Domain Knowledge Assisted Data Processing for Florida Public Hurricane Loss Model

Yilin Yan*, Samira Pouyanfar†, Haiman Tian†, Sheng Guan†, Hsin-Yu Ha†,
Shu-Ching Chen†, Mei-Ling Shyu*, and Shahid Hamid‡
*Department of Electrical and Computer Engineering
University of Miami, Coral Gables, FL, USA
†School of Computing and Information Sciences
‡Department of Finance
Florida International University, Miami, FL, USA
*y.yan4@umiami.edu, {spouy001, htian005, sguan005, hha001, chens}@cs.fiu.edu,
shyu@miami.edu, hamids@fiu.edu*

***Abstract*—Catastrophes have caused tremendous damages in human history and triggered record high post-disaster relief from the governments. The research of catastrophic modeling can help estimate the effects of natural disasters like hurricanes, floods, surges, and earthquakes. In every Atlantic hurricane season, the state of Florida in the United States has the potential to suffer economic and human losses from hurricanes. The Florida Public Hurricane Loss Model (FPHLM), funded by the Florida Office of Insurance Regulation, has assisted Florida and the residential insurance industry for more than a decade. How to process big data for historical hurricanes and insurance companies remains a challenging research topic for cat models. In this paper, the FPHLMs novel integrated domain knowledge assisted big data processing system is introduced and its effectiveness of data processing error prevention is presented.**

***Index Terms*—Big Data Processing; Catastrophe Modeling; Pre-processing; Post-processing; Florida Public Hurricane Loss Model (FPHLM)**

I. Introduction

A catastrophe is a sudden and widespread disaster that causes tremendous damages. Among all kinds of catastrophes, tropical cyclones are the most deadly hazards threatening the state of Florida. Based on different wind speeds located around the circulation center, they can be classified into different types including tropical depressions, tropical storms, and hurricanes. Hurricanes, also known as typhoons in the northwest Pacific Ocean, are the most powerful tropical storms. The Labor Day hurricane in 1935 was the strongest hurricane with 1-minute sustained 185 mph to make landfall in the United States recorded history. It killed 408 people and severely destroyed parts of the Florida east coast railway. The most intense hurricane in recent decades, Andrew in 1992, caused about $26 billion damage in South Florida and destroyed over 25-thousand houses in Miami-Dade County, while nearly 100-thousand more houses were severely damaged. Just ten years ago in 2005, Hurricane Katrina caused around $108 billion property damages in the United States, which is roughly four times the damage wrought by Andrew [1].

The tremendous loss is a strong incentive to build hurricane catastrophe models for loss estimation. Some groups put their efforts into single hurricane case studies. In [2], the authors present their work on visualization of scientific data from Hurricane Katrina, originating from computational simulations. With remotely sensed observations including the 3-D terrains, their model simulates the development of Katrina in the atmosphere interacting with the ocean which caused deadly storm surge. For the same hurricane, local researchers [3] estimated the effects of Katrina on the Louisiana coastlines. As a result of these studies there has been increased attention from the government to develop policies to mitigate extensive damages by hurricanes.

To understand hurricane risk and estimate losses, cat models are essential. Researchers in both developing [4] and developed [5] countries have tried to build general estimation models. The Florida Public Hurricane Loss Model (FPHLM) [6] is the first and only public wind hurricane loss projection model in the United States. With a multi-disciplinary team of experts in the hurricane catastrophe modeling fields, FPHLM was developed as an advanced, automated, and systematic model to help the State of Florida [7]–[13].

Supported by Florida Office of Insurance Regulation (FL OIR), the FPHLM project aims to help the residential insurance industry with the rate-making process and evaluate solvency of insurance companies. FPHLM is considered novel in the following three ways [14], [15]: (1) It is a systematic model with different functional components, including pre-processing, Wind Speed Correction (WSC), Insurance Loss Model (ILM), post-processing, verification, and documentation; (2) With multiple attributes frequency distribution tables, it also includes features for data analysis, model configuration and event notification by proper permission management using access control tables based on the fine-grained system actions; and (3) Experts from different fields, including meteorology, engineering, actuarial science, GIS, statistics, finance, and computer science, collaborate together. The current wind model project includes over 30 researchers and students from Florida International University, University of Miami, University of Florida, Florida State University, Florida Institute of Technology, National Oceanic and Atmospheric Administration (NOAA) Hurricane Division, and AMI Risk Consultants.

