

# Applying a Dynamic Data Driven Genetic Algorithm to Improve Forest Fire Spread Prediction\*

Mónica Denham, Ana Cortés, Tomàs Margalef, and Emilio Luque

Departament d' Arquitectura de Computadors i Sistemes Operatius, E.T.S.E.,  
Universitat Autònoma de Barcelona, 08193 - Bellaterra (Barcelona) Spain

**Abstract.** This work represents the first step toward a DDDAS for Wildland Fire Prediction where our main efforts are oriented to take advantage of the computing power provided by High Performance Computing systems to, on the one hand, propose computational data driven steering strategies to overcome input data uncertainty and, on the other hand, to reduce the execution time of the whole prediction process in order to be reliable during real-time crisis. In particular, this work is focused on the description of a Dynamic Data Driven Genetic Algorithm used as steering strategy to automatic adjust certain input data values of forest fire simulators taking into account the underlying propagation model and the real fire behavior.

## 1 Introduction

Forest fires are one of the most important threats to forest areas in the whole world. In the last years, thousands of hectares were lost by wildfires action. Forest areas losses damage the nature, attempting on ecological balance. Death of different species of animals and plants, profitable areas loss, air pollution, floods, water contamination, are some of the consequences of forest fires. At the same time these problems cause different diseases, famine, animals and vegetables extinction, etc. These facts attempt our standard of living.

Some forest fires occur by nature itself: long dry seasons, elevated temperatures, electric storms, could generate wildland fires. These type of fires can help in the ecological balance: young plants take place where there were old and perhaps unproductive trees. Nature keeps number of fires limited, but in the last years this number was increased by human factors. More than 90% of forest fires are provoked by human hand (accidents, carelessness and negligence) [6]: 16000 hectares were burned in Gran Canaria last summer (2007) during 6 days of strong fires [11]. In Tenerife 15000 hectares were burned at the same time [11]. More than 60 deaths were occurred during forest fire in Greece last July [12].

---

\* This work has been supported by the MEC-Spain under contracts TIN 2004-03388 and TIN 2007-64974.

Whereas the number of wildfires had increased in the last years, human work in this area had increased too: studies, strategies and tools to prevent forest fires as well as to reduce the fire during a disaster were developed these last years. Nowadays, several forest fires simulators exist for helping and improving this work. Most of these simulators are based on Rothermel mathematical model [6]. This model describes fire behavior through mathematical equations.

Simulators are used to predict fire behavior, improving the accuracy of actions and reducing fire effects. Several of these simulators use a large number of inputs for describing the environment where fire occurs. It is very difficult to dispose of exact real-time values of these parameters. Several of them change with time: air and fuel humidity change along the day (day-night cycle), weather changes due to elevated temperatures provoked by fire, fires generate very strong gust of winds, etc. Furthermore lots of these parameters have their own behavior pattern and it can change very strongly within a wildland fire. This input data uncertainty causes predictions that are far from the real fire propagation.

During a real hazard such as wildland fire both, the accuracy of the prediction propagation and the prediction response time are crucial key points. For this reason, we propose a fire propagation prediction system, which explores multiple fire propagation scenarios obtained from combining different values of the input parameters by dynamically adapting those scenarios according to observed real fire evolution. These characteristics directly match to the Dynamic Data Driven Application Systems (DDDAS) definition: *DDDAS is a paradigm whereby application/simulations and measurements become a symbiotic feedback control system. DDDAS entails the ability to dynamically incorporate additional data into an executing application, and in reverse, the ability of an application to dynamically steer the measurement process* [7].

An optimal framework for a reliable DDDAS for Wildland Fire Prediction must consider, among others, the following issues: the ability to dynamically couple models from different disciplines; real-time data assimilation strategies for being further injected into the running system; steering strategies for automatic adjusting either models or input data parameters and to have access to enough computational resources to be able to obtain the prediction results under strict real-time deadlines. Some current work on this area could be found in [10,2].

Our currently work consists on the first step toward a DDDAS for Wildland Fire Prediction where our main efforts are oriented to take advantage of the computing power provided by High Performance Computing systems, on the one hand, to propose computational data driven steering strategies to overcome input data uncertainty and, on the other hand, to reduce the execution time of the whole prediction process in order to be reliable during real-time crisis.

