviation of speed
0000–0200	19.2	7.3
0200–0400	17.7	6.8
0400–0600	18.0	7.7
0600–0800	17.7	7.1
0800–1000	16.2	6.7
1000–1200	18.3	7.6
1200–1400	18.9	6.9
1400–1600	18.4	7.3
1600–1800	17.0	7.6
1800–2000	18.9	7.3
2000–2200	18.5	7.5
2200–2400	18.8	7.2
0500–2000 (daylight)	18.2	7.3
2000–0500 (dark)	18.6	7.3
0800–0900 (morning rush hour)	14.3	5.2
1630–1730 (evening rush hour)	18.2	7.6
All hours	18.3	7.3

Source: p. 178

Intra-region average travel distances and time

Considered 'one of the major results of our research' (p. 177), the study determined a simple method of calculating intra-regional average travel distances and time. This calculation has clear relevance when considering policies that alter the geographical pattern of fire houses.

The average could be determined by estimating location-to-location distances for each possible combination of fire house to fire event, and then applying a weighting based on the proportion of the total trips within a region it represents. Rather than computing the distance to possible fire events, the study used the only data available to them at that time – the location of alarm boxes. However, the calculation would be time and resource consuming given the level of technology available. The researchers derived a simple mathematical expression to show the expected distance between the points in the region at which fires occur and the location of the closest available engine company. This is given by:

$$E\left(D_n\right) = k_n\sqrt{A/E(N)} \qquad \text{(4) p. 185}$$

where

k_n = a constant of proportionality

A = the area of a region

$E(N)$ = the number of fire houses having at least one engine available to respond.

A number of assumptions were made:

- alarms are distributed randomly but with uniform probability density throughout the region of interest;
- fire houses are spread either in a regular pattern or randomly throughout the region of interest;
- boundary effects are insignificant;
- units are always available to respond;
- fire companies travel either in a straight line between two points or on a dense rectangular grid of streets.

Whilst, as with any general model, each assumption may not hold true for a particular region the researchers, using simulations and empirical data, concluded that the square-root law still provided a useful estimate of average response times.

Estimating intra-regional travel time is similar to that for location-to-location times discussed previously. Using historical data and equation (1) above, the average regional travel time for the first arriving engine company can be estimated by

$$\overline{T}_1 = \sum_b \sum_m G\left(d_{mb}\right) p_{mb} \qquad \text{(5) p. 199}$$

A simpler approximation was derived by using the square-root law and equation (1):

$$E(T_1) = \begin{cases} c\sqrt{k_1}\left[A/E(N)\right]^{1/4} & k_1\sqrt{A/E(N)} \le d \\ a + bk_1\left[A/E(N)\right]^{1/2} & k_1\sqrt{A/E(N)} \ge d \end{cases} \qquad \text{(6) p. 200}$$

The general form of this equation was shown to be valid against simulated data from several regions in New York City.

Models of demand for fire services

Knowledge of the potential demand for a service has important implications for the provision of a service, in terms of both allocating resources and dispatching adequate resources to an event. But the occurrence of a fire event is only part of the issue for a fire service; ideally some indication of the likely severity is also available to those making the dispatch decisions. The project was particularly concerned with what the authors termed ‘serious events’, defined in terms of the likelihood that the event occurs in an occupied structure and requires at least one engine company and two ladder companies to attend.

The researchers therefore considered two issues:

- the probability of an alarm occurring;
- the likelihood that an alarm was serious.

It is clear that the number of variables that affect alarm rates is large, so the RAND team adopted a principle of ‘parsimony in parameters’. This suggested that variables should only be included in the model if they made a *significant* contribution to the accuracy of prediction. This intellectual ‘sleight of hand’ is not uncommon but does need to be borne in mind.

That the situation required modelling can be understood by considering how an alarm is raised. If the alarm is raised via the telephone a dispatcher can at least gather some rudimentary information to enable them to decide how many, and of what type, of resources to send. However, alarms can also arrive from a street box, and in such circumstances the dispatcher will only know the general location. In 1970 street box alarms accounted for 60 per cent of all alarms, and 43 per cent of structural fire alarms, in New York City. The lack of information available to a dispatcher can be seen to be an increasing area for concern as the number of incidents increased.