As in a big data era, a large amount of hurricane catastrophe data is collected by remote sensors and provided by insurance companies. Such large collections of data can create lots of opportunities but at the same time pose great challenges and difficulties in data processing in order to mine and retrieve useful information efficiently. For the FPHLM project, how to automatically process the big data from the meteorology group, the engineering group, and the insurance companies is one of the major research problems. About 57-thousand years of wind simulation data is the largest dataset in the meteorological component, and a single dataset from an insurance company can contain data on ten thousands of insurance policies. There are about 29 million grid points mapped in the state. As a result, for each geographic location in Florida, it takes an extremely long time for data processing. Besides, datasets from different property assessments can also be in different formats and suffer from data-entry errors. Therefore, in this paper, we focus on how to process the deluge of hurricane catastrophe data efficiently and correctly.

The rest of this paper is organized as follows. In section 2, previous work on catastrophe modeling including those on other hurricane models and corresponding data processing steps are introduced. Section 3 gives an overall view of our FPHLM design and discusses the details of our system components including pre-processing, wind speed correction, insurance loss model, post-processing, and verification. Section 4 shows the implementation of our system. The last section presents the conclusion and future work.

II. Related Work on Catastrophe Modeling

Catastrophe modeling is one kind of disaster data modeling which adopts information technology to estimate the losses caused by a catastrophic event. It has attracted considerable attentions from the governments, industries, and research groups due to the huge damages caused by the catastrophes such as floods, earthquakes, and hurricanes. In this section, we first introduce different kinds of hurricane catastrophe modeling, and then their pre-processing and post-processing tools. Also, a literature review of catastrophe modeling for the insurance industry is presented since it is very suitable for risk analysis.

The Cat models typically involve the efforts of experts from multiple disciplines, making domain knowledge assisted modeling and data processing a necessity. In [16], a complete hurricane catastrophe model is divided into four different modules including a stochastic model that randomly generates a hazard model for local severity (peak wind gust), a vulnerability model that estimates the loss given the local severity, and a financial model that calculates the flow of money between two parties (industry and economic losses). In [17], HAZUS-MH, a hurricane model methodology, incorporates experts from multiple disciplines which includes hurricane hazard, wind load, terrain, physical damage and loss components. To date, several HAZUS-MH tools such as Comprehensive Data Management System (CDMS), Inventory Collection Survey Tool (InCAST), Risk Assessment Tool, and Flood Information Tool (FIT) have been developed, which enable the users to manage statewide datasets. However, all these tools require users to have the domain knowledge and also to undertake huge human efforts.

A. Pre-processing and Post-processing

Data pre-processing is a vital component in many catastrophe models. For example, Coastal Louisiana Risk Assessment (CLARA) [18] is an analytical model which evaluates flood damages of storm surges in Lousiana's coastal region. CLARA includes three primary components: (1) The pre-processing module including spatial data and storm data pre-processing; (2) The flood depth module estimating the depths of a flood in the protected areas; and (3) The economic module calculating the damage losses for each flood depth.

The pre-processing is responsible for getting the trimmed model input, and post-processing serves to visualize and interpret the results. Viviroli et al. [19] list three major challenges to illustrate adequate necessity of their pre-processing and post-processing tools: (1) time-consuming, (2) the need to compile, process, and interpolate data series from the meteorological station networks, and (3) extra software need to interpret model outputs. Tasks in their proposed framework typically involved with pre-processing and post-processing are handled by adopting the tailored tools with Graphical User Interface (GUI) to speed up the data processing. In [20], the authors mention how they select the observation data from the global integrated public dataset in the pre-processing and how to analyze the model bias in the validation and post-processing part. In [21], for the post-processing part, the authors emphasize the model output statistical analysis together with screening regression procedure (which can select the predictors to be included in an equation), are capable of estimating the surface wind, the probability of precipitation, maximum temperature, cloud amount, and the conditional probability of frozen precipitation. Analyzing big data in pre-processing and post-processing is extremely important and leads to more accurate hurricane forecasting. A parallel aggregate risk analysis algorithm is employed in [22], which utilizes large-scale input data and improves the speed of the aggregation process. The algorithm and its aggregate risk engine, implemented on both multi-core CPU and GPU platforms, generate the aggregate analysis in an efficient and real-time manner. However, more techniques need to be included for efficient data processing.

B. Catastrophe Risk Modeling for Insurance Industry

Catastrophe modeling is especially applicable to risk analysis in the insurance industry and is at the confluence of actuarial science, engineering, meteorology, and computer science. Various private loss modeling systems have been developed to assist the insurance industry in the insurance rate process.