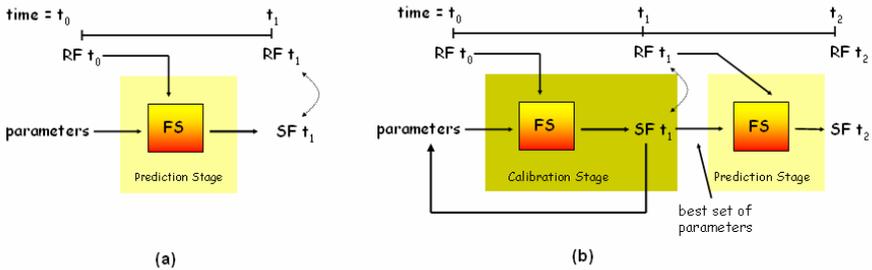
In particular, this work is focused on the first working line. A Dynamic Data Driven Genetic Algorithm (GA) is used as steering strategy to automatic adjust certain input data values taking into account the underlying propagation model and the real fire behavior.

This work is organized as follow. Next section describes the proposed dynamic data driven forest fire prediction methodology compared to the classical

prediction scheme. Section 3 is focused to describe the Calibration stage of the proposed prediction methodology. The experimental results are shown in section 4 and, finally, the main conclusions are reported in section 5.

## 2 Forest Fire Spread Prediction

Traditionally, forest fire spread prediction has been performed using a particular forest fire simulator as a prediction tool. Forest fire simulators need to be fed with several input data such as the initial fire line (what we call  $RFt_0$ : Real Fire at time  $t_0$ ), meteorological data, vegetation data, etc. The simulator uses this input data to obtain the simulated fire line for a subsequent instant  $t_1$  (we will refer to as  $SFt_1$ : Simulated Fire at time  $t_1$ ). This classical forest fire prediction method is showed in Fig. 1 (a).



**Fig. 1.** (a) Classical Forest Fire Propagation Prediction method. (b) Two Stage Propagation Prediction method.

As we can see in this figure, the only information required for the classical prediction method consists of a unique set of values for the input parameters needed by the simulator. Once this input data has been set up, the simulator is executed providing, as part of its output data, the predicted fire line after a certain period of time. Despite of its simplicity, this method has an important drawback: simulation accuracy depends on the unique set of input values, and, as we had mentioned, to dispose their real values is very difficult.

In order to overcome this drawback, we propose a wildfire propagation prediction methodology based on a two stages prediction scheme: Calibration stage and Prediction stage. The DDDAS bases are included in the Calibration stage where taking advantage of the computer power provided by High Performance Computing systems, a GA is used to explore a huge number of input parameters combinations (called scenarios). For this stage being operative, it is necessary to be able to assimilate data about the forest fire front at two consecutive time instants ( $t_0$  and  $t_1$  in Fig. 1(b)).

Each one of the generated scenarios will be fed into the underlying wildland fire simulator, as well as the forest fire front at time  $t_0$ . The obtained propagation results will be compared to the real fire propagation at time  $t_1$  and the results

obtained from this comparison plus the observed real fire propagation at time  $t_1$  will be used as feedback information at the Calibration stage.

By summarizing, we propose to use a Dynamic Data Driven GA as steering strategy within the calibration stage whose internal operations will be dynamically changed according to the data periodically assimilated from the real fire behavior. As a result of this Dynamic Data Driven Calibration stage, we will obtain a set of input data that minimizes the propagation prediction error from  $t_0$  to  $t_1$ . Assuming that the environmental characteristics will keep nearly constant during a certain time (from  $t_1$  to  $t_2$ ), this set of input data will be used in a classical prediction scheme to predict fire spread at  $t_2$  (see Fig. 1(b)). In the following section Calibration stage will be described.

### 3 Calibration Stage

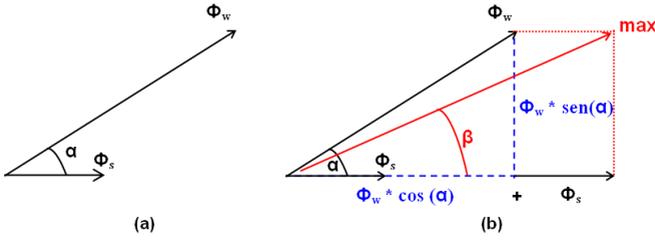
GAs are inspired in the evolutionary theory: a population of individuals (scenarios) is evolved by allowing good individual features to survive through out generations. Therefore, the Calibration stage consists of an iterative process where, at each iteration a new generation is obtained taking into account the goodness of the involved scenarios (similarity degree between simulation and real maps).

