

## **IDD AND MESOSCALE MODEL INTEGRATED SYSTEM**

Lia Martins Costa do Amaral, C. T. Homann y Y. Yamasaki

Federal University of Pelotas, Pelotas, RS, Brazil

(Manuscrito recibido el 17 de marzo de 2010, en su versión final el 7 de noviembre de 2010, aceptado el 20 de noviembre de 2010)

### **ABSTRACT**

As part of the group of research institutions and universities participating in the UNIDATA program, the Federal University of Pelotas (RS-Brazil) has implemented an automatic data acquisition with forecasting and synoptic analysis system. It is based on the data distribution system (IDD) of the UNIDATA (University Data Project) and uses the Local Data Management module (LDM). This is an integrated system which has been primarily implemented to use the GFS/NCEP global forecasting model data for the processing of the regional MM5 mesoscale model, providing the necessary initial and lateral boundary conditions data. The MM5 has been configured to provide the forecast of the mesoscale systems over the State of Rio Grande do Sul (RS). The use of the data and model integrated system contributes to improve the skill analysis of the model output, aiming to investigate the benefits of its association in the evaluation of the performance of the integrated system.

*Keywords:* mesoscale, data, MM5, IDD, LDM.

## **IDD Y EL SISTEMA INTEGRADO DEL MODELO DE MESOESCALA**

### **RESUMEN**

Como parte del grupo de instituciones de investigación y universidades que participan en el programa UNIDATA, la Universidad Federal del Pelotas (RS-Brasil) ha implantado un sistema automático de adquisición de datos, pronóstico y análisis sinóptico. Se basa en el sistema de distribución de datos (IDD) de la UNIDATA (University Data Project) y utiliza el módulo de gestión de datos local (LDM). Este es un sistema integrado que se ha implementado para usar los pronósticos de modelos globales GFS/NCEP para la inicialización y regeneración de los modelos regionales; utilizando las facilidades del IDD y del LDM. El MM5 se configuró para los pronósticos de mesoescala sobre el estado de Rio Grande del Sur (RS). El uso del sistema de datos y el modelo integrado contribuye a mejorar el análisis de habilidad de la salida del modelo, con el objetivo de investigar los beneficios de su asociación en la evaluación del desempeño del sistema integrado.

*Palabras Clave:* mesoescala, datos, MM5, IDD, LDM.

## 1. INTRODUCTION

Modern world's demands as well as the increasing support requirements for society's economical planning are increasingly strengthening the technical and scientific investments to supply the needs for monitoring and weather forecasting systems advance, development and improvement.

The numerical weather prediction is the primary and essential component of these systems, despite innumerous technical factors which still been in a relatively incipient development regarding the parameterizations, and is providing forecasting with a high level of skill and reliability.

The fact that allowed the advance of the modeling and numerical weather forecasting fields is the rapid development and evolution of telecommunications and information technology fields. This is associated, as a consequence of its assessment, with the increasing computing power and reduced costs, which has been facilitating its acquisitions by academic and personal environments. However, despite technical and scientific advances, as well as the capacity of processing, numerical models processing is still restricted to being processed with relatively low space resolution in a global scale level. Even though, it is providing conditions for several universities and research centers to routinely process their own regional mesoscale numerical models, with very high space resolution (e.g. 18km, 6km and 2km used by UFPEL) compared with global forecasting models (54km).

The numerical weather prediction models, on a regional level, demands meteorological observational data with high space and time resolution, in order to forecast the mesoscale phenomena which cannot be captured by the coarser grid resolutions of the large scale global models. These meteorological data were, until recently, practically monopolized by data acquisition centers, usually governmental institutions, which have been justifying such fact, not only due to the economic value of data, but also due to geopolitical questions. Although there are still some restrictions from some institutions, several data are presently available for public domain but not for commercial purpose.

According to Almeida et al. (2005), to facilitate the data access, for the development of atmospheric research, particularly for modeling and numerical

weather prediction, the American universities established in the 80's decade a program called "University Data Project" (Unidata) – in the University Corporation for Atmospheric Research (UCAR).

The Unidata program (<http://www.unidata.ucar.edu/>) has fastly developed. Its mission is: to provide data services, tools and cybernetic infrastructure, acting as a leader in the advance of Earth system's science and in the creation of educational opportunities, as well as in the increase of its user's participation (Yoksas 2006). Presently, it has more than 25,000 users registered from about 150 countries; comprising 1,500 academic institutions and 7,000 organizations. Within this program it was implemented, the project: Internet Data Distribution System (IDD), which evolves the Local Data Management (LDM) software, to enable the network data distribution. As an integrate member, participating along with a great number of research institutions and universities in the UNIDATA program, the Federal University of Pelotas (RS-Brazil) installed an automated integrated data acquisition and processing system. The integrated system has been developed, among others, to use the forecasts of the GFS/NCEP, global forecasting model data, in the regional MM5 mesoscale processing, to fulfill the primary requirements of initial and lateral boundary conditions, using the facilities of the IDD and LDM. This system, besides being used to contribute to the improvement of the skill of the model, as well as providing the weather forecast for the southern region of the country, is used to investigate the benefits of the collective combination - models to predict mesoscale events, and the evaluation of the performance of the integrated system with data acquisition. The products of this integrated system are currently available on the Internet network (<http://200.132.99.3/mm5/yy.html>).