Recently, research studies of catastrophic insurance have attracted attentions from many governments as the loss caused by catastrophes is becoming more and more serious. The 2008 Wenchuan earthquake, the 21st deadliest earthquake, alone with a severe snow storm rarely seen before, cost no less than 1 trillion China Yuan, or $150 billion USD over the next three years to rebuild the ravaged areas.

For the aforementioned Wenchuan earthquake, experts of the China Insurance Regulatory Commission (CIRC) estimated that the actual compensation was no more than 5% of the losses, which is far below the global average catastrophe insurance compensation level of 30% and 60%-70% in developed countries [23]. As catastrophes result in very high damages, developing the catastrophe insurance model is a necessary and difficult task.

Currently, catastrophe risk models have been extensively used for risk and loss analysis, and can be further leveraged to conduct stress tests on insurance companies and make strategic plans. Paper [24] presents various catastrophe modeling technologies which offer significant values to the insurance companies. One such model is used for water bloom prediction in [25], where data pre-processing techniques are used to unify the data dimension and fit the data into the model. However, the pre-processing is limited to the historical data statistics such as yearly mean value.

III. FPHLM DATA PROCESSING SYSTEM COMPONENTS

Since FPHLM initially released in 2006, it has been used for over 1000 occasions to evaluate rate change requests from the insurance companies. The FPHLM data processing framework is a novel, automated, and modern system integrating the following five components:

(1) Pre-processing
(2) Wind Speed Correction
(3) Insurance Loss Model
(4) Post-processing
(5) Verification

Figure 1 shows an overall view of these components, their interactions, and the groups involved in each component.

A. *Pre-processing*

As discussed earlier, the pre-processing part for FPHLM is a very important component and faces a plethora of challenges. It receives the input data from both FL OIR and private insurance companies, most likely in different formats. In addition, the data may include up to millions of policies with a large number of attributes and in most cases come with missing or erroneous values. Here is a list of common problems that FPHLM faces:

- Missing zip code
- Zip code provided does not match corresponding physical address
- Address for a particular policy is not from Florida
- Missing geographic coordinates
- Invalid city or county name
- Value shift

It takes a lot of efforts to take care of the potential issues with the zip codes, for instance, changing the zip code format, removing the last 4 bits of the zip codes, and updating the invalid ones. If the zip code fields provided by the insurance companies have missing values or just simply cannot be matched with the actual physical address, the FPHLM model is able to automatically update these attributes without manual efforts. Furthermore, in FPHLM, a preliminary module run is performed after pre-processing in order to identify any inconsistency between attributes, e.g., zip code and county. That is, all inconsistencies are automatically resolved before the vulnerability calculation in FPHLM; while the unmapped zip codes are either manually revised or are omitted from the pre-processing in the other models [26].

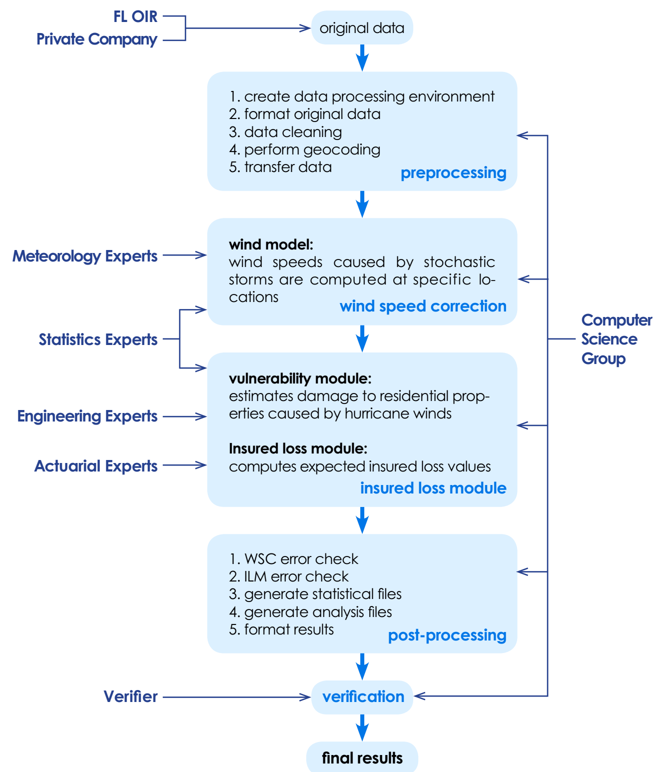

Fig. 1. Overview of FPHLM system.