Consequently, to be able to properly set up the whole Calibration stage we need, on the one hand, to determine the data that must be assimilated from the observed fire propagation for being posteriorly injected in the calibration process. And, on the other hand, the particular implementation of the dynamic data driven GA. The following sections will focus on these topics.

#### 3.1 Fire Spread Model

Several studies have demonstrated that the parameters that most affect forest fire propagation are wind speed, wind direction and slope characteristics. There exist other factors that also influence forest fire spread such as fuel moisture, fuel type and so on, however, it has been shown that the fire rate of spread is less sensitive to these parameters [1]. Therefore, the proposed fire propagation prediction methodology shown in Fig. 1(b) uses this principle to determine what data to assimilate from the observed real fire behavior. According to this fact, we applied reverse engineering to the underlying wildland fire simulator to analytically determine how wind speed, wind direction and slope interact to generate fire propagation. In this work, we use the fire simulator *fireLib* [5].

Fig. 2 shows in a simplified way, how wind and slope are composed to obtain the direction of the maximum propagation spread, being  $\Phi_w$  the wind effect,  $\Phi_s$  the slope effect and  $\alpha$  the difference between wind and slope directions (Fig. 2 (a)). Both vectors are composed to obtain a third vector, which represents the maximum propagation ratio and its direction (max and  $\beta$  in Fig. 2 (b)). This new vector has a component perpendicular to the slope equal to  $\sin(\alpha) * \phi_w$ , and a component parallel to the slope equal to  $\cos(\alpha) * \phi_w + \phi_s$ , and the angle  $\beta$  is the angle between these two vectors.



**Fig. 2.** (a) Wind and slope effect vectors. (b) Composition of spread vector.

Although the scheme shown in Fig. 2 models in a simplify way how *firelib* proceeds to determine fire spread rate, we can conclude that identifying two of the three vectors involved in the process, we have the key to obtain the third one. Therefore, in the next section, we will discuss how from the analysis of the observed fire behavior we assimilate the needed data to identify each one of these three vectors. Finally, in a later section, we will discuss how this information is injected in the steering strategy.

### 3.2 Data Assimilation

Taking into account slope and wind interactions and their role in fire behavior, we need slope features and real fire progress to obtain the main wind characteristics for performing steering method through Calibration stage. The information related to the slope could easily being obtained from the topography terrain description. Real fire progress is obtained by the analysis of real fire front progress ( $RFt_1$  in Fig. 1 (b)), where point  $(x, y)$  in Fig. 3 is calculated and it represents the maximum rate of spread point in the map.

Applying "reverse engineering" to the fire spread model commented in the previous section, and using trigonometry, we can obtain the following equations:

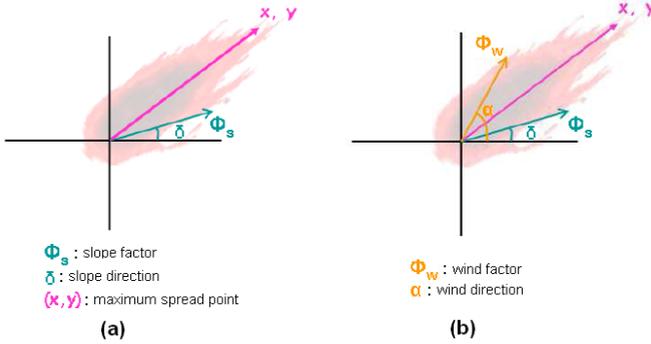
$$x = \Phi_s * \cos(\delta) + \Phi_w \cos(\alpha) \quad y = \Phi_s * \sin(\delta) + \Phi_w \sin(\alpha) \quad (1)$$

Then, as  $x$  and  $y$  values are observed values, it is possible to isolate the wind values from equation 1 as follows:

$$\Phi_w = \frac{x - \Phi_w \cos(\delta)}{\cos(\alpha)} \quad \alpha = \arctan\left(\frac{y - \Phi_s \sin(\delta)}{x - \Phi_s \cos(\delta)}\right) \quad (2)$$

Once wind factor ( $\Phi_w$ ) is obtained, the "ideal" values for the wind speed and wind direction for the observed  $x, y$  point could be deduced. This data will be injected as feedback information inside the dynamic data driven GA to guide the operations applied to a given generation.

