International Conference on Hydroinformatics

2014

# HPC For DRM - Operational Flood Management In Urban Environment

Evžen Zeman

Pavel Tachecí

Johan Nicolai Hartnack

Vít Vondrák

Jan Martinovič

11th International Conference on Hydroinformatics
HIC 2014, New York City, USA

# HPC FOR DRM - OPERATIONAL FLOOD MANAGEMENT IN URBAN ENVIRONMENT

EVŽEN ZEMAN (1), PAVEL TACHECÍ (1), PAVEL ŠTROF (1), JOHAN NICOLAI HARTNACK (2), OLE RENE SØRENSEN (2), JESPER CARLSON (2), VÍT VONDRÁK (3), JAN MARTINOVIČ (3), ŠTĚPÁN KUCHAŘ (3)

*(1): DHI a. s., Prague, Czech Republic*

*(2): DHI, Hørsholm, Denmark*

*(3): IT4Innovations, VŠB-Technical University of Ostrava, Czech Republic*

## Abstract

A regional flood warning system based on a combination of data processing, modelling and communication tools is proposed. The system is founded on a common framework MIKE CUSTOMISED, giving great flexibility in tailoring solutions as needed. Where more conventional flood warning systems focus mainly on discharge predictions in the main rivers the proposed system considers the whole catchment area – flood plain, as well as tributaries. Local floods on smaller streams and tributaries may cause high damages, particular in urban areas. Such cases require timely reasonably accurate forecasts for proper decision making.

The MIKE SHE modelling system is used to ensure simulation of flood danger map. It also provides runoff hydrographs for the detailed hydrodynamic models of the river channel network and flood plains based on 1D / 2D approximations (MIKE FLOOD). Generated flood maps are post processed and ported to the required forms and delivered via communication channels to users. Dissemination of results is done through web pages automatically maintained by the system.

The complete forecast simulation must be run at a frequency of tens of minutes making the system demanding with respect to computational power and transmission capacity. For this reason the system is being adapted to a High Performance Computer (HPC) solution, see *Hager et al.* [1]. The use of HPC also allows parallel variant computations and real-time probabilistic assessments. IT4Innovations National Supercomputing Centre is a research institute at the VŠB – Technical University of Ostrava. This centre provides the perfect platform for applications of HPC for Disaster Risk Management (DRM), utilised for life protection and damage minimization.

The paper contributes both to the theory of application of HPC for standard hydrodynamic modelling but also to a real life application. A concept for operational Flood Risk Mapping projects has been developed and offered for several potential users in the region where such efficient solution could solve the operational issues of DRM.

## Introduction

Risk and uncertainty have become fundamental aspects of most Decision making processes. Qualitative risk assessment is usually based upon perceptions, opinions, judgment, consent, public or political consensus (or a combination here of), whereas quantitative risk assessment is based on the results of modelling techniques. Modelling is used to express various component of the risk, while taking into account the probability of whether the loss or damage will occur. Disaster Risk Assessment can be regarded as a process of determining qualitative and quantitative aspects of risk related to an actual situation. Therefore, the main purpose of Disaster Risk Management (DRM) is to obtain a comprehensive understanding of the disaster risk before identifying mitigation initiatives that are likely to reduce the risk (see Figure 1).

Some of the threats are combined or multiplied with other hazards because the event is a combination of parallel or consecutive dynamical effects e.g. the flooding of a flood plain and damaging water resources by the flooding of open wells. Evaluating the combined risk processes requires integrated modelling tools. Modelling tools may be used on-line or off-line. Models could be integrated only if the models are prepared for such integration. Integration of models in principle influence computing time in negative way.

Often analysis of disaster risk phenomena requires modelling utilizing on-line monitored data. In this context an important aspect is the relation between computing time and real time of the modelling run. Even powerful personal computers are not able to run the models fast enough to get results in time basically for prediction and forecast but mainly for the response phase.