Using historical data collected from incident records the area is divided into fire demand regions. In constructing the regions a number of constraints should be fulfilled for each:

- the distribution of types of alarm is the same throughout a region;
- a region has the same, or similar, type of hazard present;
- there are sufficient fire companies to provide meaningful analysis.

The definition of a demand region is likely to be a compromise between this ideal and existing administrative districts. Where data are required for a smaller region, the researchers considered that using a proportion of the demand region data would be adequate. The historical data collected showed a temporal and seasonal variation in alarm rates within any region, and that this differed between regions.

To model the arrival of incoming incidents, the team showed that a modified Poisson distribution was theoretically appropriate. The influence of location, time of day and season were taken into account to produce the probability distributions. Empirical testing of data for the Bronx, covering the period 1967–70, justified the use of the Poisson distribution.

To arrive at a model for estimating the seriousness of an alarm the researchers:

1. estimated the probability that an alarm was for a fire in an occupied structure;
2. modelled the temporal and seasonal effects, and the trend for an alarm to be serious for a demand region;
3. combined points 1 and 2.

Data from the Bronx for the period 1967–69 were used to develop the estimates. The combined model was then compared against data for 1970 and shown to be a good fit.

Models of availability and dispatch

The distance and time models above go only so far in aiding the analysis of options. What they do not show is the effect of the daily operation of the fire service, i.e. they do not take account of the fact that fire companies fight fires and therefore may not always be available to answer a call out. To understand the effects of daily operation the study used queuing theory and Markov decision theory.

Queuing theory is concerned with establishing how a service copes with demand from customers. A simple example is the model of a single checkout in a supermarket, with customers arriving at a specific rate. In modelling this situation queuing theory uses estimates of service time (the time to put the customer's shopping through the till) to predict the length of the queue and how long customers must wait in the queue. Additional factors such as variability of the quantity of the shopping between customers, or even the effects of customers getting fed up with waiting and leaving (balking) the queue can be included in the model.

For the RAND Fire Project queuing theory was applied, taking each fire crew within a fire station as a service. By estimating the rate of calls, average time spent on a call etc., the study modelled the availability and likely delay in dispatching crews to incidents. The model was further refined by including past data on specific locations, for example if a specific location required a certain number of fire crews dispatched in the past.

The development of the queuing models was felt to be important because, although the situation was often simplified, the models provided insights for decision making. Even when situations became more complex the general insights gained from the simple models often continued to hold true. Developing the models also provided clear guidance in what types of data are needed to analyse fire deployment options.

Queuing models were developed to improve the understanding of:

- the number of crews busy in a region at any time;
- the decisions on which crews to dispatch;
- the decisions on how many crews to dispatch.

Five descriptors of the queuing system were used:

- arrival process, expressed as the probability of an alarm being received;
- service time, the time a ‘server’ is handling a customer. Depending on the situation being analysed, a server could be, for example, a fire company or a dispatcher in the communications office;
- number of servers – for most models these are the fire companies;
- maximum number of customers who can wait for service, an expression of the ability of a system to handle customers;
- states of the system, a declaration of the conditions or states of the system that are of interest to the analyst. For example, an analyst may be interested in alarms in progress. State 0 would describe a situation where no alarms were active, state 1 would be one alarm active, and so on. This is a simple example where the description of the state only includes active alarms. The greater the number of descriptions for each state the more complex the analysis.

The probabilities for both the arrival process and the service times were generated from a modified Poisson distribution. Although there were issues with the use of a Poisson distribution, it was felt overall that it provided the most suitable solution as there was no significant difference between theoretically derived results and reality. Empirical data from the Bronx indicated that alarm rates did indeed approximate a Poisson distribution.

Fire operations simulation model

The use of simulation was important in the context of the RAND studies in that it:

- helped evaluate complex deployment policies that exceeded the relatively simple constraints of the simple analytical models;
- could be used to test the validity of the analytical models that had been created;
- was useful for explaining, and gaining acceptance of, policy decisions.

For decision makers, it was an advantage that simulation gave a 'lifelike' test of policy proposals. Discussion of a simulation and its results was seen as more attractive than the discussion of a complex mathematical model, with the simulation giving added credibility to the proposed policy changes.

