Towards a Dynamic Data Driven Wildfire Behavior Prediction System at European Level

Tomàs Artés¹, Andrés Cencerrado¹, Ana Cortés¹, Tomàs Margalef¹, Darío Rodríguez-Aseretto², Thomas Petroliagkis², and Jesús San-Miguel-Ayanz²

¹ Computer Architecture and Operating Systems department, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
tomas.artes@uab.cat, andres.cencerrado@uab.cat, ana.cortes@uab.cat, tomas.margalef@uab.cat
² European Commission, Joint Research Centre, Institute for Environment and Sustainability, Via E. Fermi 2749, I-21027 Ispra (VA), Italy
dario.rodriguez@jrc.ec.europa.eu, thomas.petroliagkis@jrc.ec.europa.eu, jesus.san-miguel@jrc.ec.europa.eu

Abstract
Southern European countries are severely affected by forest fires every year, which lead to very large environmental damages and great economic investments to recover affected areas. All affected countries invest lots of resources to minimize fire damages. Emerging technologies are used to help wildfire analysts determine fire behavior and spread aiming at a more efficient use of resources in fire fighting. In this case of trans-boundary fires, the European Forest Fire Information System (EFFIS) works as a complementary system to a national and regional systems in the countries, providing information required for international collaboration on forest fires prevention and fighting. In this work, we describe a way of exploiting all the available information in the system to feed a Dynamic Data Driven wildfire behavior prediction model that can deliver results to support operational decision. The model is able to calibrate the unknown parameters based on the real observed data, such as wind condition and fuel moisture using a steering loop. Since this process is computational intensive, we exploit multi-core platforms using a hybrid MPI-OpenMP programming paradigm.

Keywords: DDDAS, Input Data Uncertainty, High Performance Computing, Calibration, Hybrid MPI-OpenMP, Applications, Evolutionary Computation, Forest Fires, EFFIS

1 Introduction
Europe suffers approximately 65,000 fires every year, which burn, on average, half a million hectares of forested areas [1]. The main direct effect of forest fires is the destruction of the natural landscape and the consequent loss of ecosystem services that have drastic economic
impact. In addition, and no less important, fires also result in the loss of human lives every year. Although being forest fires a problem present in all EU members, the most affected areas to this kind of hazards are the southern countries due to their meteorological conditions, specially during summer season. All affected countries invest lots of resources to minimize fire damages, but many times when dealing with large fires, regional and national disaster management units are lack of efficient and reliable tools to help wildfire analysts.

Wildfire spread models/simulators are the basic element for designing forest fire prediction frameworks to be used as a decision support systems during an ongoing disaster. However, it is well known that the forecast fire evolutions provided by existing fire spread simulators do not exactly reproduce the real behaviour of the fire. The reason for such a difference ranges from the input parameters uncertainty to the imprecision of the model itself. In previous works [2, 3, 4], it has been stated that a pre-processing of the simulator input parameters based on a steering loop driven by real data acquisition and fire behaviour observation, could lead to enhanced forecast fire evolutions. Since forest fires are a dynamic phenomena, which is quite affected by changing data such as meteorological information, the mentioned pre-processing data phase has been designed as a feedback loop where gathered data guides the simulation and, the simulation results at a time, could eventually drive the data collection. This way of work feeds the so called Dynamic Data Driven Application System [5, 6]. However, not all the information needed by the prediction system could be considered in the calibration process. In particular, those data considered as static information such as elevation maps and fuel data, is typically obtained from public repositories. However, there is not an unique source of this kind of data and, consequently, the fire behaviour prediction delivered by a given forest fire spread model could vary according to the selected static data sources. Furthermore, gathering dynamic data such as real fire perimeter evolution and meteorological data could be a bottleneck to the system if there is no a clear way to proceed.