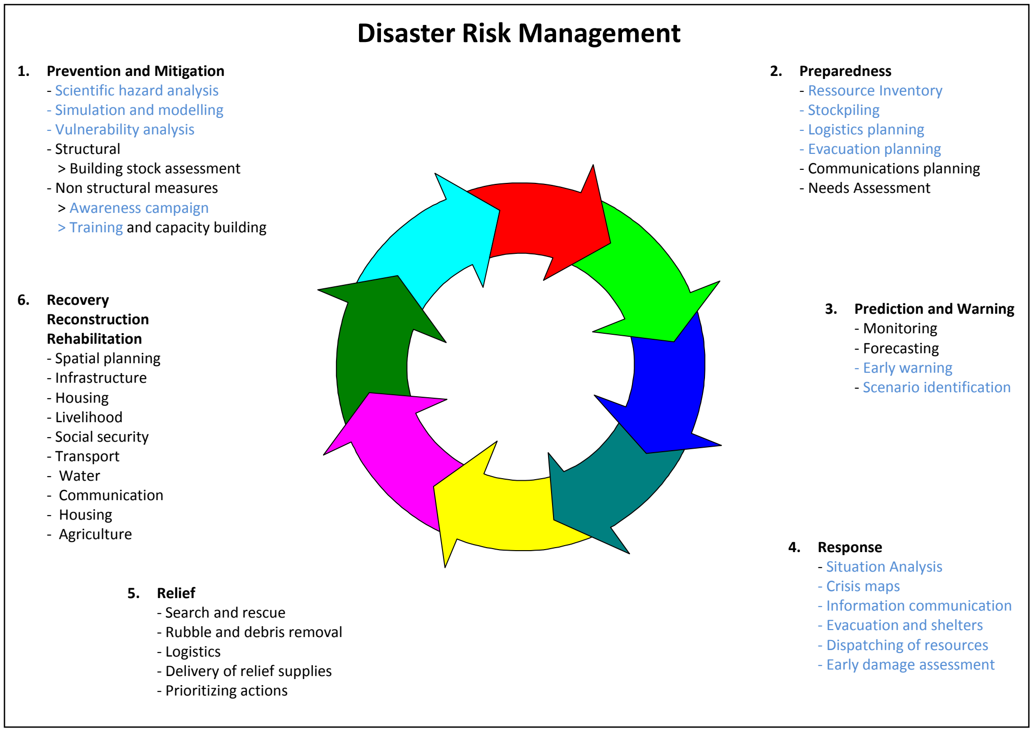

Figure 1 Phases of Disaster Management

## MIKE21 FM

DHI’s Flexible Mesh (FM) modelling system, which has been developed for applications within ocean, coastal, estuarine and riverine environments, has been parallelised based on the message passing paradigm to enable utilisation of distributed-memory systems. This extension makes it possible to utilize massively parallel computers and clusters. The distribution of work and data is based on the domain decomposition concept, and Message Passing Interface (MPI) is used for communication between processors, see *Pacheco* [2]. The computational mesh is partitioned into a number of physical subdomains, and the work associated with each subdomain is processed by an individual processor. The data exchange between processors is based on the halo-layer approach with overlapping elements.

The use of massive parallel computers makes it possible use MIKE 21 Flow Model FM (see *DHI* [3]) for real-time flood inundation forecasting and calculations of high resolution two-dimensional (2D) flow solutions over irregular floodplain topography. MIKE 21 Flow Model solves the 2D shallow water equations. The spatial discretisation is performed using a cell-centered finite volume technique and an unstructured mesh approach. An approximate Riemann solver (Roe’s scheme, see *Roe* [4]) is used to calculate the convective fluxes. Second-order spatial accuracy is achieved by employing a linear gradient-reconstruction technique. The average gradients are estimated using the approach by *Jawahar et al.* [5]. To avoid numerical oscillations, a second-order total variation diminishing (TVD) slope limiter (Van Leer limiter, see *Hirsch* [6]) is used. An explicit Euler method or a second-order Runge-Kutta method is used for the time integration. A detailed description can be found in *DHI* [3].

