S^2F^2M - Statistical System for Forest Fire Management*

Germán Bianchini, Ana Cortés, Tomàs Margalef, and Emilio Luque

Departament d'Informàtica, E.T.S.E,
Universitat Autònoma de Barcelona,
08193-Bellaterra (Barcelona), Spain

Abstract. One of the most serious problems in wildland fire simulators is the lack of precision for input parameters (moisture content, wind speed, wind direction, etc.). In this paper, a statistical method based on a factorial experiment is presented. This method evaluates a high number of parameter combinations instead of considering a single value for each parameter, in order to obtain a prediction which is closer to reality. The proposed methodology has been implemented in a parallel scheme and tested in a Linux cluster using MPI.

1 Introduction

The main goal of forest fire model developers is to provide models that explain and predict fire behavior. These models can be used to develop simulators and tools for preventing and fighting forest fires [1, 2, 7, 8]. These simulators and tools are integrated into a Decision Support System (DSS). It is possible to define a DSS as "a computer system that helps in the process of making a decision, helping users to form and explore the implications of their judgments, and, therefore to make decisions based on understanding" [14]. Therefore, this type of system should help to form judgments instead of giving general advise as, for example, an information digest does in a database. Nowadays, a DSS has the more ambitious objective of trying to supply accurate information (sometimes in real time) to achieve terrain planning, implementation of preventive rules, efficient monitoring and giving online help while the forest fire is happening.

However, most models are unable to accurately predict the forest fire's behavior. This is due to several reasons but one of the most significant ones is that there are several parameters (i.e. moisture content, wind conditions, etc.) that are difficult to estimate precisely.

It is possible to minimize this input parameter problem by using techniques such as parameter optimization [3], with the aim of determining as precisely

* This work has been supported by the Comisión Interministerial de Ciencia y Tecnología (CICYT) under contract TIC2001-2592 and by the European Commission under contract EVG1-CT-2001-00043 SPREAD.

V.S. Sunderam et al. (Eds.): ICCS 2005, LNCS 3514, pp. 427–434, 2005.

as possible the parameter values that provide the closest prediction of real behavior.

In this paper, although we also focus on processing the parameters, our goal is to develop a methodology based on statistical analysis to determine the most probable behavior of a forest fire and apply this methodology to implement a DSS.

S^2F^2M (Statistical System for Forest Fire Management) does not feed the simulation core with "known" single values, but rather carries out a set of simulations considering a range of possible values for the input parameters that are more uncertain.

This method requires a lot of computations to reach a conclusion because it is necessary to run a large number of simulations. To tackle this problem we have used a parallel scheme (master-worker), applied in a PC cluster. The method has been implemented using MPI as a message pass library and is executed in a Linux cluster. In this paper we analyze the improvements obtained by using the proposed scheme in terms of quality of the prediction and simulation speed-up for burns on experimental fields.

This paper is organized as follows: The factorial experimentation and basic concepts of the system are explained in section 2. The system's implementation is described in section 3. Section 4 includes the results obtained when the method was applied to two forest fires. Finally, the main conclusions are reported in section 5.

2 Factorial Experimentation

The methodology of this work is based on statistics. Statistics deal with collection, presentation, analysis and use of data to make, for example, decisions. There are two possible ways of collecting data about an event. In an **observational study** the researcher only takes notes without interacting in the situation. Data are obtained as they appear.

Another way is through **designed experiments**. In these kinds of experiments it is possible to make deliberate changes in the controlled variables of a system or process. Results are observed and then it is possible to either make an inference or make a decision about variables that are responsible for changes. When there are a lot of significant factors involved (i.e. weather, wind speed, slope, etc.), the best strategy is to use some kind of **factorial experiment**. A factorial experiment is one in which the factors vary at the same time [16](for example, wind conditions, moisture content and vegetation parameters). A **scenario** represents each particular situation that results from a set of values.

For a given time interval, we want to know whether a portion of the terrain (called a cell) will be burnt or not. If n is the total number of scenarios and n_A is the number of scenarios in which the cell was burned, we can calculate the **ignition probability** as:

$$P_{ign}(A) = n_A/n$$

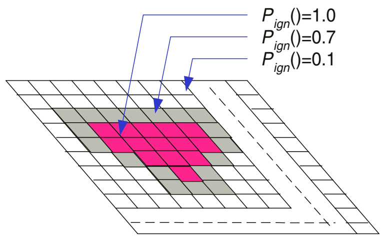

Fig. 1. Generalizing the cells analysis

The next step is to generalize this reasoning and apply it to some cell sets. In this manner we obtain a matrix with values representing the probability of each cell catching fire (Fig. 1).