For the policies with missing zip codes and/or geographic coordinates, we perform geocoding using a third-party software named ArcMAP [27] which is the main component of Esri's ArcGIS suite of geospatial processing programs to produce geographic coordinates needed in the next component in Section III-B. The processor first exports the addresses from the policies, and then feeds them into ArcMAP for geocoding. Next, the results are placed in the corresponding folder. Before continuing with pre-processing, the results are checked, formatted, and then imported back to the policies. In fact, in many cases, the raw insurance policies do not include the coordinates, which makes the geocoding sub-component necessary. If an address could not be matched by ArcMAP, the centroid of the provided county is used for further processing.

The experts in the engineering team determine the external vulnerability of the structures using different combinations of policy attributes including year built, number of stories, and various mitigation properties to generate the vulnerability matrices used as the input in Section III-C. However, it is impossible to provide the vulnerability matrices for every combination of attributes since the number of insurance policies in most companies is big. Thus, the engineering team provides a base set of matrices which cover the basic categories of combinations. The pre-processing component needs to map each value to a valid one for the vulnerability matrices in the base set. For example, the roof shape has to be one of the following types: gable-unbraced, gable-braced, gable, hip, other, and unknown. If the provided roof shape or its code is invalid, the system will try to map it to a valid value or set it to unknown. As another instance, when there are more than one county referred to the same zip code and only one of them is used to denote the corresponding vulnerability matrices, then linking a policy containing a different county name with the correct matrix may raise a problem. Thus, the county names also need to be revised.

Other fields such as losses, deductibles, and construction types also need to be cleaned and corrected. A summary file including the basic statistics of all attributes and potential erroneous values is generated automatically both before and after pre-processing. After this component is done, the summary file is formatted and stored for our record and references. The results are also used to further generate the region and year built distribution files. If a problem is ambiguous and cannot be automatically fixed, the pre-processing component can generate a summary report that identifies any major issues which require to contact the corresponding insurance company.

Furthermore, the pre-processing component can automatically check typo errors, change the invalid inputs, and solve the formatting issues in the data, which reduces a lot of manual efforts. The pre-processing component also controls the permission issues and keeps the environment clean and clear. Overall speaking, all the aforementioned steps are integrated into this component.

B. Wind Speed Correction

Derived from the meteorological discipline, the Wind Speed Correction (WSC) component simulates the storm tracks and wind speeds at each policy location with multiple events and refines the wind speed produced by the hurricane wind model on the actual terrain. Typically, tens of thousands of stochastic events that span almost sixty thousand years are simulated in this WSC component. The WSC component initiates their conditions with random historical records that come from the Atlantic tropical cyclone basin. The input of the WSC component consists of the storm information (radius of maximum winds and location of the storm center), roughness length for the open terrain, and the policy data information (policy IDs and their latitude/longitude coordinates), etc. The output of the WSC component consists of the simulation results of the wind speeds for multiple height spots at each policy location. Combined with the exposure data from the insurer policies, the characteristics of the personal or commercial residential properties and the vulnerability matrices serve as the input to the component in Section III-C. The WSC component includes the following sub-components:

(1) Storm Track and Intensity Model
(2) Inland Storm Decay Model
(3) Wind Field Model
(4) Gust Factor Model
(5) Terrain Roughness Model
(6) ArcIMS environment

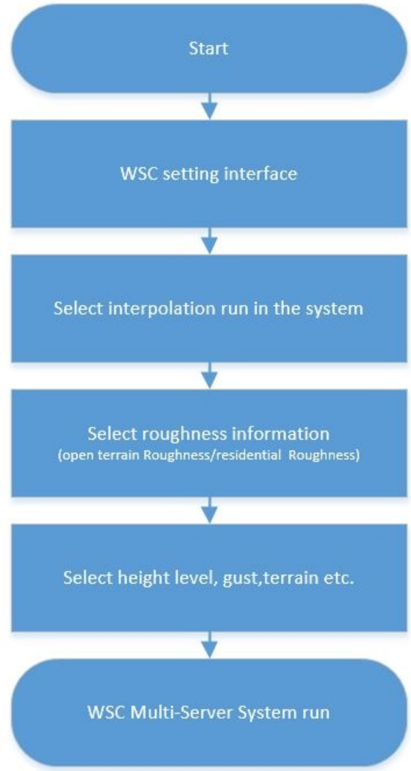

Fig. 2. Flowchart of the WSC component.