## 2. IDD AND LDM SYSTEM

The Internet Data Distribution (IDD) is a dissemination of meteorological data system, almost in real time – which depends only on communication transfer time and their availability for community users - throughout Internet. It is coordinated by Unidata Program Center (UPC),

which is the focal point of reference, for a community of more 160 universities and meteorological centers of many countries. The UPC belongs to the same structure of the National Center for Atmospheric Research (NCAR), both subjected to UCAR.

Unlike other systems, where data is accessed from a centralized point, the IDD is designed in a such a way that, data requested by a university/user is sent to their computer(s) as soon as it becomes available, that is, whenever it is collected by the observational system (Global Telecommunication System –GTS). Meteorological data in the IDD system is inserted into a computer source-node and sent through the LDM, by the key software that allows the access through the internet to data distribution system.

Each receiving node of the IDD can use the data locally, and eventually, acts as another network node sending it to other nodes. Thus, the system is not overloaded and allows the distribution to all participants in the network. Due to its closed system nature, it is now a network which sometimes have been referred as Internet-2

The topology of the network connection of the IDD/LDM system of July of 2009 (<http://www.unidata.ucar.edu/software/idd/rtstats>) is presented in Figure 1.

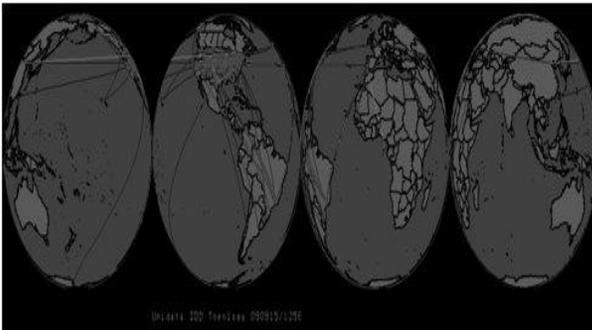


Figure 1: Topology IDD/LDM

The IDD presents numerous benefits and features, inherent to its own system: a) the load distribution allows the decentralization of IDD, avoiding the excessive concentration of data traffic on each node of the network system; so that the large volumes can be distributed without prejudice to any participant; b) a low-cost

implementation, because the required software are freely available open sources; c) permanent support to participants, which also have them from the UPC and the entire community of users; d) remote monitoring, with a set of tools which allow the monitoring, in real time, of all system's network nodes conditions. According to Almeida et al. (2004), a great aspect of the IDD is to allow for the ongoing monitoring, on the Internet, the operation of the system as a whole, or individually, in real-time with the statistics which each LDM can send to remote computers. This facility allows the verification of failures in a given node, allowing, when necessary, to re-configure routes or alternate servers, to redirect the data flow. The various types and sources of data available on the IDD system include:

- conventional meteorological data: collected by synoptic weather stations, buoys, ships, aircrafts and radio-soundings; which are in the global telecommunication system (GTS);

- numerical models data: results from a variety of weather forecasting numerical models, among others, the NCEP, ECMWF, UK, CPTEC, etc.;

- satellite images: several satellites and spectral channels, in particular the GOES series of geostationary and sun-synchronous NOAA series.

A complete list of available data source and available type, with their descriptions, is presented in the

<http://www.unidata.ucar.edu/software/ldm/ldmcurrent/basics/feedtypes/> site. In the system installed in the UFPEL, two reception nodes are included: [yy.ufpel.edu.br](http://yy.ufpel.edu.br) and the [ppn101.ufpel.edu.br](http://ppn101.ufpel.edu.br) – both, properly configured, for receiving data of interest, particularly those related to forecasting operations and research purposes. Following, without a complete description, some of the data type routinely received is shown.

The volume of EXP data type received between 00:00 UTC on 13<sup>th</sup> to 21:20 UTC on July 15<sup>th</sup>, 2009, at [yy.ufpel.edu.br](http://yy.ufpel.edu.br) node from the node CPTEC / INPE, is presented in Figure 2. The bars correspond to three computers, INPE / CPTEC, which sent the data. The highest throughout values per hour, indicate a volume flux around 300 Mbytes/hour, close to the times at which the conventional meteorological data have been collected and transmitted (6 in 6 hours).