It should be recalled that all figures and diagrams are simplified. We are dealing with physics phenomena (fire, wind, slope), where each of them has its own behavior and, in addition, there are physics interactions between them.



**Fig. 3.** Available data: slope and point of maximum rate of spread (a) and, the obtained wind vector (b)

### 3.3 Dynamic Data Driven Genetic Algorithm

The *fireLib* simulator has basically the following input parameters: terrain slope (direction and inclination), wind speed and wind direction, moisture content of the live fuel, moisture content of the dead fuel (at three different times), kind of vegetation (one of the 13 defined through [3]).

Slope and vegetation type change very few in space and time so we assume them as static parameters. Moreover, their values are easy to determine, so we assume them as known and available for our use. Therefore, GA is going to work with the remainder parameters: wind direction and wind speed, moisture content of live fuel and the three moisture content of dead fuel.

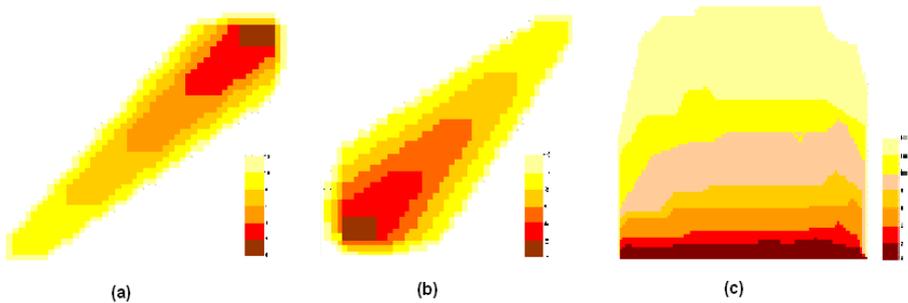
Basically, a GA uses three operations to obtain the consecutive generations: *selection*, *crossover* and *mutation*. *Selection* operation selects good quality parents (fitness function) to create children that will inherit parents good characteristics (by *crossover* operation). In order to guarantee nature diversity of individual characteristics, *mutation* phenomenon can occur for each children characteristic (under a very slight probability). *Selection* can include *elitism* where the best  $j$  individuals ( $j > 0$ ) are included in the new generation directly. Last two operations intent to guarantee nature diversity of individual characteristics.

In the previous section we had exposed how wind values are obtained in order to steer GA: we modify elitism and mutation operations forcing the assignment of these wind values instead of random values. The main goal of this data changes is to give a correct direction to the GA process, achieving the exploration of promising zones of the whole search space. We try to minimize simulation errors, avoiding the use of individuals with inaccuracy wind values.

In our application, after each simulation (using a particular individual), ideal wind values are calculated and stored together with this individual. After obtaining all the simulations using a certain population, and genetic operations take place (elitism or mutation), the obtained values are assigned to the wind velocity and direction of each individual.

## 4 Experimental Results

The experimentation main goal is to evaluate the benefit of applying a dynamic data driven GA for fire spread prediction purposes. In particular, the assimilated data will be injected into the mutation and elitism operations independently and compared to applying the prediction methodology without considering any external data. For this purpose, two kind of burns have been used: two synthetic fires (simulation results) and one real prescribe fire. Fig. 4 shows the fire front evolution through time for the three experiments. In all cases, we use the best individual (with minimum error) of calibration stage to obtain the final prediction (at prediction stage).



**Fig. 4.** (a) Experiment 1: synthetic fire case. (b) Experiment 2: synthetic fire case. (c) Experiment 3: real fire case.

All cases had used as initial population a set of 50 individuals with random values. Through GA, 5 iterations are performed (error reductions are insignificant from 3<sup>rd</sup> or 4<sup>th</sup> iteration [8]). Experiment results are the average of 5 different initial populations (in order to achieve more descriptive results).