To launch one WSC task, we need to first import the interpolation setting in the WSC setting interface and select the existing roughness option of the model. In addition, the remaining information like the height level, gust, task name, and so on need to be set at the same time. When the pre-processing component is finished, the trimmed policy dataset is generated and would be used as the input to the WSC component to estimate the wind speeds on different heights of each policy location (latitude and longitude coordinates). After the WSC task is launched, we use the roughness information and marine surface winds to calculate the terrain-corrected 3-second gust winds and 1-minute peak winds for the actual terrain and open terrain at the street level. The detailed flow diagram of WSC is shown in Figure 2.

The output of the WSC component includes the policy ID and wind speeds for at least ten-meter height level and up to fifteen levels in ten-meter increments if requested. The results together with the exposure information and the

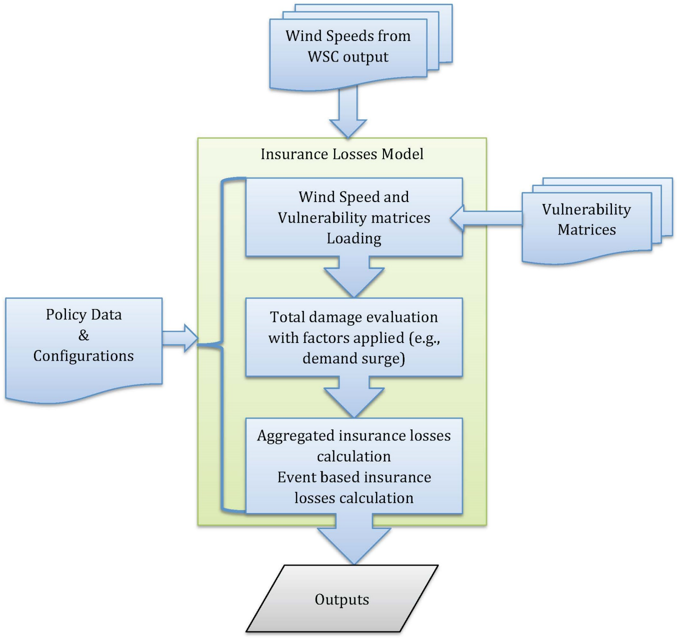

Fig. 3. Flowchart of the ILM component.

engineering vulnerability matrices are further applied to the ILM (Insurance Loss Model) component in Section III-C.

C. Insurance Loss Model

Insurance Loss Model (ILM), which is responsible for predicting the losses of the residential structures, integrates the damage ratio distribution from the engineering component with the wind speed calculation results from the meteorology component. The engineering component, which also be called as the vulnerability component, uses the Monte Carlo simulation to achieve the external damage of the buildings at various wind speeds [28]. The internal damage, including the utilities and content losses, are estimated from the external analysis results formulated as the damage matrices. The vulnerability matrices for each typical building type in Florida, including the manufactured homes, are expanded based on the Monte Carlo model simulations to cover all structural type combinations (frame or masonry), roof cover type (gable vs. hip, tile vs. shingle), region (North, Central, or South), and sub-region (high wind velocity zone, wind-borne debris region, and others). The ILM program automatically decides the matrices usage for each property in order to estimate the total building loss.

The ILM component receives the building stock information from the pre-processing output that determines the set of exposure, zip codes, and construction types, along with the wind speed data which is the result of the previous step for each storm, to perform the simulation and quantify the total damage as the output. Figure 3 gives an overview of the ILM component.

D. Post-processing

As mentioned in the Introduction section, the integrated data processing is vital because of the deluge of insurance data. Also, due to different data sources and requirements, a comprehensive post-processing component is designed by using a series of software, tools, and languages. An effective and automatic post-processing component is very critical to ensure a fast and coherent system. The results from Sections III-B and III-C are first checked by a set of codes to meet the "zero error tolerance" requirement. These codes check the errors and exceptions from the log files and also check the potential problems. If a problem is found, the results are considered unreliable and immediate attention messages are sent to the corresponding data processor.

To generate user-friendly results for FL OIR and insurance companies, raw results from ILM need to be formatted into Excel tables. However, formatting tens of files (from the data analysis) and millions of policies is a tedious and time-consuming job. Furthermore, manually dealing with this kind of big data is error-prone and is definitely not a smart choice. Using the proposed post-processing system, statistical files from the pre-processing component and analysis files from the ILM component are automatically verified and formatted.

Additionally, a data transfer element is designed for insurance data transformation between components. It simplifies the automation process and is flexible to new technologies and model changes. FPHLM also includes a well-designed email notification system and a web-based processing monitoring system. Finally, the results are formatted, reviewed, and packaged by the post-processing component before the submission. While all data processing jobs are associated with organizations, with the web-based processing monitoring system, permission control is as simple as creating a new user account and associating it with the corresponding organization.