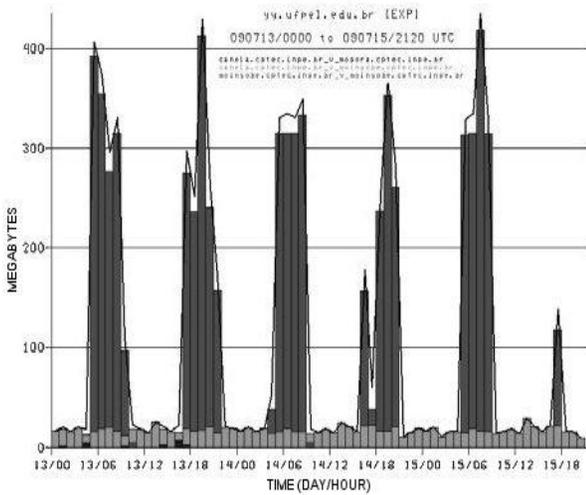


Figure 2: Volume of EXP data received from July, 13th 00:00UTC to 15th 21:20UTC, 2009.

The volume of the conventional HDS (High resolution Data Service) data type, received by yy.ufpel.edu.br node, from July, 13<sup>th</sup> 00:00UTC to 15<sup>th</sup> 21:20UTC, 2009, is presented in Figure 3. This data type comprises the binary products, containing centrally generated analysis and forecast fields in binary GRIB formats. The maximum value of shaded lines indicates the total volume flux received, in Mbytes per hour, from all computers of CPTEC / INPE which sent data, indicated by several black and white shaded bars. Each gray tone refers to one computer of the node which sent the data type.

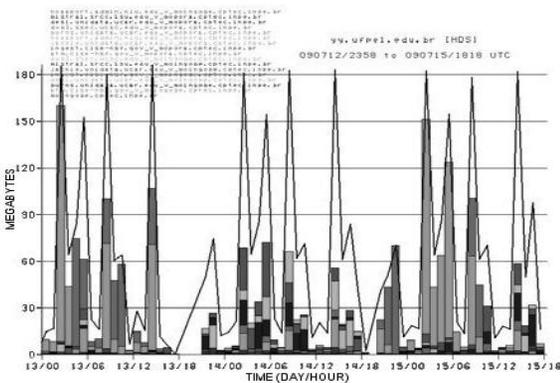


Figure 3: Volume flux of HDS data type received from July, 13<sup>th</sup> 00:00UTC to 15<sup>th</sup> 21:20UTC, 2009.

The IDS/DDPLUS data flux comprises the text bulletins, containing weather information or observations from outside of the United States. The volume of conventional data of IDS|DDPLUS type (also known as NOAAport text products), received by the node yy.ufpel.edu.br, from July, 13<sup>th</sup> 00:00UTC to 15<sup>th</sup> 21:20UTC, 2009 is shown in Figure 4. The upper line indicates the total volume of data received showing volumes around 30 Mbytes per hour. The gray tones have the same meaning as before.

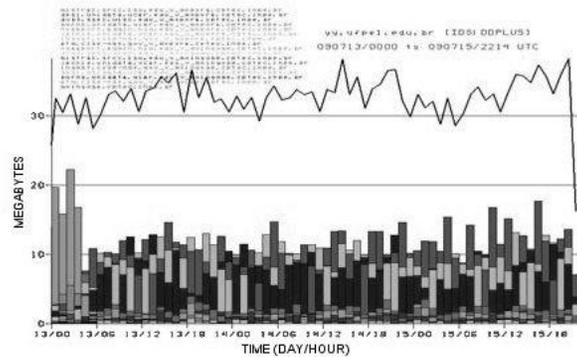


Figure 4: Volume of IDS/DDPLUS data received from July, 13<sup>th</sup> 00:00UTC to 15<sup>th</sup> 21:20UTC, 2009.

Figure 5 shows the total flux volumes of EXP, HDS and IDS|DDPLUS data type received from July 14<sup>th</sup> 00:00UTC to 15<sup>th</sup> 18:00UTC, 2009 in the yy.ufpel.edu.br node. The shaded bars indicate the corresponding three types of products received.

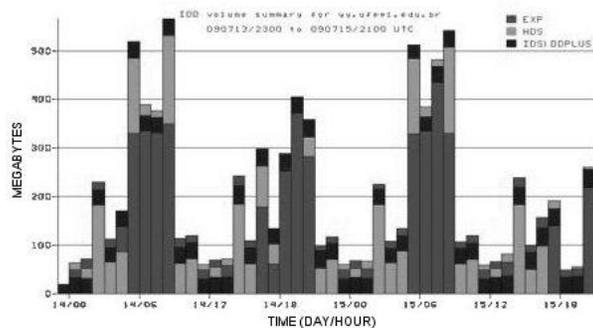


Figure 5: Total volume received in the yy.ufpel.edu.br node from 00:00UTC of July 14<sup>th</sup> to 18:00UTC of July 15<sup>th</sup>, 2009.

The highest volume flux, which can be clearly seen, is related to EXP data, which corresponds to forecasting data of numerical weather forecasting global models, excluding the GFS/NCEP which presents much more and the highest data volume flux.