### 4.1 Experiment 1: Synthetic Case

Experiment 1 concerns with the map depicted in Fig. 4(a) (109 x 89  $m^2$  and cells of 1  $m^2$ ). The terrain had 18° slope and vegetation was fuel model 7 [3].

Figs. 5 (a) and (b) show Calibration and Prediction stages results respectively. In  $x$  axis different time intervals are shown and each vertical bar represents one of the three different methods tested: classical GA (non guided), dynamic data driven elitism operation and dynamic data driven mutation.

For both stages (Calibration and Prediction), we can observe that the error has been significantly reduced when any of the two data driven methods have been used. Furthermore, both strategies provide very similar results. Prediction stage results shows errors slightly higher than in Calibration stage, however this is an expected result because we are using the best individual obtained at the Calibration stage to obtain the prediction for a later time interval at the Prediction stage.

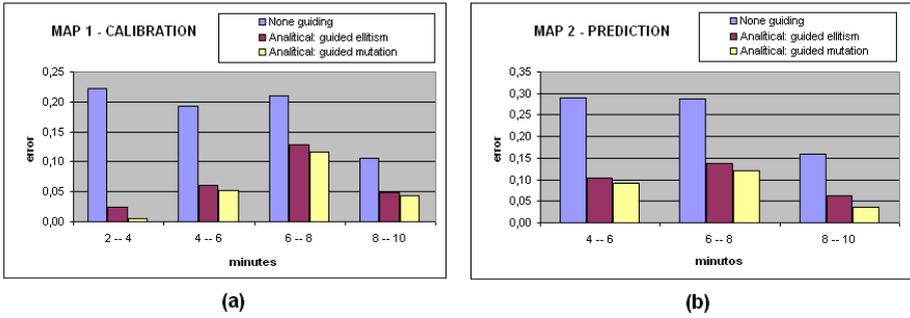


Fig. 5. Experiment 1: (a) Calibration stage results, (b) Prediction stage results

## 4.2 Experiment 2: Synthetic Case

Another synthetic map was used to carry out the second test (concerning with map illustrated in Fig. 4 (b):  $33,22 \times 27,12m^2$  size map and cells of  $1 f^2$ ). In this case, we consider  $27^\circ$  slope and vegetation is the same than the first burning case.

The error values obtained for this experiment are shown in Fig. 6. This experiment denotes a similar behavior as the first one. As it happens in the previous experiment, applying any dynamic data driven strategy reduces the calibration error as well as the prediction error.

## 4.3 Experiment 3: Real Case

The last experiment concerns with Fig. 4(c) ( $89 \times 109m^2$  map size and cells size were  $1 m^2$ ). This burn has been extracted from a set of prescribe burns performed in the frame of an European project called SPREAD. In particular, these burns were done in Gestosa (Portugal) during 2004 [6].

Fig. 7(a) and 7(b) show the calibration and prediction stages results respectively. The 2 first time terms in Calibration stage we can see that results were similar through all methods, there weren't significant differences between different configurations. But in the 3 last steps, different methods had varied in a significant way. Each method had been adapted to fire features in a different way. For Prediction stage, we can see that in general, all the errors are bigger than observed errors in synthetic fires. Again, the behavior of different methods were different for each slice of time.

We can see that errors from experiment 3 were larger if we compare them with previous experiment errors. This characteristic may be explained considering real fire map features: instead of each step lasts the same amount of time (2 minutes), fire propagation was different in each step. This changing fire behavior may be due to wind or slope or other nature characteristics influences, because of their dynamic behavior. It is important to take into account that the microsystem generated by fires may enlarge the changing characteristics of wind, weather, etc.

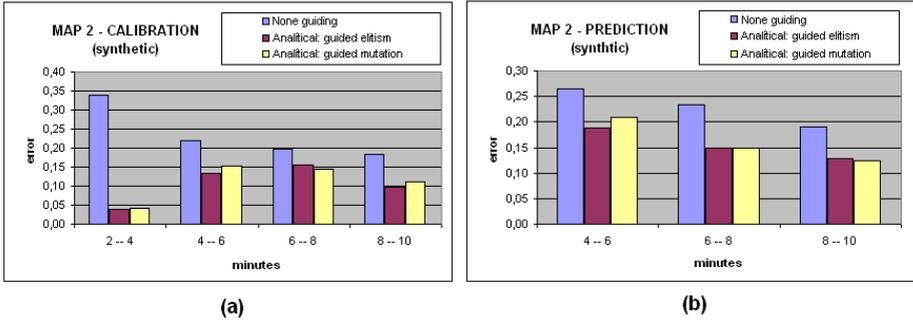


Fig. 6. Experiment 2: (a) Calibration stage results, (b) Prediction stage results