> Simulation may also be useful in providing governmental decision makers with the necessary confidence that a proposed new policy will actually work in the way predicted. For example, in Denver the simulation model was used after other methods of analysis had suggested that a reduction from 44 to 39 fire companies, together with construction of six new fire stations, could provide almost the same level of fire protection as the then-current configuration. However, these conclusions were based on the assumption that every fire company would almost always be available for dispatch at its station. At this point, the simulation model was run to see whether the real performance of the new configuration would be as good as anticipated, taking into account the unavailability of companies.
>
> Denver's Deputy Finance Director has said that 'one of our main concerns was what effect the changed configuration of fire companies would have on response times if the number of fire alarms continued to increase.' The RAND simulation told us that, even at double our present peak alarm rate, there would be no significant deterioration in service. This was important in presenting our recommendations to the mayor and the City Council.
>
> (pp. 497–498)

Simulation was seen as an experimental tool rather than as a means of producing optimum solutions. It used either a sequence of pre-programmed events (e.g. sequence of events in a real-life situation) or a set of computer-generated events. In the case of the computer-generated set of events, a computer program would sequence events based on statistical probabilities, often a Poisson distribution.

The advantages of the simulation were that:

- it allowed a detailed analysis of how a sequence of events affected the fire service system;
- it allowed a detailed representation of the region including road layouts, traffic densities, number and types of fire stations, high risk buildings etc;
- different situations could be studied, for example the steady state operation of the fire service or specific disaster scenarios;
- policy decision makers could take into account multiple causality factors in any situation.

The main disadvantage of the simulation was its cost. This was not only related to the hardware and personnel costs associated with the system; the lack of readily available data required that some additional expenditure was necessary for data acquisition. However, the costs were much less than if a modelling and simulation approach had not been used.

The fire operations simulation model comprises three phases:

- an event generator – consisting of input variables, deployment policies, an incident generator or exogenous events tape;
- the simulation program itself;
- a post-simulation analysis program which could provide data for a number of questions such as average response time, response adequacy, coverage, workload, relocation and many others.

Three main categories of input exist for the Event Generation stage. The input variables are the main contextual data relating to the situation being simulated. This will include geographical data, travel times for combinations of origins/destinations, incident categories, characteristics of delays etc. Even quite sophisticated criteria could be added, for example ‘mutual aid responses’, where fire crews are moved to incidents outside the region being studied. Static data are not the only kind that could be entered. In the case of fire severity it was also possible to include a relationship between speed of arrival by a fire crew and severity of the fire.

The input of the deployment policy is perhaps the most crucial element of the simulation. This was effectively the independent variable for which the effects of changes were being studied.

The final category of setup data was the pattern of events to be included in the simulation. The simulation could be run in two ways. If a historical pattern of events needed to be entered then a computer tape of ‘exogenous event’ was prepared and then used for controlling the simulation. In other circumstances it was more appropriate to enter probability distribution data (e.g. Poisson distribution) for various categories of event, and then produce a computer-generated set of events.

The simulation program itself consisted of a number of sub-routines that operated in a specific order and resulted in an imitation of the response of a fire department to a sequence of incidents for a particular deployment policy. Figure 16 shows the logical flow of these sub-routines and Table 3 explains the purpose of each sub-routine. Tables 4–6 show the sub-routines that controlled the types of event generated; they give a flavour of how rich the simulation was in producing a wide range of lifelike events.

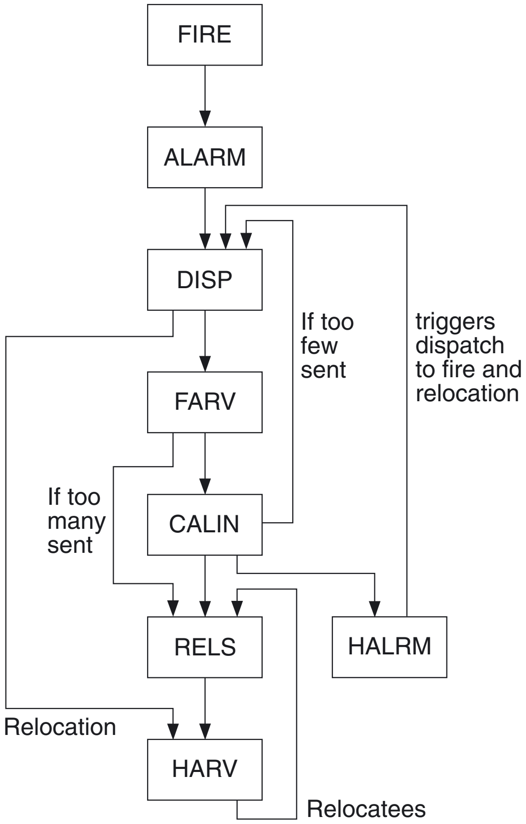

Figure 16 Flow of events in the fire operations simulation model (Source: p. 503)