## Anselm HPC Cluster Architecture

The Anselm HPC cluster consists of 209 computational nodes. Each node is a powerful x86-64 computer, equipped with 16 cores (two eight-core Intel Sandy Bridge processors), at least 64GB RAM, and local hard drive. These nodes are interlinked by high speed InfiniBand (3600MB/s) and Ethernet networks. All nodes share a Lustre parallel file system with a throughput of 6 GB/s.

There are four types of compute nodes:

- 180 compute nodes without any accelerator,
- 23 compute nodes with GPU accelerators (NVIDIA Tesla Kepler K20),
- 4 compute nodes with MIC accelerators (Intel Xeon Phi 5110P) and
- 2 fat nodes with larger RAM and faster storage (512GB RAM and two SSD drives).

Total theoretical peak performance of the cluster is 82 Tflop/s with maximal LINPACK performance being 66 Tflop/s.

## Performance of MIKE21 FM

Mesh based hydrodynamic model parallelization on HPC clusters allows users to increase the mesh resolution significantly while maintaining useable runtime. Larger problems can be processed by the increased available memory and added processing capabilities.

The MIKE21 FM simulations were run on a number of standard compute nodes without accelerators on the Anselm HPC cluster to see how the problem and models scale on different number of cores. This model was tested on a processor range of between 16 and 2,048 cores and the resulting speedup is shown in Figure 2. (The speedup experimentation was done for a 2D hydrodynamic model simulating the tides in the Mediterranean Sea.)

The results show that the speedup is linear to the number of cores for up to 512 cores used. For higher numbers of cores the speedup shows a slower increase due to the high amount of communication between individual processes and small size of each subdomain of the mesh.

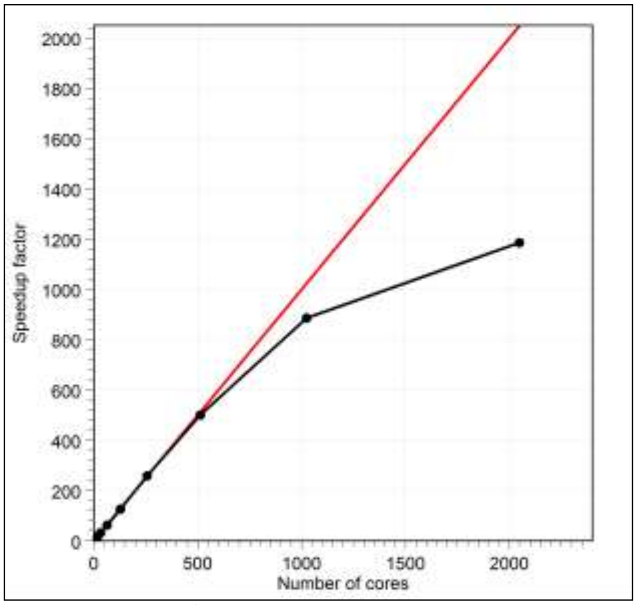

Figure 2 Scalability of MIKE21 FM model

The performance of MIKE 21 Flow Model FM is furthermore evaluated using a high resolution model setup for river flood forecasting and modelling. The mesh contains 995,019 triangular elements with an average element area of approximately 50 $m^2$. The simulation period for the benchmarks is 1 hour and the average time step is 0.13 s. At the upstream boundary a discharge boundary is applied and at the downstream boundary a level boundary is applied. For the spatial and time discretisation, the first-order scheme is applied.

The simulations were performed on the Anselm Cluster. Using 512 processors (cores) the computational time was 57.33 s. This corresponds to a speed up of a factor 21 compared to a simulation with 16 processors.