As discussed before, our task involves lots of tables and faces the big data issues. Without this post-processing component, a lot of human efforts will be needed for checking and formatting the results and it will take more than two hours for a single company. On the other hand, by integrating a set of scripts together, simply clicking a single button is needed for all the post-processing steps and it takes only two minutes for checking and formatting the final results. By switching the manpower from routine and tedious work to monitoring the overall system, the FPHLM model successfully achieves error prevention - nearly leaving no room for mistakes from data processors.

E. Verification

A minor fault in a technical system may result in costly and critical consequences. Thus, verification is essential to alleviate the unintended faults made by the processor. However, the system complexities make the verification bothersome and difficult. In the proposed system, verification is applied in an automatic manner under the supervision of the verifier. As soon as the data processing is completed by the processor, an email is automatically sent to the verifier to check the correctness of the results and the processing steps. The whole FPHLM verification process is automated and includes the following routines:

- Pre-processing verification: In this step, several prepared SQL functions are executed to check the data inconsistencies. First, the processed insurance portfolio is checked, as well as its number of policies and the order of the data columns. Then, the verifier checks if there is any null value and the pre-processing component has not assigned any value to it. Finally, the geographic information is checked to ensure there is no GIS errors.
- WSC and ILM verification: The main purpose of this step is to verify the model parameters and the values assigned by the processor. In addition, the log files are automatically checked using a shell script that looks for all errors, warnings, exceptions, etc. during the process execution. These kinds of errors may happen due to server problems or parameter inconsistencies. If no fault is found in this step, the verifier runs another script which checks if the policy numbers in WSC and ILM are consistent with the original portfolio. The final script checks the correctness of the soft links, such as WSC, vulnerability matrices, and roughness information, to name a few.
- Results verification: In the final step of the verification, the pre-processing and ILM results are verified. The format and completeness of the ILM results including the analysis files, expected losses, and PML, and pre-processing results such as the summary file, distribution, and exposure files are all checked to avoid any incompatibility.

Running these powerful functions and scripts makes the verification process easy and straightforward. Once the verifier approves the processing and the results, a notification is sent to the approver by the system. At this stage, the approver checks the results to ensure there is no actuarial inconsistencies and logical errors. There are various distribution tables generated by the system to assist the approver for further verification. As soon as the results are validated by the approval, an email notification with a link to the page containing all the result files is automatically sent to the client (usually an insurance company). Thereafter, the client can easily download all the desired files.

IV. System Implementation

As discussed in the previous section, the data flows as follows. The original insurance data is first cleaned and formatted by the pre-processing component. Next, the data-transfer element delivers the processed data to the WSC component as well as the ILM component. The WSC component (Section III-B) generates the wind speeds, and the ILM component (Section III-C) uses the results from WSC together with the engineering vulnerability matrices as its input. Finally, the results after post-processing are verified and submitted. These complicated and computational intensive components from cleaning the data to formatting the final results are now automated by the proposed FPHLM integrated domain knowledge assisted big data processing system.

In addition, the monitoring of this integrated data processing system is web-based and user-friendly. Starting from receiving the data from FL OIR or the insurance companies via secured e-mails or postal services using the encrypted media, an FPHLM task is created by the computer science group and the original data is uploaded to the Data Processing Manager (DPM). When the processing stage changes, the DPM notifies all the corresponding members via secured e-mails. This web-based system is developed using a suite of popular software tools, including JavaScripts (JS) and Cascading Style Sheets (CSS).

The logic tier of FPHLM is written in Java and hosted in an Apache Tomcat server with a cluster of powerful computing nodes running Linux; while the data tier is composed of two relational database management systems with several file servers. One of the database management systems manages system metadata such as users, roles, and client organizations. The second one houses hundreds of databases processed through FPHLM. Each database contains several processing stages of the input policy file along with the auxiliary information and scripts used during the pre-processing of the data. For the post-processing stage, Visual Basic for Applications (VBA) is used to build the user-defined functions, automate the processes, access the Windows API, and provide other low-level functionality through dynamic-link libraries.

While FPHLM manages the production usage of a long-term research project, domain knowledge including the application domain requirements and solution domain technologies changes frequently. For example, the requirements like new regulations from the Florida Commission on Hurricane Loss Projection Methodology (FCHLPM) may change. Therefore, FPHLM has both a production version and a development version. These requirements evolve new data and research findings; while the solution domain changes as new implementation technologies become available. In other words, the system is designed to be flexible and extensible. The development version is submitted and certified by FCHLPM, and becomes the next production version upon certification every two years. Since our system provides services to both FL OIR and private insurance companies, the whole system is designed for computationally complex components with stability and robustness from various disciplines .