Table 3 Major events in the fire operations simulation model

Event	Definition and discussion
FIRE	The start of the incident. The delay until the incident is reported appears on the input tape. The tape can be modified if, for example, the introduction of detectors reduces the delay.
ALARM	The receipt of the alarm at the dispatching office. Schedules the dispatch of fire companies. Therefore, it can reflect the possible dependence of the delay, from alarm receipt until dispatch on the number of incidents then active (Section 14.5). Or, it can reflect the possible changes in this delay that would occur if communications office operations were changed (for example, by the installation of a computer-aided dispatch system).
DISP	The notification of companies to respond to an alarm or to relocate.
FARV	The arrival of fire companies at the scene. A fire escalation model or a dollar-damage calculation would use this company arrival information together with the FIRE event time.
CALIN	The first report to the communications office from a company at the scene of the incident. If too many units are sent, those not needed will be turned back. If too few are sent, others will be sent (see HALRM).

Table 3 (continued)

Event	Definition and discussion
HALRM	The request from the scene for a second-alarm, third-alarm, and so forth. Will induce dispatch of companies to the scene and relocations of companies into empty houses.
RELS	The release of companies from the scene. They are usually available during the trips back to their houses. (If it is necessary to recognize standby status – a company on the scene but available for dispatch to another incident – a new type of event would be needed.) The release of companies is dependent on the work times specified. A unit's work time can begin when it arrives or when the fastest of the needed companies arrives, as specified by the user.
HARV	The arrival of a company at a fire house. It can be a company returning from an incident or a relocated company returning home. Or, it can be a relocated company getting to the house it will cover. The company may be available for dispatch at this time, or it may be given some time to recover.

Source: p. 504

Table 4 Other important events in the fire operation simulation model

Event	Definition and discussion
AVAIL	The return to service of a company that is granted a recovery period after return from an incident.
DOWN	The dispatch of companies to an incident outside the area of primary concern. The units to be sent are specified by the exogenous events tape (rather than by the ALARM event).
ACC	The occurrence of an accident during a company's trip.
INSER	The return to service of a company, following DOWN or ACC.
SMPCV	Used to determine the coverage being provided. Called periodically for output purposes.

Source: p. 505

Table 5 Major sub-routines in the fire operations simulation model

Sub-routine	What it does, and why
DATAT	Used for aggregating alarms for output purposes. Determines the output class in which the response times and response adequacy for this alarm belong.
DCDE	Chooses the companies to dispatch to an incident in response to a call from the ALARM, CALIN, HALRM, or ACC.
SEND	Determines the length of every trip by a company.

Table 5 (continued)

Sub-routine	What it does, and why
LCTCO	Calculates the location of a company that is currently travelling. Can be used to consider dispatching companies that are returning from incidents, or relocating.
STOPCO	Changes the destination of a company that is currently travelling. Used to turn back excess companies if the first-arriving unit finds no need for them, to redirect companies (that are relocating or returning) to a new incident, and in case of accident.
TESTA	Determines whether an accident will occur on this trip.
DECDR	Makes the relocation decisions.
DORLT	Does the relocations (if any) that DECDR picks.
UNRLCT	Returns relocated companies to home stations. (If any are working at alarms, assures that they will return home rather than to the houses they responded from.)

Source: p. 505

Table 6 Major function sub-programs in the fire operations simulation model

Function	Purpose
DIST	Determines a travel distance given the (x, y) coordinates of the starting point and the destination.
TRVLT	Determines the travel time of a trip on the basis of the distance to be covered. If a matrix of travel times between points is available, it can replace this function and DIST.
DSPT	Determines the time required to make the dispatch decision and notify the companies.
AVLT	Determines the length of time a unit will be out of service following an accident.

Source: p. 506

The final part of the model was the post-simulation analysis. This involved producing statistics on almost any variable that the RAND team wished to study. The most common variables considered were:

- average response time;
- standard deviation of response time;
- response adequacy;
- coverage (the potential of the service to respond to incidents in a particular region at any point in time);
- workload of crews;
- number and frequency of fire crew re-locations.