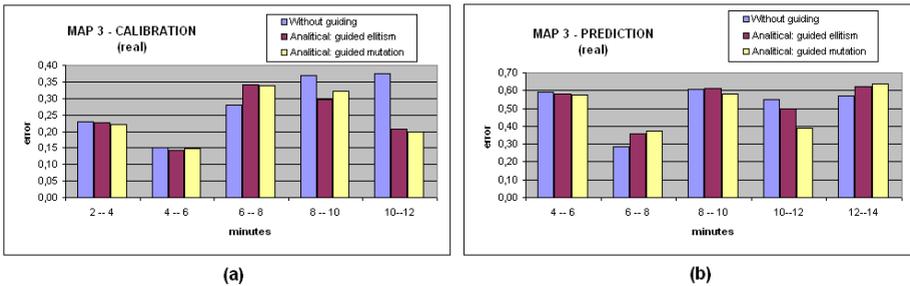


Fig. 7. Experiment 3: (a) Calibration stage results, (b) Prediction stage results

Instead of actual dynamic characteristics of nature, *fireLib* considers slope, wind, fuel type invariant in time [5]. This is an important simulator limit, because it can't adjust its fire advance in favor of actual dynamic behavior of its inputs.

## 5 Conclusions

This work is focused to improve forest fire spread prediction by using a dynamic data driven GA. For that purpose, certain internal operations of the GA are dynamically adapted according to the observed fire evolution and the underlying fire spread model.

In this work we present three cases of study, two synthetic fires and one real prescribe fire. The obtained results show that the inclusion of the principle of the dynamic data driven systems in the calibration stage improves the quality of the propagation predictions. In particular, we propose two alternative data driven GAs and the results obtained in both cases reduce the prediction error compared to not using any dynamic data driven strategy.

However, we know that this method is strongly coupled to the underlying simulator, therefore, we are currently working in generalize the main idea of the proposed methodology to be completely independent on the underlying fire forest simulator.

## References

1. Abdalhaq, B.: A methodology to enhance the Prediction of Forest Fire Propagation. Ph. D Thesis. Universitat Autònoma de Barcelona (Spain) (June 2004)
2. Allen, G.: Building a Dynamic Data Driven Application System for Hurricane Forecasting. In: Shi, Y., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) ICCS 2007. LNCS, vol. 4487, pp. 1034–1041. Springer, Heidelberg (2007)
3. Anderson, H.E.: Aids to Determining Fuel Models For Estimating Fire Behaviour Intermountain Forest and Range Experiment Station Ogden, UT 84401. General Technical Report INT.122 (April 2002)
4. 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, p. 130 (1986)
5. Bevins, C.D.: FireLib User Manual & Technical Reference (1996) (accessed January 2006), <http://www.fire.org>
6. Bianchini, G.: Wildland Fire Prediction based on Statistical Analysis of Multiple Solutions. Ph. D Thesis. Universitat Autònoma de Barcelona (Spain) (July 2006)
7. Dynamic Data Driven Application Systems homepage (November 2007), <http://www.dddas.org>
8. Denham, M.: Predicción de Incendios Forestales Basada en Algoritmos Evolutivos Guiados por los Datos, MsC thesis. Universitat Autònoma de Barcelona (Spain) (July 2007)
9. FIRE.ORG - Public Domain Software for the Wildland fire Community (accessed May 2007), <http://www.fire.org>
10. Mandel, J., et al.: A Dynamic Data Driven Wildland Fire Model. In: Shi, Y., van Albada, G.D., Dongarra, J., Sloot, P.M.A. (eds.) ICCS 2007. LNCS, vol. 4487, pp. 1042–1049. Springer, Heidelberg (2007)
11. Canarias 7 (accessed October 2007), <http://www.canarias7.es>
12. El País (accessed October 2007), <http://www.elpais.com>