The statistics made over one node of the UFPEL, presented a maximum value of data flux, received from July, 13<sup>th</sup> 00:00UTC to 15<sup>th</sup> 21:20UTC, 2009, of 565,806 Mbytes/hour - with an average of 194,081 Mbytes/hour and a total of 38,413 products/hour. The highest data volume, which has been daily received in the other node of the UFPEL, is due to the global forecasting numerical model GFS of NCEP products. The data are transmitted in the conduit flux data type and is received by ppn101.ufpel.edu.br node. All forecasting produced by GFS model has been received, with space resolution of 0.5 degrees, in grib2 format, every 6 hours. The total volume received every 6 hours is around 3 Gbytes; in other words, approximately 12Gbytes per day, as shown in Figure 6.

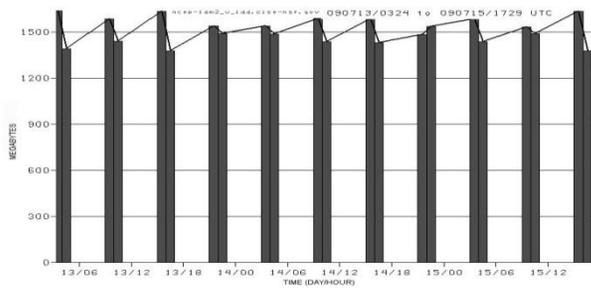


Figure 6: Conduit data flux received from 00:00UTC of July 13<sup>th</sup> and 12:00UTC of July 15<sup>th</sup>, 2009

Figure 7 shows the latency of the conduit data flux reception, by ppn101.ufpel.edu.br, from July 13<sup>th</sup> 03:24UTC to 15<sup>th</sup> 17:01UTC, 2009. According to the monitoring system, during the period between 12:00UTC and 18:00UTC, the highest latencies and the most critical periods of the day occur. These latencies are related to the time period on which the UFPEL communication system presents the highest number of users. However, even with these high latencies, during some periods of the day, it lasts nearly 1800 seconds to get all the

requested data, with a delay of 30 minutes approximately, considering the whole period depicted in Figure 7. These facts are not too critical, at least for the operations and developments which have been conducted with these data flux in the UFPEL.

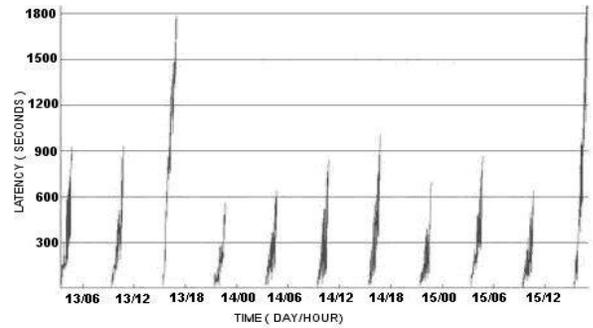


Figure 7: Conduit data flux latency reception from July 13<sup>th</sup> 03:24UTC to July 15<sup>th</sup> 17:01UTC, 2009

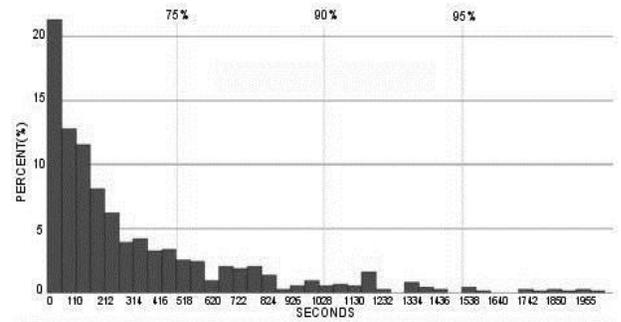


Figure 8: Conduit data reception rate from July 13<sup>th</sup> 03:24UTC to July 15<sup>th</sup> 15:00UTC, 2009

The values, as shown on the top of vertical dashed lines in Figure8 represent the total of conduit flux data percentages being received. The horizontal axis, in seconds, exhibits the time that lasted the requested data reception. Therefore, as shown, 75% of them were received in less than 520 seconds, and 90% in less than 1200 seconds. The statistic of the conduit volume data rate received, by ppn101.ufpel.edu.br node, indicates that it has a maximum volume of reception of 1633,936 Mbytes/hour, with an average of 1511,026 Mbytes/hour.

The UPC also provides analysis and visualization software, as GEMPAK, IDV, McIDAS and NetCDF. All of them are extremely useful for several operational and research applications. They are, with the LDM and their decoders, an integrated system with the data acquisition, decoding, analysis and visualization –allowing user’s to generate their own graphics and figures, as well as immediate display for their applications.

### 3. MM5 MODEL AND THE BASIC CONFIGURATION

The mesoscale modeling system MM5 (Dudhia et al. 2005) is a limited area, non-hydrostatic model, developed in sigma coordinates to simulate or predict the atmospheric mesoscale circulation. The Figure 9 presents the modules of the MM5 system.