V. Conclusions and Future Work

In this paper, an integrated and automated big data processing system assisted by domain knowledge for the FPHLM project is presented and introduced in details. FPHLM together with the proposed integrated data processing system are a novel computing framework that integrates functional components from various disciplines, and are able to provide long-term and up-to-date services for new insurance regulations. It can also provide the insights and help other researchers who are interested in working on hurricane catastrophe loss projection models.

With the development of the advanced hurricane research work, more factors can be considered. From the computer science point of views, the proposed integrated domain knowledge assisted data processing system can incorporate more domain knowledge and consider more advanced big data techniques like multithreading and Spark.

VI. Acknowledgment

This work is partially supported by the Florida Office of Insurance Regulation (OIR) under the Hurricane Loss Projection Model project. Opinions and conclusions expressed in this paper are those of the authors and do not necessarily reflect those of the Florida Office of Insurance Regulation.

References

[1] National Oceanic and Atmospheric Administration, “National hurricane center,” http://www.nhc.noaa.gov/, accessed June 2016.

[2] S. Venkataraman, W. Benger, A. Long, B. Jeong, and L. Renambot, “Visualizing hurricane katrina: Large data management, rendering and display challenges,” in *Proceedings of the 4th International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia*, ser. GRAPHITE ’06. New York, NY, USA: ACM, 2006, pp. 209–212. [Online]. Available: http://doi.acm.org/10.1145/1174429.1174465

[3] W. Hawkins and W. Zhang, “The effects of the hurricane katrina on louisiana’s gulf coast,” in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, vol. 1, July 2008, pp. I–434–I–437.

[4] E. Michel-Kerjan, S. Hochrainer-Stigler, H. Kunreuther, J. Linnerooth-Bayer, R. Mechler, R. Muir-Wood, N. Ranger, P. Vaziri, and M. Young, “Catastrophe risk models for evaluating disaster risk reduction investments in developing countries,” *Risk Analysis*, vol. 33, no. 6, pp. 984–999, 2013. [Online]. Available: http://dx.doi.org/10.1111/j.1539-6924.2012.01928.x

[5] M. G. Donat, T. Pardowitz, G. C. Leckebusch, U. Ulbrich, and O. Burghoff, “High-resolution refinement of a storm loss model and estimation of return periods of loss-intensive storms over germany,” *Natural Hazards and Earth System Sciences*, vol. 11, no. 10, pp. 2821–2833, 2011. [Online]. Available: http://www.nat-hazards-earth-syst-sci.net/11/2821/2011/

[6] Florida Public Hurricane Loss Model, “Florida public hurricane loss model 6.1,” http://fphlm.cs.fiu.edu/, accessed June 2016.

[7] K. Chatterjee, K. Saleem, N. Zhao, M. Chen, S. C. Chen, and S. S. Hamid, “Modeling methodology for component reuse and system integration for hurricane loss projection application,” in *Proceedings of the IEEE International Conference on Information Reuse Integration*, Sept 2006, pp. 57–62.

[8] S.-C. Chen, M. Chen, N. Zhao, S. Hamid, K. Chatterjee, and M. Armella, “Florida public hurricane loss model: Research in multi-disciplinary system integration assisting government policy making,” *Government Information Quarterly*, vol. 26, no. 2, pp. 285–294, 2009.

[9] S.-C. Chen, M. Chen, N. Zhao, S. Hamid, K. Saleem, and K. Chatterjee, “Florida public hurricane loss model (fphlm): Research experience in system integration,” in *Proceedings of the 2008 International Conference on Digital Government Research*, ser. dg.o ’08. Digital Government Society of North America, 2008, pp. 99–106. [Online]. Available: http://dl.acm.org/citation.cfm?id=1367832.1367851

[10] S.-C. Chen, S. Gulati, S. Hamid, X. Huang, L. Luo, N. Morisseau-Leroy, M. D. Powell, C. Zhan, and C. Zhang, “A web-based distributed system for hurricane occurrence projection,” *Software: Practice and Experience*, vol. 34, no. 6, pp. 549–571, 2004.

[11] S.-C. Chen, S. Gulatit, S. Hamid, X. Huang, L. Luo, N. Morisseau-Leroy, M. D. Powell, C. Zhan, and C. Zhang, “A three-tier system architecture design and development for hurricane occurrence simulation,” in *Proceedings of the International Conference on Information Technology: Research and Education*, Aug 2003, pp. 113–117.