Issues addressed

Using the techniques described above, a number of application areas emerged from the study. The project demonstrated that systems analysis techniques could be applied to the fire services and that quantitative models could be developed and used to support decision making regarding:

- the allocation and location of fire services;
- the dispatch of services to a fire event;
- relocating units to cover for busy units.

Allocation and location

Perhaps the fundamental question for a fire service is 'how many fire crews are needed?' In a perfect world this may be calculated based on the expectation of the number and types of calls. It is perhaps more likely that the economic imperative will lead to the number of fire crews being based on the available budget. Given such constraints, the important question is 'how should these crews be allocated across the region?', given that regions have different needs for fire protection and that allocation is a compromise between potential and realized demand. Once the allocation issue has been resolved, the next step is to decide where to locate the crews within an area.

To address the problem of crew allocation, the RAND project rejected the use of a crude cost–benefit analysis of the problem. Balancing the cost of deployment of crews against the cost reduction due to improved fire protection was seen as impossible.

It was recognized that a best, or most equitable, solution could not be defined mathematically owing to the complex nature of the decision-making process. The team considered that the issue was concerned with the compromise between minimizing a loss function and equalizing that loss function across a region. The parametric allocation model was developed, whereby a 'trade-off parameter' was identified, such as average travel time to an incident, indicating the emphasis to be placed on the satisfaction of objectives. By varying the value of the parameter it was possible to change the emphasis between satisfying potential or realized demand. A parameter value of 0 emphasized the potential demand whilst a value of 1 emphasized realized demand. Taking travel time as the key indicator and objectives for allocating crews across a region stated in terms of minimizing travel time, two extreme strategies are possible. First, crews could be allocated to the high-risk areas, implying that travel time for most fires would be minimized. Unfortunately, for a fire in a low-risk area, travel time would be compromised. The second strategy was to provide an equitable distribution of crews across the region, so maintaining a consistent average travel time. The problem with this second approach is that it would lead to higher workloads and travel times for crews in high-risk areas. Consequently the RAND team chose to develop a model that initially provided a base level of crew allocation across a region that would meet average demand, often based on the 4 pm to midnight data, and would then allocate the remaining crews proportionately based on the area's typical demand and an assessment of hazard. The model is then run several times to produce a number of alternatives for consideration at the decision-making point.

Figure 17 shows the steps involved. Notice that the decision point is made by a user not the model.

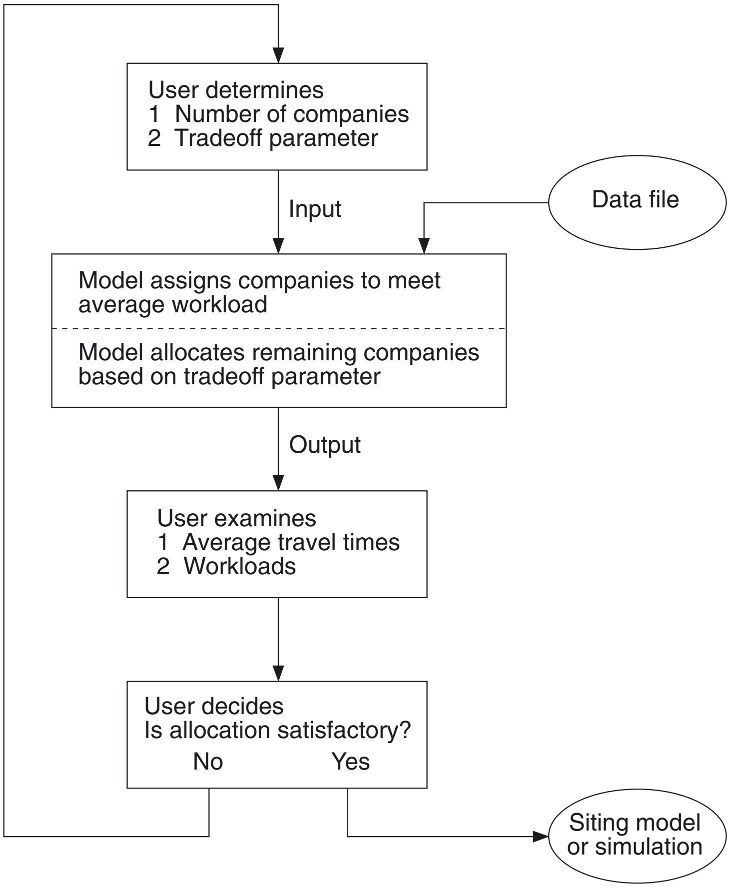

Figure 17 Steps in the use of the parametric allocation model (Source: p. 335)

To provide data, the area was first split into demand regions and then the following information was input for each region:

- average alarm rate for different types of alarm, such as structural fires and false alarms;
- the existing number of fire engines and ladders;
- estimates of potential demand based on levels of risk in particular areas.