So, in order to provide the EU community with a reliable forest fire spread prediction system, it would be mandatory to have standard actuation protocols at European level, specially when dealing with cross-boundaries events, that specify a common data source repository to be used for fire spread prediction purposes. EFFIS (European Forest Fire Information System) raises as the EU common platform for this kind of assessment. Although EFFIS provides comprehensive information on forest fires at EU level, the component devoted to the prediction of forest fire spread is still at an initial stage. This work is a prove of concept to update the system to support near-real time forest fire spread forecast which can help on operational decisions specially on trans-boundary forest fires, which need international collaboration. Two kind of strategies are adopted to test the viability of including a prediction module on EFFIS. First of all, a basic prediction scheme has been tested using as source data all the information available at EFFIS and, the second approach consists of applying a DDDAS for forest fire spread prediction where certain meteorological data values are tuned on-line. Since the prediction results delivered by the DDDAS scheme must be obtained in a short time in order to be useful for mitigation purposes, a hybrid programming scheme has been adopted to exploit the computing capabilities of the current multi-core architectures.

Subsequently, in subsection 2 the calibration strategy based on a hybrid DDDAS scheme is detailed. In section 3 an overview of the European Forest Fire Information System is reported. A prove of concept is presented in section 4 based on a forest fire that took place in Greece in 2011. And finally, the main conclusions of this work are described in section 5.
2 Dynamic Data Driven Wildfire Behaviour Prediction

In this section, we describe in more detail, how the above mentioned dynamic data-driven prediction scheme used to calibrate certain input parameters of the underlying simulator works [7]. Figure 1 sketches how the two-stage approach is implemented. As it can be observed, in order to launch an adjustment stage, it is necessary to have access to two consecutive real fire perimeters. Each time the system could be feed with a new observed fire behaviour, the calibration process could be initiated again. The goal is to have the system working in a continuous fashion, to provide fire evolution forecasts on continuous preset times horizons (steps $t_i$, being the first two-stage prediction delivered at a certain time between $t_1$ and $t_2$). So, the input data set used for the prediction stage is calibrated in the corresponding adjustment stage.

This methodology is simulator independent and it is flexible enough to change the simulator in a plug&play way. In this work, the used forest fire spread simulator is FARSITE [8], which is based on the Rothermel model [9]. The calibration/adjustment strategy used is Genetic Algorithms (GA), which has been demonstrated to perform well for this problem [7, 3]. The GA starts from an initial random population of individuals, each one representing a scenario to be simulated. A scenario/individual is composed of a number of different genes that represent input variables such as dead fuel moisture, live fuel moisture, wind speed and direction, among others. Each individual is simulated and it is evaluated comparing the predicted and real fire propagation. The representation of fire evolutions is carried out using a cell map that indicates whether a cell has been burnt or not at a certain time (time of arrival). Therefore, map differences are evaluated by the fitness function described in equation 1. This fitness function computes the symmetric difference between predicted and real burned areas.

\[
\text{Difference} = \frac{\text{UnionCells} - \text{IntersectionCells}}{\text{RealCells} - \text{InitCells}}
\]  

(1)

In equation 1, $\text{UnionCells}$ is the number of cells which describe the surface burned considering predicted fire and the real fire. $\text{IntersectionCells}$ is the number of cells burned in the
real map and also in the predicted map, and $\text{RealCells}$ are the cells burned in the real map. $\text{InitCells}$ is the number of cells burned at the starting time. This difference takes into account the wrong burned cells and the mistaken for burned cells. According to this fitness function the whole population is ranked and the genetic operators selection, elitism, mutation and crossover are performed over the population, producing an evolved population which will have, at least, the best individual of the last generation (elitism). The new population is then evaluated in the same way. This iterative process allows us to find a good input parameter set, but it involves high computational cost due to the large amount of simulations required. Therefore, it is essential to speed up the execution keeping the accuracy of the prediction. For this reason, an implementation of the two stage methodology has been developed using High Performance Computing techniques. Since the GA fits the Master/Worker paradigm, an MPI implementation has been developed. At the first stage, the master node generates an initial random population which is distributed among the workers. Then, the workers simulate each individual and evaluate the fitness function. The errors generated by the workers are sent back to the master, which sorts the corresponding individuals by their error before applying the genetic operators and producing a new population. This iterative process is repeated a fixed number of times. The last iteration (generation) contains a population from which the best individual is taken as the best solution, and then it is used in the prediction stage. Since every simulation can be carried out in a parallel way, the individual whose simulation takes longer determines the elapsed time for that particular generation. In order to shorten simulation times, FARSITE has been analyzed with profiling tools such as OmpP and gprof to determine which regions of the code could be parallelized with OpenMP. The result of such analysis determined the particular loops that could be parallelized using OpenMP pragmas. The results of such parallelization have been presented in [13]. The parallelized loops represents about 60% of one iteration execution time.