To achieve a good speedup it is important that the subdomains have enough wet elements to calculate. Dry elements uses negligible calculation time and a full subdomain of dry elements would be idle most of the calculation time. To get optimal performance the subdomains should have an even distribution of wet elements. The discrepancies between the speedup achieved with the 2D Mediterranean Sea model and the flood forecast model comes from the partitioning of the two models. The Mediterranean Sea model contains very few dry elements whereas the flood forecast model contains a large amount of dry elements. The inherent changes in the number of dry/wet elements in a flood forecast model as the time step progresses can cause a partitioning which was ideal to be suboptimal.

## Warning system description

A regional flood warning system based on a combination of data processing, modelling and communication tools is proposed. Where more conventional flood warning systems focus mainly on discharge predictions in the main rivers the proposed system considers the whole catchment area – flood plain, as well as tributaries. Local floods on smaller streams and tributaries may cause high damages, particular in urban areas. Such cases require timely reasonably accurate forecasts for proper decision making. Flash floods are caused by locally concentrated precipitation events, higher than available retention capacity. Runoff response is determined by two categories of factors – those with small temporal dynamics (which are encapsulated to calibrated rainfall-runoff models) and dynamic factors e. g. precipitation input and initial soil water content. Adopted modelling approach is based on distributed running-on simulation of soil water balance.

Three levels of flood warning are adopted (see Figure 3). 1st level of warning (Caution) is based on regional weather forecasting services results with a lead time in the order of days (e.g. 24 hours). Areas of low retention storage in soil, where high rainfall totals may cause floods are delimited. 2nd level of warning has a lead time of 1-2 hours and uses precipitation field forecast based on on-line data measurement and weather radar products. Maps of flash flood danger are produced. 3rd level of warning (Flood Alarm) is based directly on on-line monitoring data from selected rain gauges.