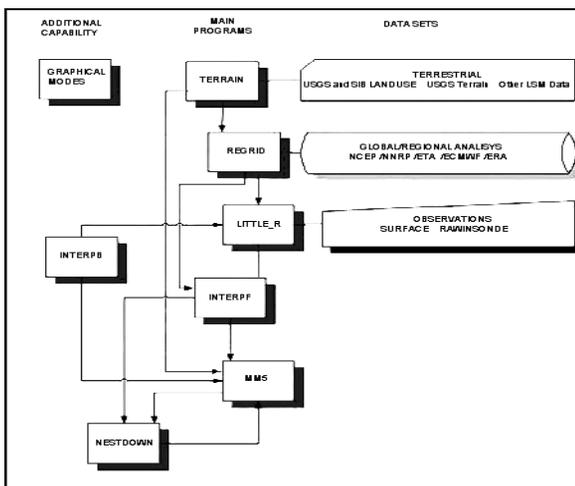


Figure 9: Components of the MM5 modeling system

The integrated system implementation - namely the MM5 mesoscale forecasting model and LDM (conduit data flux) have been made using the forecasting of the global model GFS/ NCEP data. The data has been obtained in GRIB2 format, with space resolution of 0.5 degrees. It enables the processing of the MM5 system, starting with the MM5 system REGRID module (Figure 9). As the LDM also presents, among others, the synoptic and rawinsonde data, they can also be used in the

Little\_R module, using the 3DVAR assimilation technique, to process the MM5.

The integrated system has been implemented configuring the MM5 with 3 nested domains, D1, D2 and D3, employing two way integration procedure. The area coverage of each domain is shown in Figure 10. The model has been configured with 23 vertical sigma layers, with higher resolution on lower levels. The top of the model atmosphere has been delimited to 100 hPa.

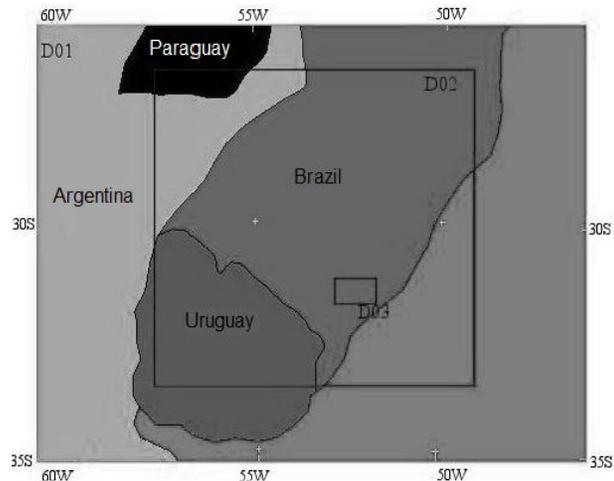


Figure 10: MM5 Domains of integration

Table I shows the horizontal space resolution, pre-established, for each domain of integration; the respective numbers of grid points along the east-west and north-south directions and the time step used in the model integration.

The parameterizations used were established according to Yamasaki and Orgaz (2003). No cumulus parameterization is employed in domains D2 and D3, since the model explicitly treats the convective processes for these space resolutions.

Some of MM5 output, as well as ECMWF global model forecasting products, which have daily been generated and or received in the integrated system will next be presented. The MM5 products will be shown only as an example, for the domain D1. It will not be discussed any weather forecasting analysis of the atmospheric condition, which is beyond the present scope.

MM5 Domain	D1	D2	D3
Horizontal resolution (km)	18	6	2
East-West grids	80	139	52
North-South grids	60	130	31
Time step (s)	54	16	5.3333

Table I: MM5 configurations

#### 4. GEMPAK SYSTEM

The GEMPAK (The **G**eneral **M**eteorology **P**acKage) is a system which generates, for analysis and visualization of meteorological data products. When integrated with an IDD system, besides processing all received data, decoding and providing an efficient data storage, it has been used by several operational weather forecasting centers (Fulker 1997). Some products which have been received in UFPEL are presented below, without being committed to all details of this powerful system for products generation. The technical details on the data collection, processing, as well as installation procedure are presented in <http://www.unidata.ucar.edu/software/gempak>.

Figure 11 shows the CPTEC\_ETA forecast of the accumulated rainfall, in 6 hours, for July 17<sup>th</sup> for 12:00UTC, 2009. The gray shaded bar on the right side indicates the intensity of precipitation in mm. It is overlapped by (channel 4) infrared geostationary GOES-10 satellite image as well as with the METAR data.

Figure 12 shows the ECMWF (European Centre for Medium-Range Weather Forecasts) model forecast for the 850 hPa temperature field.

The UKMET model forecasting for the mean sea level pressure (hPa), overlapped with channel 4 infrared GOES-10 satellite image, is presented in Figure 13.