The next step was to consider the specific location of fire stations. To achieve this, the RAND Institute developed a descriptive model, the firehouse site evaluation model, whose output could not show an optimum location but could make it easy to eliminate poor choices. A later, more sophisticated model, the fire station location package (FSLP), was able to produce an optimum solution.

By 1976 the two models had been adopted by several jurisdictions. Interestingly, 52 of the more complex FSLP models were implemented compared with 10 of the siting model. The advantages of the FSLP model were that it had potential use for many public service applications, perhaps making the overhead in building the street network more cost effective.

The result of this application of modelling to the problem of fire station allocation was not to provide a mechanistic approach to siting but to provide additional information for the decision-making process, again showing the value of models that provided a simplified rather than an exact representation of a situation.

Initial dispatch

The previous two application areas of systems analysis were concerned with the physical design of the fire service, namely density of resources and specific sites for fire stations in a region. The next two applications were concerned with the design of system processes. Prior to the 1960s the rules governing the dispatch of fire services to calls was relatively simple. The low level of calls meant that a standard response to all calls was acceptable and practical. The dispatching policy was therefore simple. The dispatcher would either use information provided in the alerting phone call to provide an adequate response or, in the case of calls alerted via alarm boxes, a standard response was dispatched. Low levels of false alarms meant that there was not excessive wastage using this policy. The relatively low frequency of fires meant that the chances of several calls occurring simultaneously were minimal and so losses to life and property due to non-availability of resources were correspondingly rare and not seen as an important issue.

At the time of the RAND project, the conditions under which the fire service needed to respond were changing, as detailed earlier. The pressure on local authorities to monitor and control public spending had the effect of cutting available budgets. The problem of dispatching was becoming a central issue, especially in terms of:

- ensuring that services would always be available, even during peak periods of demand;
- ensuring that services were not wasted on false alarms;
- maintaining overall services within shrinking budgets.

A balance was needed between the losses of life and property due to lack of availability of fire services and the direct/indirect cost of wasted resources. This led to the need to categorize alarms in some way that would help guide the dispatcher’s decision.

A simple categorization of fire severity was needed to distinguish between serious and non-serious fires. This would provide a simple basis for dispatch.

Unfortunately, the decision on whether an alarm signals a serious fire is dependent on the amount of information available. In reality the dispatcher depended on a small range of information:

- the seriousness of the fire;
- the time of day;
- the weather conditions.
- the location.

In addition, other factors such as the current deployment at the time of the call may affect the decision on how to deploy services to an incident.

The result of RAND's study was the identification of the role that different dispatching policies may have. Three generic policies were identified.

- *Fixed* This is the simplest type of policy and involves a simple algorithm that on the assessment of the fire as either serious or non-serious will define the number of resources and from which station(s). Other factors may also be taken into account, such as time of day or whether the location of the alarm indicates high risk, e.g. a built-up industrial area. For a region with a low alarm rate this is probably the simplest and most effective policy.
- *Variable* This type of policy is sensitive to the availability of resources when the alarm is raised. This type of policy is more appropriate for a busy city; it reduces workload and increases coverage of a region.
- *Adaptive* This policy uses characteristics from the two other policies. By mixing predetermined responses with adjustment on the basis of resource availability, a hybrid policy may be created.

In developing an adaptive policy built on the use of queuing theory, the RAND team created a simulation varying the number of companies on duty, number dispatched and the alarm rate. The simulation exercise tested three basic strategies:

1. add companies;
2. reduce response time;
3. a combination of 1 and 2.

The effect of traditional dispatch policies was to increase workloads under strategy 1, increase response times of second- and third-arriving engines for strategy 2, while the third strategy had positive effects on both workload and response times.