Hence, we can focus our analysis on the procedure of generating possible scenarios.

2.1 Scenarios Generation

Our system uses a forest fire simulator as a black box which needs to be fed with different parameters in order to work. A particular setting of the set of parameters defines an individual scenario. These parameters correspond to the parameters proposed in the Rothermel [9] model.

For each parameter we define a rank and an increment value, which are used to move throughout the interval. For a given parameter i (which we will refer to as *Parameter_i*) the associated interval and increment is expressed as:

[Inferior_threshold_i, Superior_threshold_i], Increment_i

Then, for each parameter i, it is possible to obtain a number C_i (parameter domain cardinality), which is calculated as follows:

$$C_i = ((Superior\_threshold\_i - Inferior\_threshold\_i) + Increment\_i)/Increment\_i$$

Finally, from each parameter's cardinality it is possible to calculate the total number of scenarios obtained from variations of all possible combinations.

$$\#Scenarios = \prod_{i=1}^{n} C_i$$

where n is the number of parameters.

3 S^2F^2M Implementation

The concepts described above has been implemented in an operational system that incorporates a simulation kernel and applies the methodology to evaluate the fitness function. This system has been developed on a PC LINUX cluster using MPI as message passing library.

3.1 The Simulator

S^2F^2M uses as a simulation core the wildland simulator proposed by Collin D. Bevins, which is based on the fireLib library [4]. **fireLib** is a library that encapsulates the BEHAVE fire behavior algorithm [1]. In particular, this simulator uses a cell automata approach to evaluate fire spread. The terrain is divided into square cells and a neighborhood relationship is used to evaluate whether a cell will be burnt and at what time the fire will reach the burnt cells.

As inputs, this simulator accepts maps of the terrain, vegetation characteristics, wind and the initial ignition map.

The output generated by the simulator consists of a map of the terrain in which each cell is labeled with its ignition time.

3.2 The Fitness Function

To evaluate the system's response we defined a fitness function. Since S^2F^2M uses an approximation based on cells, the fitness function is defined as the quotient between the number of cells in the intersection between the simulation results and the real map, and the union of the simulation results and the real situation (Fitness = (cells in the intersection) / (cells in the union)).

Figure 2 shows an example of how to calculate this function for a terrain made up of 5x5 cells. In this case, the fitness function is 7/10 = 0.7.

A fitness value equal to one corresponds to the perfect prediction because it means that the predicted area is equal to the real burned area. On the other hand, a fitness equal to zero indicates the maximum error, because in this case our experiment did not coincide with reality at all.

3.3 Parallelisation

S^2F^2M has to make a large quantity of calculations because it uses a sequential simulator as a kernel [4], and for this reason it needs to make a simulation for each resulting combination of parameters (*#Scenarios*). This high number of simulations requires a lot of time.

To reduce the execution time we used multiple computational resources working in parallel to obtain the desired efficiency. Keeping in mind the nature of the problem that S^2F^2M tries to solve, we believe a master-worker architecture is

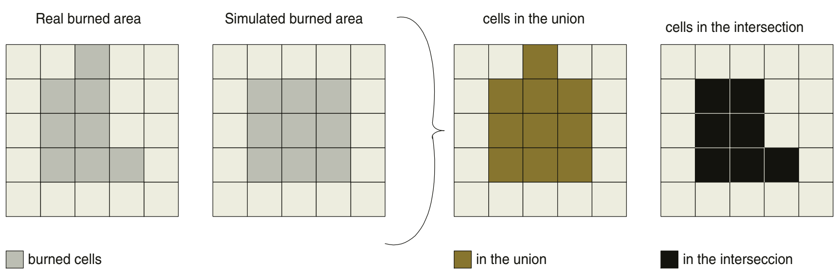

Fig. 2. Calculating the fitness for a 5 x 5 cell terrain

suitable to achieve this aim, because a main processor can calculate each combination of parameters and send them to a set of workers. These workers carry out the simulation and return the map to the master. This resulting map indicates which cells are burned and which are not.