The spatial coverage and distribution of the Upper Level Weather Stations (EMA), with the wind vector direction and intensity for 500hPa are presented in Figure 14.

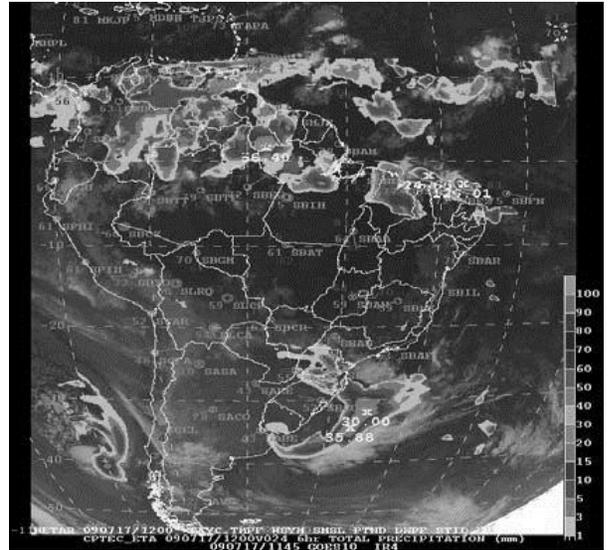


Figure 11: METAR data, 24h total accumulated precipitation of CPTEC\_ETA forecast (mm) for 170709:12UTC and the IR-4 GOES-10 satellite image

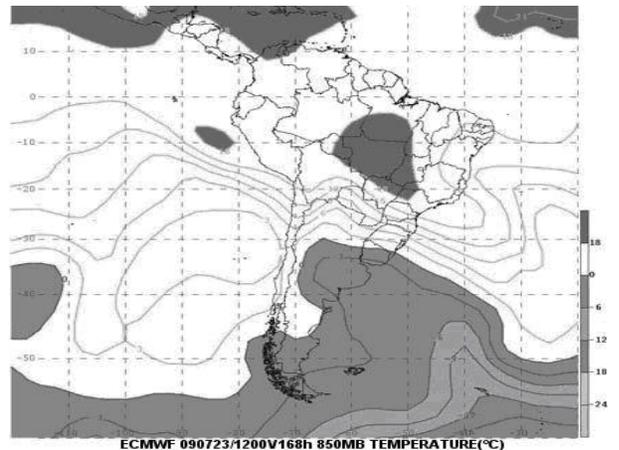


Figure 12: ECMWF model forecasting of temperature (°C) at 850hPa for 12:00 UTC of July, 23<sup>rd</sup>, 2009.

### 5. THE INTEGRATED LDM/MM5 SYSTEM RESULTS FOR RIO GRANDE DO SUL STATE (RS) - BRAZIL

The numerical model forecasting of the integrated LDM/MM5 system for Rio Grande do Sul State (RS) - Brazil is presented only, as an example, for the D1 integration domain.

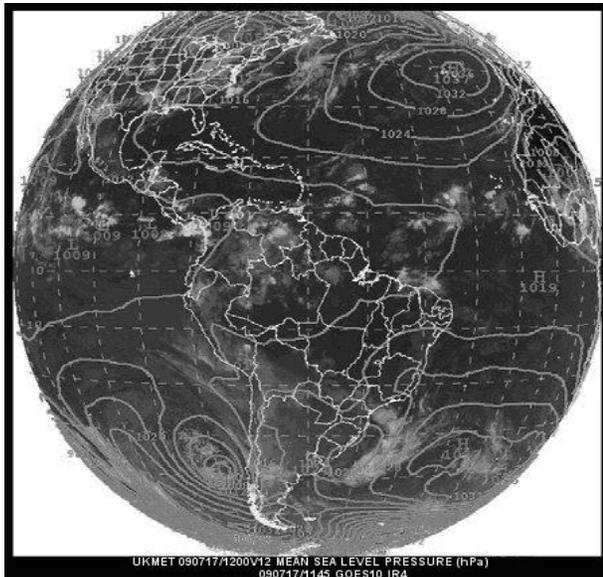


Figure 13: UKMET model forecasting of the mean sea level pressure (hPa), for July 17<sup>th</sup> 12:00UTC, 2009, over the IR-4 GOES-10 satellite image of 17<sup>th</sup> 12:00 UTC, 2009