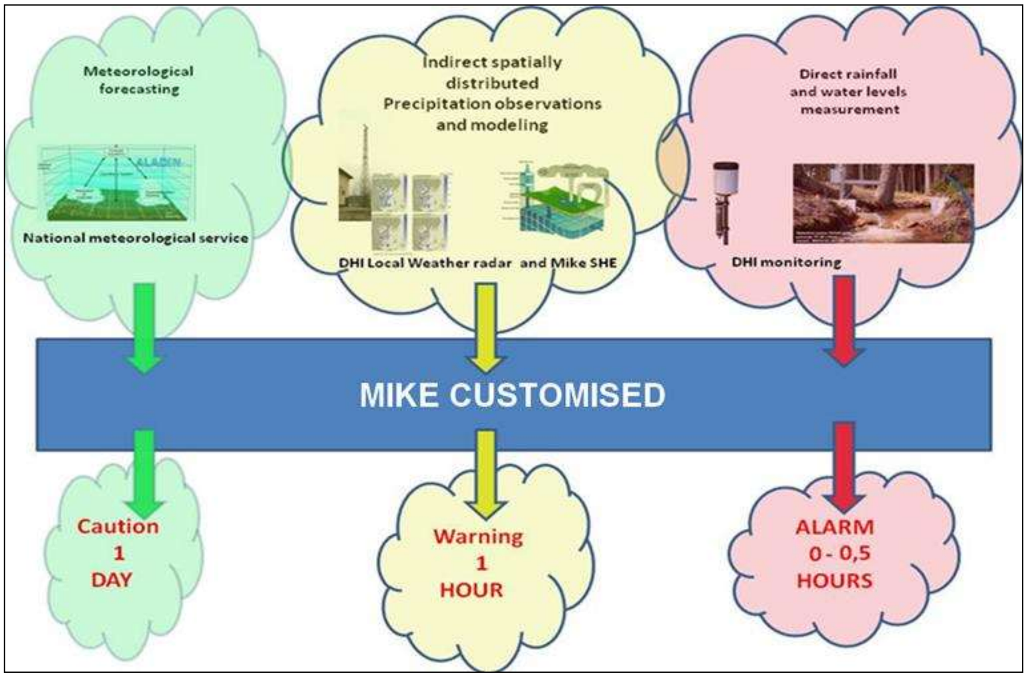

Figure 3 Scheme of three levels of warning

The system is founded on a common framework MIKE CUSTOMISED, giving great flexibility in tailoring solutions as needed. MIKE CUSTOMISED serves a platform with wide range of functionality (time series tools, GIS processing tools, scenarios building, job scheduling, scripting, assessment of uncertainty etc. The data handling tool DIMS.CORE is part of the platform, covering data-import, -storage and -management. Ground based observations (precipitation and water level observations) are combined with meteorological radar data, which may originate from previously installed device, or newly established local weather radar with high resolution (DHI LAWR solution).

A lumped (NAM see *DHI* [7]) or distributed (MIKE SHE) rainfall-runoff models are used for runoff hydrograph simulation. The distributed deterministic modelling system MIKE SHE is used to update initial conditions for the next run of the model. It is also used to simulate instantaneous soil retention capacity and water balance for the forecast period. Finally, from the results flash flood danger maps are determined.

At the same time rainfall-runoff models also provides runoff hydrographs for the detailed hydrodynamic models of the river channel network and flood plains based on 1D / 2D approximations (MIKE FLOOD). Generated flood maps are post processed and ported to the required forms and delivered via communication channels to the users. Dissemination of results is done through a set of web pages automatically maintained and updated by the system.

The need for high computational power is required when utilizing spatially distributed hydrological modelling and combined 1D and 2D hydrodynamic simulations.

## Conclusion

The use of HPC is the key in delivering timely detailed hydrodynamic modelling results for use in disaster risk management. The HPC improves the execution time of the individual runs but also empowers the system to evaluate multiple scenarios when needed. The prerequisite for harvesting these HPC benefits are:

- HPC framework which can support and maintain real time simulations,
- HPC enabled hydrodynamic modelling software i.e. use of parallelization techniques which scale well on computer clusters,
- flexible software framework which can adapt to the data input and required output.

The proposed system covers all of the above and is thus well suited for adaptation in disaster risk management processes.

The time consuming performance of combined 1D and 2D hydrodynamic simulation could be effectively eliminated with the usage of HPC, and then fully dynamic flood hazard and risk simulation could be utilised for the Decision Support Systems. The new pathway is opened by this way for all stakeholders in DRM to use downscaled operational on-line system for flood prediction with all necessary details and links to contingency and evacuation planning in densely populated area or in very complex industrial installations where the environmental risk could be high.

## AKNOWLEDGEMENT

This work was supported by the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070), funded by the European Regional Development Fund and the national budget of the Czech Republic via the Research and Development for Innovations Operational Programme, as well as Czech Ministry of Education, Youth and Sports via the project Large Research, Development and Innovations Infrastructures (LM2011033).

## REFERENCES

[1] Hager G. and Wellein G., "*Introduction to High Performance Computing for Scientists and Engineers"*, Boca Raton, FL: CRC Press, (2011).

[2] Pacheco P. S., "*Parallel Programming with MPI",* San Francisco, Calif.: Morgan Kaufmann Publishers, (1997).

[3] DHI, *"MIKE 21 & MIKE 3 FLOW MODEL FM, Hydrodynamic and Transport Module, Scientific Documentation"*, DHI, (2013).

[4] Roe P. L., “Approximate Riemann solvers, parameter vectors, and difference schemes”, *Journal of Computational Physics*, Vol. 43, (1981), pp. 357-372.

[5] Jawahar P. and Kamath H., “A high-resolution procedure for Euler and Navier-Stokes computations on unstructured grids”, *Journal Comp. Physics*, Vol. 164, (2000), pp. 165-203.

[6] Hirsch, C., “Numerical Computation of Internal and External Flows”, *Computational Methods for Inviscid and Viscous Flows*, Wiley, Vol. 2, (1990).

[7] DHI, *“MIKE 11 Reference Manual”*, DHI, (2013).