3 European Forest Fire Information System (EFFIS)

The European Forest Fire Information System (EFFIS) was established by the Joint Research Centre (JRC) and the Directorate General for Environment of the European Commission, in close collaboration with the Member States and neighbour countries, to provide harmonised information required for international collaboration on forest fire prevention and fighting, especially in cases of trans-boundary fire events. The EFFIS system supports the services in charge of the protection of forests against fires in the EU countries and provides the European Commission services and the European Parliament with updated and reliable information on wildland fires in Europe. Since its first conception as a system providing fire risk forecast and fire danger assessment, the EFFIS system has evolved to become a fire monitoring system that comprises all the phases of forest fires, from the pre-fire assessments to the post-fire estimation of damages and the analysis of vegetation recovery [1]. However, one missing component in this current composition of the system is a fire spread prediction module, which allows to forecast, in near real-time, the areas threatened by the fires in process across Europe. Figure 2 shows the modules that are currently operative at EFFIS. Related to Active Fire Detection, one can found two modules Fire Detection and Burnt Area Maps. These modules are feed with data providing from EU Fire Databases and satellite images (MODIS images). The Fire Detection geoparses the fire news with the MODIS hot spots, in order to identify fire burnt areas perimeters. Those perimeters are available at EFFIS from the Burnt Area Map module. The proposed new EFFIS module will be directly connected to this modules, specially to the Burnt Area Map as it is shown by the red cicle in figure 2. The inclusion of a DDDAS prediction scheme will lead to a
Decision Support System, which could support operational decisions at European Level.

The dynamic data driven wildfire behaviour prediction system described in section 2 requires, for being operational, near past information for calibration purposes and, forecast data at the prediction stage. As it has been previously stated, input data stems from different sources of errors that will affect the predictions results. Since the nature of the input data is quite different going from static information such as the topography of the terrain, to the very dynamic meteorological data such as the wind speed and wind direction, it is very helpful to rely on common data sources when dealing with trans-boundaring events. However, working within a common data framework does not relive the system from input data errors, but at least homogenizes the error treatment and the information delivered to the wildfire analysts. In the following, it will be described how each of these information is considered by EFFIS:

1. **Topography:** This data is obtained by processing Digital Elevation Maps (DEM). A DEM defines the height of the terrain in every cell of the map. This is a discretization of a continuous surface, taking into account measures in certain points of the terrain. Those maps are obtained from the ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) imaging instrument onboard NASAs Terra satellite that takes high-resolution images of the earth. These images are processed and raster files are extracted with the information needed to perform the fire spread simulations (elevation, aspect and slope). The ASTER map resolution is 30m.

2. **Fuel map:** The vegetation map (or fuel map) is a raster file that describes the predominant vegetation in every cell. The fuel model used for fire simulation purposes, is the standard
fuel model defined by [14]. This information has been obtained from the fuel type map of Europe developed at the JRC. The classification scheme adopted for the fuel map encompasses 42 fuel types representing the variety of fuel complexes found in European landscapes. A crosswalk to the original set of 13 fire behaviour fuel models tabulated by Anderson fire spread model is done at the JRC [14].

3. Meteorological data: The meteorological data used is the ECMWF (European Centre for Medium-Range Weather Forecasts) operational high-resolution single global deterministic model (ec16), with a horizontal resolution of about 16 km [15]. The model is initiated on both the 00 and 12 UTC analysis fields reaching to a 24-hour forecast horizon with archived time-step of 3 hour. It is worth mentioning, that is the configuration that it has been used in this work for the experimental proposes, but it is not the unique source of meteorological data processed at the JRC.

4. Fire Perimeter: One key point in the DDDAS scheme for wildfire spread prediction previously described, is the evolution of the fire, that is, the capability to feed the system with real observed fire perimeters in a systematic time step. As much accurate the image/perimeter processing is, more effective will be the steering loop. The EFFIS Burnt Area Map module is in charge of this data. To obtain the fire perimeter information, the JRC relies on the MODIS (Moderate Resolution Imaging Spectroradiometer) sensors systems, which are on board both the NASAs Terra and Aqua satellites. Each satellite requires to complete 3 orbits (approximately 3 hours) to cover the whole Europe area, so it could be possible to obtain fire perimeters twice a day, one from each satellite [16]. The image resolution provided by the MODIS system is of 250m.