Figure 15: Major cities with MM5 forecasting

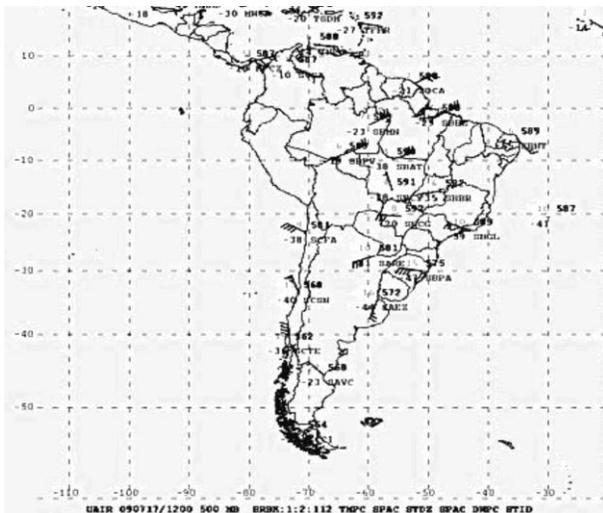


Figure 14: EMA stations coverage map and wind data for July, 17<sup>th</sup> 12:00UTC, 2009

The model has been integrated for a 5 day period, and hourly forecasting results have been archived. The most commonly used meteorological variables to weather forecasting operations are plotted in graphical forms. One of these forms is presented, as a forecasting product for Rio Grande do Sul State major cities- located on the map shown in Figure 15 – and available on web page <http://200.132.99.3/mm5/yy.html>.

The forecasting in this particular form, presents the MM5 forecasting for the selected city - providing the hourly forecast of: the temperature and rainfall as shown in Figure 16; pressure and precipitable water as in Figure 17 and the wind and clouds as shown in Figure 18 (Pelotas City, respectively). This particular form has been made to attend the demand of the general public and, since the plotting form is self-explained, it is very easy to make the interpretation of the forecasting by itself.

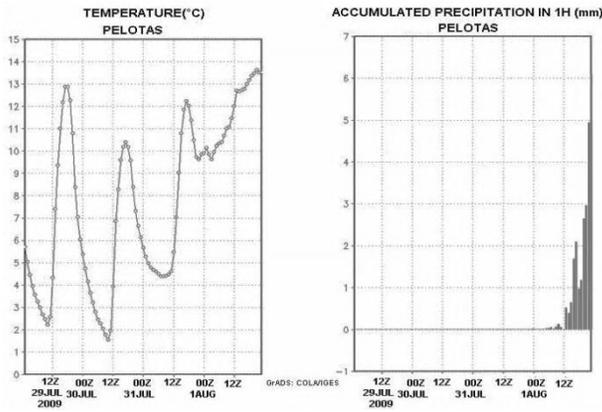


Figure 16: Temperature (°C) and one hour accumulated precipitation (mm) forecasting for 5 day period.

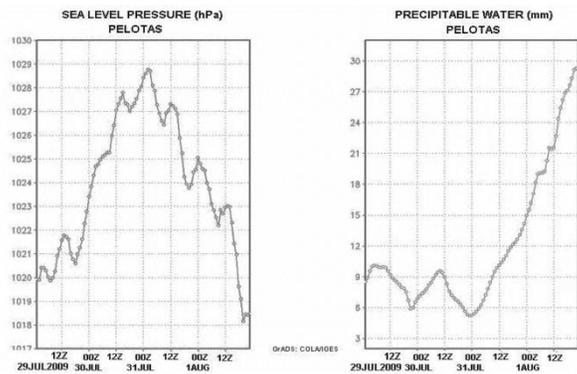


Figure 17: Pressure (hPa) and precipitable water (mm) forecasting for 5 day period

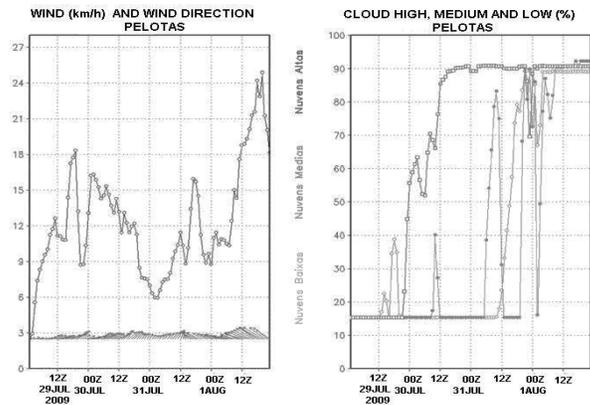


Figure 18: The wind (km/h) and cloud coverage (%) forecasting for 5 day period

Another forecasting presentation is in a loop form. The hourly sequential images, with up to 48h of prognostic fields, are presented as in Figure 19, for the 2m temperature field, predicted for July 30<sup>th</sup> 07:00UTC, 2009, with model analysis started on July 29<sup>th</sup> 00:00UTC, 2009. This module allows, therefore, the sequential visualization of different meteorological fields, with hourly forecasting, over the whole area coverage of the MM5 model integration domain.