Our system has a well defined structure. The Master process has a data reception stage (parameter files, terrain files, simulation time, etc.). After this there is an initialization stage for data structures. In the main loop, the Master process distributes scenarios to the workers, waits for results, receives results and distributes more data to idle workers (if there are more scenarios to simulate). Finally, it gives a graphical output.

The Worker structure is complementary. Each one has a data reception stage (to initialize terrain size, slope). Following this, it enters a loop to receive scenarios from the Master process to activate the simulation function for calculating fire spread.

4 Experimental Results

To test the system we used two experiments in the field. Both burns took place in Serra da Lousã (Gestosa, Portugal (40°15'N, 8°10'O)) , at an altitude of between 800 and 950 m above sea level. The burns were part of the SPREAD project [15]. In the Gestosa field experiments [10], terrain was divided into dedicated plots in order to carry out different sorts of tests and measurements. In particular, we worked with plots 513 and 519, which had the following characteristics:

Experiment 1 (Plot 513): the plot was represented by means of a grid of 58 columns x 50 rows (each cell was 2.989 x 2.989 feet).

Experiment 2 (Plot 519): the plot was represented by means of a grid of 89 columns x 91 rows (each cell was 2.989 x 2.989 feet).

In order to gather as much information as possible about the fire-spread behavior, a camera recorded the complete evolution of the fire. The video obtained was analyzed and several images were extracted every 2 minutes in the first experiment and every 2.5 minutes in the second. From the images the corresponding fire contours were obtained and converted to cell format in order for S^2F^2M to interpret them.

4.1 Experiment 1

The first case is very complicated, because in a field experiment it is not possible to control environmental conditions. Nevertheless, we fixed certain known values (*slope* and *moisture* in 1, 10 and 100 hours) and let the others vary.

To make comparisons we fixed the initial time to 0 and a limit value of 12 minutes.

In table 1, we can see that the fitness for this experiment has values between 0.7 and 0.91. This indicates that our statistic output is very close to reality. It is important to note that in the Fitness table only those cells with 100% ignition probability (i.e. cells burned in 100% of the scenarios) are considered.

Table 1. Fitness of experiment 1 in each interval

Initial time	Final time	Fitness
0:00	2:00	0,749420
2:00	4:00	0,690152
4:00	6:00	0,864360
6:00	8:00	0,953166
8:00	10:00	0,826158
10:00	12:00	0,915669

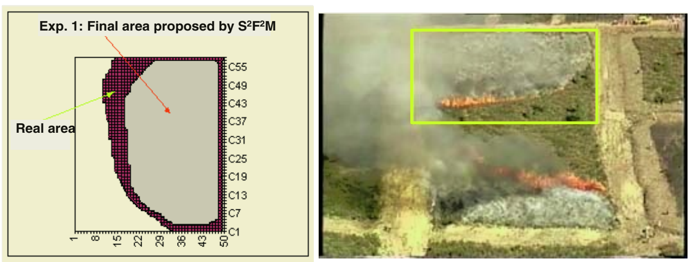

Fig. 3. S^2F^2M output for minute 12 of Experiment 1

In the figure 3 we can see that the S^2F^2M result is always inside reality, that is, the result does not exceed the real propagation. We only include the last step of the simulation (in this case at minute 12), as the previous steps are included in the real perimeter.

4.2 Experiment 2

The second experiment has a rank file equal to plot 513. This is because the plots are located very near each other, and therefore the terrain features can be taken to be equivalent. Table 2 shows the resulting fitness.

Finally, using the same criterion, we show the state proposed by S^2F^2M at minute 12.5 in figure 4. It is possible to identify clearly that the S^2F^2M area is inside the real perimeter.

4.3 Speed-Up Improvement

In a real case the system works under real time constraints and, therefore, it is necessary to analyze the speed-up obtained by using different numbers of

Table 2. Fitness of experiment 2 in each interval

Initial time	Final time	Fitness
2:30	5:00	0,451988
5:00	7:30	0,486521
7:30	10:00	0,425703
10:00	12:30	0,774615