[12] S.-C. Chen, S. Hamid, S. Gulati, G. Chen, X. Huang, L. Luo, C. Zhan, and C. Zhang, “Information reuse and system integration in the development of a hurricane simulation system,” in *Proceedings of the IEEE International Conference on Information Reuse and Integration*, Oct 2003, pp. 535–542.

[13] S.-C. Chen, S. Hamid, S. Gulati, N. Zhao, M. Chen, C. Zhang, and P. Gupta, “A reliable web-based system for hurricane analysis and simulation,” in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, vol. 6, Oct 2004, pp. 5215–5220 vol.6.

[14] F. Fleites, S. Cocke, S.-C. Chen, and S. Hamid, “Efficiently integrating mapreduce-based computing into a hurricane loss projection model,” in *Proceedings of the IEEE International Conference on Information Reuse and Integration*, Aug 2013, pp. 402–407.

[15] S. S. Hamid, J.-P. Pinelli, S.-C. Chen, and K. Gurley, “Catastrophe model-based assessment of hurricane risk and estimates of potential insured losses for the state of florida,” *Natural Hazards Review*, vol. 12, no. 4, pp. 171–176, 2011.

[16] D. Lohmann and F. Yue, “Correlation, simulation and uncertainty in catastrophe modeling,” in *Proceedings of the Winter Simulation Conference*, ser. WSC ’11. Winter Simulation Conference, 2011, pp. 133–145. [Online]. Available: http://dl.acm.org/citation.cfm?id=2431518.2431534

[17] P. J. Vickery, J. Lin, P. F. Skerlj, L. A. Twisdale Jr, and K. Huang, “Hazus-mh hurricane model methodology. i: Hurricane hazard, terrain, and wind load modeling,” *Natural Hazards Review*, vol. 7, no. 2, pp. 82–93, 2006.

[18] J. R. Fischbach, D. R. Johnson, D. S. Ortiz, B. P. Bryant, M. Hoover, and J. Ostwald, *Coastal Louisiana risk assessment model*. Rand Corporation, 2012.

[19] D. Viviroli, M. Zappa, J. Gurtz, and R. Weingartner, “An introduction to the hydrological modelling system prevah and its pre- and post-processing-tools,” *Environ. Model. Softw.*, vol. 24, no. 10, pp. 1209–1222, Oct. 2009. [Online]. Available: http://dx.doi.org/10.1016/j.envsoft.2009.04.001

[20] B. A. Colle, J. B. Olson, and J. S. Tongue, “Multiseason verification of the mm5. part i: Comparison with the eta model over the central and eastern united states and impact of mm5 resolution,” *Weather and Forecasting*, vol. 18, no. 3, pp. 431–457, 2003.

[21] H. R. Glahn and D. A. Lowry, “The use of model output statistics (mos) in objective weather forecasting,” *Journal of Applied Meteorology*, vol. 11, no. 8, pp. 1203–1211, 1972.

[22] A. K. Bahl, O. Baltzer, A. Rau-Chaplin, and B. Varghese, “Parallel simulations for analysing portfolios of catastrophic event risk,” in *Proceedings of the SC Companion: High Performance Computing, Networking, Storage and Analysis (SCC)*. IEEE, 2012, pp. 1176–1184.

[23] Y. m. Lin and Y. b. Li, “Defects and development of catastrophe insurance in china,” in *Proceedings of the International Conference on Information Systems for Crisis Response and Management*, Nov 2011, pp. 389–394.

[24] K. M. Clark, “The use of computer modeling in estimating and managing future catastrophe losses,” *The Geneva Papers on Risk and Insurance. Issues and Practice*, vol. 27, no. 2, pp. 181–195, 2002.

[25] C. Yun-Feng, Y. Fu-Cai, and L. Gen-Fa, “A catastrophe model for water bloom prediction: A case study of china’s lake chaohu,” *Human and Ecological Risk Assessment*, vol. 13, no. 4, pp. 914–921, 2007.

[26] Applied Research Associates, Inc., “Hurloss version 7.0.a,” https://www.sbafla.com/method/Portals/Methodology/ModelSubmissions/, accessed June 2016.

[27] Esri, “Arcmap,” http://desktop.arcgis.com/en/arcmap/, accessed June 2016.

[28] S. Hamid, B. G. Kibria, S. Gulati, M. Powell, B. Annane, S. Cocke, J.-P. Pinelli, K. Gurley, and S.-C. Chen, “Predicting losses of residential structures in the state of florida by the public hurricane loss evaluation model,” *Statistical methodology*, vol. 7, no. 5, pp. 552–573, 2010.