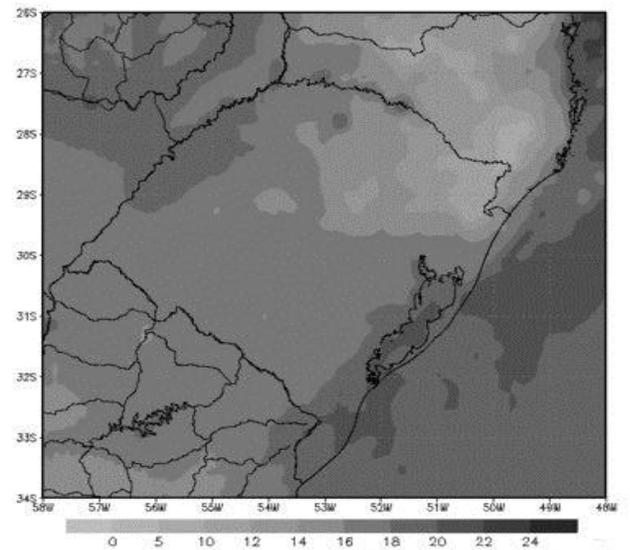


Figure 19: The 2m level temperature (°C) for the MM5 D1 domain

## 6. CONCLUSION

The ever increasing advent of new methodologies, software, graphic display, formatting standardization and data processing is a present reality, which has been going along with the development of communication and computing - particularly those of "desktop" type. Despite the fast development, the fact is that numerical modeling of the atmosphere, which had been during long time, limited to the numerous physical parameterizations developed for processing of global atmospheric models with relatively low space resolution - both horizontally and vertically -, are still being used in several mesoscale models processing, particularly for

limited areas and grids with a space resolution bigger than about 10 km. Thus, many studies, which were developed in the past for mesoscale models and applied to spatial resolutions above 10 km, have been explored due to the access to new technology capacities and facilities available.

Thus, it is important to implement an efficient system not only for processing, but also for data acquisition, model processing as well as the tools for diagnostic and prognostic tests as an integrated system.

As a result of new developments it was implemented an integrated IDD / LDM / GEMPAK with the MM5 mesoscale model systems. This integrated modeling system implemented since mid-2007 has been processing data 24 hours a day. It has revealed that all modules operate in a consistent way with the needs of both - research and developments, also meeting the needs for operational analysis. The huge volume of data available within the IDD/LDM system - parts of which are of private interests - which are obtained on a daily basis, has proven the efficiency of data gathering. In addition, the system integration with display systems, including one of them, presented here - the GEMPAK - clearly shows the ease with which it is possible to have at hand any product that meets the needs for both research and operation.

One of the important advantages of integrating the system with a display, as GEMPAK, is that it allows the quick detection of any faults in numerical forecasts - and therefore helps as a diagnostic tool also studying involving developments to produce better forecasts of numerical models, as shown for some special cases of the MM5 prognostic model.

Acknowledgments: To all institutions which provided the essential data and supports for carrying out the research: NCEP, UCAR; INPE; UFPEL

## REFERENCES

- Almeida W. G., Carvalho L. A., Coelho D. G., Cutrim E., Oliveira I. C., Silva M. G. A. J., Ferreira S. H. S., Yoksas T., Spangler T., 2004. Testes no Brasil com o Sistema de Distribuição de dados meteorológicos pela internet (IDD). XIII Congresso Brasileiro de Meteorologia, Fortaleza. Online at [www.cbmet.com/edicoes.php?pageNum\\_Records\\_et\\_busca=12&totalRows\\_Recordset\\_busca=694&cgid=22](http://www.cbmet.com/edicoes.php?pageNum_Records_et_busca=12&totalRows_Recordset_busca=694&cgid=22)
- Almeida W. G., Carvalho L. A., Ferreira S. H. S., Coelho D. G., Justi M. G., Cutrim E. and Yoksas T., 2005. IDD-Brasil: Distribuição de dados meteorológicos para ensino e pesquisa. Boletim da Sociedade Brasileira de Meteorologia, 29(2), 33-38
- Dudhia J., Gil D., Kuo Y. R., Burgeois A., Wang, W., Bruyere C., Wilson J. and Kelly S., 2005. PSU/NCAR Mesoscale Modeling System. MM5 Modeling System V3. NCAR TN, p.1-3. Online at [www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes.html](http://www.mmm.ucar.edu/mm5/documents/tutorial-v3-notes.html)
- Fulker D., Bates S. and Jacobs C., 1997. Unidata: A Virtual Community Sharing Resources via Technological Infrastructure. Bull. Amer. Meteor. Soc., 78, 457-468
- Yamasaki, Y. and Orgaz M. L. M., 2003. A MM5 Extreme Precipitation Event Forecast Over Portugal. 13th. PSU/NCAR Mesoscale Model User's Workshop, NCAR, Boulder, p. 85-89.
- Yoksas T., Emmerson S., Chiswell S., Schmidt M. and Stokes J., 2006. The Unidata Internet Data Distribution (IDD) System: A Decade of Development. Preprints, 22<sup>nd</sup> International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology, American Meteorological Society, January. Online at [www.unidata.ucar.edu/publications/uni.bibliosave.html](http://www.unidata.ucar.edu/publications/uni.bibliosave.html)