All these data inputs have been harmonized to fit a simulation grid map with a basic cell of 100mx100m square. Consequently, this resolution will be the one used in the forest fire spread prediction model when applying the dynamic data driven wildfire prediction strategy. In the following section, the DDDAS approach for forest fire behaviour prediction is applied to a case study by using all the above described data provided by EFFIS.

4 Case of Study

As it has been previously stated, the experiments reported in this work, are the first step towards a methodology to exploit the paneuropean data available at the Joint Research Centre within EFFIS, in order to couple into the system an operational forest fire spread prediction component. For that reason, we have retrieved the information of a past fire that took place in Greece during the summer season of 2011 in the region of Arkadia. In figure 3(a), the images provided by the MODIS system are shown for three different time instants. The corresponding burnt areas (shapes) once the images have been processed, are shown in figure 3(b). These shapes are the information available at EFFIS and they are the perimeters used to apply two prediction approaches: a basic forest fire prediction scheme where no calibration is applied and the Dynamic Data Driven prediction scheme described in section 2 where a Genetic Algorithm is used to calibrate certain input simulator parameters. In order to simplify the initial tests, the forecast meteorological data used for the simulations are the wind components (wind speed and wind direction), dew point and temperature of the pinpointing centroid of the observed fire. These data values are considered homogeneous distributed through out the terrain for a given time step. This time step has been set to 3 hours because of the arrival frequency of the meteorological data. The prediction time horizons have been set according to the exact time of
the MODIS images have been obtained and the prediction scheme applied. On the one hand, when applying a basic prediction scheme, the initial fire perimeter is the one observed by the Terra satellite on August 26th at 09:43 pm. The time horizon has been set to almost one day (24 hours) because the perimeter to be predicted is the one obtained on August 27th at 08:49 am. On the other hand, when dealing with the DDDAS prediction scheme the time horizon varies because of the inclusion of the calibration stage. The simulations executed during the calibration stage took as a reference perimeter for calibration purposes, the one obtained from the Aqua satellite around 11:27 am of August 26th, so those simulations have been set to a time horizon around 2 hours. However, the simulation executed at the prediction stage has a time horizon of 22 hours. Subsequently, the two prediction schemes shall be compared according to their quality results.

We want to show the viability of annexing a prediction module to the EFFIS system by analyzing the prediction results delivered by the two prediction schemes previously described. Therefore, the main goal consists of verifying that EFFIS data could properly be used to feed an eventually EFFIS prediction module providing this a feasible output independently on the prediction strategy applied. For that reason, two prediction strategies have been tested, the Basic prediction scheme and a DDDAS approach where the observed fire evolution is used in a steering loop to improve the prediction results.

For both strategies, different settings of the meteorological data has been used to analyze the influence on the prediction quality. Those settings define the following test cases:
1. **Homogeneous Spatial and Homogeneous Temporal (HSHT) data**: This setting corresponds to use as meteorological data, the forecast information available at the fire perimeter acquisition time and in the pinpointing fire centroid. Those values (wind speed, wind direction, dew point and temperature) are considered constant throughout the terrain and invariable through time.

2. **Homogeneous Spatial and Variable Temporal (HSV T) data**: In this case, the pinpointing meteorological information keeps constant at each cell of the terrain, but the forecast data is injected into the system every 3 hours, that is, the meteo data varies as the prediction time goes on.

3. **Variable Spatial and Homogeneous Temporal (VSHT) data**: In this case, a spatial distribution of the wind data is used by executing a wind field model (WindNinja [17, 18]) for the forecast wind data at the image perimeter capture time.

4. **Variable Spatial and Variable Temporal (VSV T) data**: This parameter setting implies the variation in time and throughout the terrain of the meteorological data.

The prediction results obtained when applying the four above described meteorological data settings to the Basic and DDDAS prediction approaches are shown in figure 3. Each individual picture of figure 3 shows, in green colour, the burnt area on August 27th at 08:49am, which is the perimeter to be predicted and, with dotted lines, the forecast perimeter. The images located at the left column of the figure correspond to the results obtained when applying a Basic prediction scheme where no input parameter calibration is performed. On the other hand, the propagation results depicted on the right column images reflect the results provided by the DDDAS scheme. As it can be observed, the worst prediction results are the ones obtained when the effect of the terrain in the meteorological (wind speed and wind direction in this case) data is not considered.