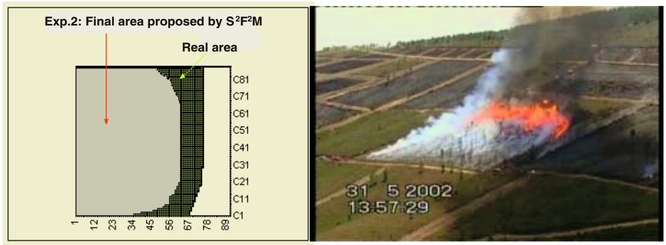

Fig. 4. S^2F^2M output for minute 12.5 of Experiment 2

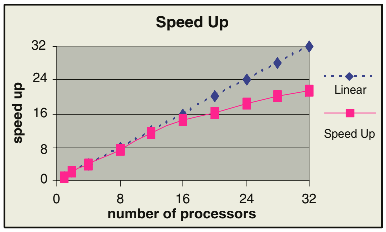

Fig. 5. Speed-up curve for experiment 1

processors. The number of processors used in the successive experiments were 1, 2, 4, 8, 12, 16, 20, 24, 28 and 32. Figure 5 shows the speed-up for a particular example compared with a linear speed-up (the ideal case).

It can be observed that the speed-up is close to being linear until 16 processors are used. From this point, an increase in the number of processors continues being profitable, but it can be observed that the speed-up is not so linear.

5 Conclusions

In this paper we have described a tool with the objective of offering an alternative to the normal use of a forest fire simulator. With this methodology we can obtain a prediction of the ignition probability of a terrain without knowing exact data about climatic factors, and without waiting for the fire to start.

From the experimental studies we can conclude that the area that S^2F^2M indicates with 100% probability of being reached by fire in a time interval is always included in the real burned area. Furthermore, since each output proposed by S^2F^2M needs a lot of calculations, we have used the parallel scheme of a master-worker programming paradigm in order to speed-up the whole process.

References

1. Andrews P. L. “BEHAVE: Fire Behavior prediction and modeling systems - Burn subsystem, part 1”. General Technical Report INT-194. Odgen, UT, US Department of Agriculture, Forest Service, Intermountain Research Station; 1986.

2. Andrews, Patricia L.; Bevins, Collin D.; Seli, Robert C. “BehavePlus fire modeling system, version 2.0: User’s Guide”. Gen. Tech. Rep. RMRS-GTR-106WWW. Ogden, UT: Department of Agriculture, Forest Service, Rocky Mountain Research Station. 2003.
3. Baker Abdalhaq, G. Bianchini, Ana Cortés, Tomàs Margalef, Emilio Luque: “Improving Wildland Fire Prediction on MPI Clusters”. LNCS 2840, pp. 520-528, 2003.
4. Collins D. Bevins, “FireLib User Manual & Technical Reference”, 1996. http://www.fire.org.
5. E-FIS A ten telecom project. http://www.e-fis.org/
6. Eftichidi G., Varela V. 1999. SAFES: Safe Fire Expert System. Presentation in the International Scientific Conference “Fires in the Mediterranean forests: Prevention -Suppression - Soil Erosion - Reforestation” organised by UNESCO in Athens, 3-6 February 1999.
7. Finney, Mark A.. “FARSITE: Fire Area Simulator-model development and evaluation”. Res. Pap. RMRS-RP-4, Ogden, UT: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 47 p. 1998
8. ADAI Products: FIRESTATION http://www.adai.pt/products/firestation/
9. Rothermel, R. C., “A mathematical model for predecting fire spread in wildland fuels”, USDA FS, Ogden TU, Res. Pap. INT-115, 1972.
10. ADAI - CEIF (Center of Forest Fire Studies) http://www.adai.pt/ceif/Gestosa/
11. MPI: The Message Passing Interface Standard http://www-unix.mcs.anl.gov/mpi/
12. Prometheus. http://kentauros.rtd.algo.com.gr/promet/schedule.htm
13. Reinhardt, E.D.; Keane, R.E.; Brown, J.K. “First Order Fire Effects Model: FOFEM 4.0, User’s Guide”. General Technical Report INT- GTR- 344. 1997.
14. Sixto Ríos Insua, Concepción Bielza Lozoya, Alfonso Mateos Caballero, “Fundamento de los Sistemas de Ayuda a la decisión”, RaMa 2002 ISBN 84-7897-494-6
15. Project Spread, Forest Fire Spread Prevention and Mitigation http://www.adai.pt/spread/
16. Douglas C. Montgomery, George C. Runger, “Probabilidad y Estadística aplicada a la Ingeniería”, Limusa Wiley 2002 ISBN: 968-18-5914-6