In November 1972 FDNY implemented the adaptive response policy, after piloting it in South Bronx for three years, to help it cope during busy periods. Prior to establishing the experiment the RAND team simulated the effect using a more sophisticated program which included a representation of the geography of the Bronx area. The policy eventually decided upon meant that generally fewer units were sent than traditionally, and that response times to serious fires were reduced (Ignall *et al.*, 1975)

Relocation

A problem related to dispatching fire crews is the decision on how to relocate fire crews. Relocation is required so that areas that are normally covered maintain cover when their 'home' crews are busy.

Relocation is normally straightforward in times of low alarm rates and, prior to the RAND project, was handled through the use of predetermined policies held on a card system. During the process of dispatching crews these relocation cards would then be used for directing crew movements so that all areas would maintain cover. Unfortunately, as the rate of fires increased, especially when several fires occurred simultaneously, predetermined relocation systems tended to break down as many of the units to be relocated might themselves be fighting fires.

The RAND group saw that the application of analysis might yield a system that could produce 'good' relocation decisions and might even allow the relocation process to be done by a computer in real time. This solution was particularly attractive, as the problem of relocation requires a complete view of the fire service deployment at any moment in time. However, in common with the allocation problem, what constitutes a 'good' relocation decision was hard to define.

The team decided that the problem of relocation was in fact a series of sequential decisions that needed to be made (p. 464):

1 When should a relocation be made?
2 How many companies should be relocated?
3 Which available companies should be relocated?
4 To which house should each relocating company go?

Two methods of solving the problem were presented: one was a mathematical solution and the other a heuristic algorithm. The algorithm was tested against contrived situations of interest, a simulation model of over 3600 alarms, and historical data from one evening in the Bronx. In tests, the heuristic algorithm produced the optimal solution.

The relocation algorithm was implemented in the Brooklyn dispatch office in mid-1977 as part of a real-time system solution. The role of the algorithm was to recommend relocations in response to changing information regarding alarm and company status. Should a response neighbourhood not have cover, the system will suggest a solution. The dispatcher then decides whether to implement the suggestion in full, part, or not at all, based on any relevant information the program is unaware of.

Where the algorithm is not used as part of a real-time system, it can be used to generate a number of pre-planned relocation responses. Unlike those produced previously, the use of a computer can generate a greater number of more sophisticated potential responses, thereby giving the dispatcher more flexibility.

Conclusion

Systems analysis techniques seem to have had an important role in the support of decision makers. The use of quantitative techniques, sometimes for the first time, provided very clear guidance for some key decisions. This guidance, though, is perhaps very narrow, as in building the models the real-world situation had to be represented in much simpler terms.

It is important that we recognize that the models created were only models; they remain separate from the reality of fire protection in New York – it was never intended that events on a specific day would follow a predicted pattern. Their relevance, though, is in identifying some of the key variables and making decision makers sensitive to their effects on a situation. They were intended to guide decision making, not to replace it, for example regarding manning levels in stations.

Within the context of this course, was the RAND–FDNY project an example of a systems engineering project? The answer to this is probably not! For many reasons the RAND team looked at the problem situation of the New York Fire Service and were very selective in the areas that they chose to study. For political reasons it was necessary to have short-term gains realized as well as longer-term gains. The cost of the study also had to be contained within limits, and this placed constraints on the amount and range of data available. Finally, the range of systems analysis techniques available was suitable for relatively well-defined problem situations. For example, the use of queuing theory requires the consideration of a small set of quantifiable variables. This meant that considering the wider impact of changes to the system, e.g. relationships between dispatching policy and morale, was not possible. The principle of systems engineering being holistic is not fulfilled in this case. Many of the wider organizational and social aspects of the fire protection service had to be ignored, so the study could be accused of focusing only on the aspects of the situation for which possible solutions could easily be defined.

References

Blum, E.H. (1971) ‘Studies of the operations of the New York City Fire Department’, *Fire Research Abstracts and Reviews*, vol. 13, no. 1, pp. 23–35.

Ignall, E.J., Kolesar, P., Swersey, A.J., Walker, W.E., Blum, E.H. and Carter, G. (1975) ‘Improving the deployment of New York City fire companies’, *Interfaces*, vol. 5, no. 2 pt. 2, February 1975.

The National Commission on Fire Prevention and Control (1973) *America Burning*, Washington D.C., Government Printing Office.

Walker, W.E., Chaiken, J.M. and Ignall, E.J. (eds) (1979) *Fire Deployment Analysis: A Public Policy Analysis Case Study*, New York, Elsevier North Holland, Inc.