Both predictions strategies perform poor when the wind is considered homogeneous throughout the terrain. As it can be observed from the pictures, meteorological settings where the terrain is taking into account, in general provides prediction results with a wider burnt area that the predictions obtained without considering the topography of the terrain. That is a key point when dealing with real wildfires because the slope of the terrain directly affects the rate of spread of the underlying propagation model and, dismissing such a effect could lead to a catastrophe. Another aspect to outline, for this test case, is that in general, the burnt area predicted is wider than the real observed burnt area. These results corresponds to one test case, so we can not extrapolate direct conclusions. However, one can not forget that any wildfire used as study case, have a spread behaviour influenced by mitigation actions and human intervention. This aspect has a clear influence on the fire final burnt area.

For that reason, those results that predict more burnt extension than the real burnt observed area, could not be considered as bad as it could look like. In this study case, we can stated that using a meteorological setting where those variables which dynamic behaviour could be incorporated to the prediction system, that is Variable Spatial and Variable Temporal (VSV T), provides better results. It is worth to remark, that in this particular case, the wind speed and wind direction keep quite similar during the entire prediction interval, therefore, the DDDAS scheme performs almost the same that not using a calibration stage prior to the prediction stage.
Figure 4: Prediction results applying Basic and DDDAS approaches for all meteorological settings HSHT, HSVT, VSHT, VSVT.
Although the advantages of the DDDAS scheme has been widely demonstrate for prescribed fires and small real fires, when moving to large wildfires as the one presented in this work, a deeper analysis of more study cases should be reported to be able to determine the improvements in terms of quality results when using the DDDAS approach. However, the main objective of this work has been accomplished since it appears to be viable to design a prediction module for EFFIS that uses standard data sources for provided wildfire behaviour predictions.

5 Conclusions

Forest fires is one of the most proud natural hazard across Europe and, in particular, in the Mediterranean area. For this reason, any effort oriented to support wildfire analyst when a hazard occurs is welcome. Forest fire spread simulators provides wildfire evolution forecasts based on the observed and gathered environmental data, therefore, the quality and reliability of such data is a critical point in fire simulation. A major aspect that must also be considered, specially when dealing with trans-boundary fires, is the agreement in the databases used to retrieve the input data required for the phenomena simulation.

At European level, EFFIS was established to provide harmonized information on such as hazards to be used on international forest fire management, therefore, it can be considered the meeting point for fire management efforts during European international wildfire events. So, EFFIS becomes the facto, the standard database for European disasters. However, one missing module in EFFIS is a prediction module. This work is a prove of concept to test the viability of exploiting EFFIS data for eventually annexing a prediction module to the system, to be used as a support tool during an ongoing forest fire. So, it is the first step towards a prediction system at European level. Although being this a step forward to a standard platform for forest fire management, one can not dismissed that the accuracy and precision of the data is limited due to the different kind of errors introduced when generating/collecting it.

For that reason, in this paper we have analyzed two alternatives for the above mentioned EFFIS’ prediction module: a Basic fire spread simulation scheme and a Dynamic Data Driven wildfire behaviour prediction approach. In the Basic scheme the fire spread simulator FARSITE is used to perform a raw fire spread simulations with the available data. In the DDDAS approach, a calibration stage is introduced to take advantage of all this data to drive the simulations to better prediction results. Both schemes have been tested using as a study case a forest fire that took place in Arkadia (Greece) in 2011. In order to consider any kind of data available at EFFIS, both schemes have been proved either considering or discarding the injection of dynamic meteorological data. The results show that using a DDDAS scheme will lead to better prediction results.

It is worthy mentioning, that any forest fire analyzed in a post-mortem way, as it has been done in this paper, suffers from the penalty of not being a completely free burning fire, because of the human contention actions. Therefore, better predictions are the ones that overestimate the observed burnt area, as it happens when the DDDAS prediction is performed.

Acknowledgments

This research has been supported by MICINN-Spain under contract TIN2011-28689-C02-01.
References


