A Scalable and Automatic Validation Process for Florida Public Hurricane Loss Model

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

***Abstract*—The Florida Public Hurricane Loss Model (FPHLM) is a public catastrophe model that integrates and regulates all key components, such as meteorology, engineering, and actuarial components, by following a certain workflow in the execution phase. The validation phase governed by an Automatic Data Validation (ADV) program simulates each modeled execution component with a large number of historical insurance data of specific hurricane events. The differences between the actual losses and the modeled losses of the insurance portfolios are evaluated to validate the model. The original validation process is time-consuming and error-prone when handling large data sets. This paper presents how the automated computer program efficiently and correctly incorporates the key components and produces useful reports for the validation purposes. By considering sixty-six combinations (i.e., the combination of one company and one hurricane) of the claim data, the FPHLM model adopts the largest set of portfolios comparing to the other four private models, which makes the validation process more challenging.**

***Index Terms*—FPHLM; validation; catastrophe model.**

I. Introduction

Hurricane is one of the most severe natural disasters, which can cause significant losses of properties and endanger human life. When hurricanes come onto lands, the heavy rains, strong winds, and large waves can cause huge damages. The state of Florida is particularly vulnerable since all the insured properties in the state are exposed to hurricane storms. For instance, Hurricane Ivan, in 2004, landed in Florida and caused widespread damage with an estimated loss of about $22.6 billion, which makes it the fifth costliest hurricane to strike the United States. Hurricane Charley, in the same year, landed near Tampa and caused $15 billion loss, and imposed evacuation order on 1.9 million people living along the west coast of Florida. Hurricanes cause widespread damages to residential properties and infrastructure, uproot communities, curtail economic activities, and put severe stress on the insurance industry. Thus, it is important to be able to predict the economic and insurance consequences of hurricanes.

Funded by the Florida Office of Insurance Regulation (FL OIR), the Florida Public Hurricane Loss Model (FPHLM) aims to evaluate the wind risk and predict losses for personal or commercial residential properties located in Florida [1]–[3]. At the same time, the FPHLM model functions as a reliable tool for state regulators to better formulate rate-making policy for residential property insurance. The FPHLM model indeed has a considerable impact on the community since state regulators use the model results as a benchmark for purposes of evaluating and regulating home insurance rates.

The FPHLM model is the first and only hurricane catastrophe model open to public and is the result of multidisciplinary collaboration among experts in the fields of meteorology, structural engineering, computer science, and actuarial, statitics, etc. [4]–[6]. The FPHLM model was first launched in 2006 and has been used more than 1000 times by the state and over 130 times by insurance companies to help them forecast the potential insured wind losses for personal or commercial residential properties.

The FPHLM model consists of three major components: wind hazard (meteorology), vulnerability (engineering), and insured loss cost (actuarial) [7]–[9]. The major components are developed independently before being integrated, such that they are independently and theoretically sound without compensation for potential bias. The contribution of meteorology part is to simulate hurricane activities, to forcast wind tracks intensity and wind fields, and to answer questions on how hurricanes form; how these hurricanes decay, and how the terrain will affect their tracks. The engineering component produces vulnerability functions whose main purpose is to estimate the physical damage at different windspeeds to the exterior parts and interior parts of various types of buildings, as well as their contents. The actuarial component consists of a set of algorithms which uses the relevant output produced by the meteorology and engineering components to generate expected annual losses in the aggregates or by construction type, county, or ZIP code level, policy form, etc. or combinations thereof.

The FPHLM model runs on an integrated and scalable framework that consists of two phases: (a) an execution phase that includes pre-processing, storm forecast and wind field module, engineering vulnerability module, actuarial loss module, and post-processing parts; and (b) a validation phase that verifies the model output [10]. Because the FPHLM model has been used extensively by state regulators, by the insurance companies, and impacts Florida residents, the proper and efficient validation of the FPHLM model is important. In the validation phase, we carefully select nine historically well-observed hurricanes that passed by or made landfall in Florida. The FPHLM model can add new storms and quickly conduct new validation studies as new validation data becomes available, which makes the validation process scalable. The FPHLM model validation uses the HURDAT, Rmax, and Holland B databases for wind track construction [4], [11]. The validation suite includes Hurricane Andrew in 1992 and storms Charley, Frances, Jeanne, Ivan, Dennis, Katrina, Rita, and Wilma in the years 2004 and 2005.

Based on the model validation process, the FPHLM is considered novel due to the following characteristics: (1) It serves as the very first hurricane loss estimation model open to public and has a systematic validation plan which takes a number of parameters and factors into consideration; (2) The FPHLM model is able to keep the maximum number of insurer data to be used for the model validation, such that the obtained statistical analysis result is very good; (3) The modeled probability distributions of hurricane parameters and characteristics are consistent with historical hurricanes, and the modeled results are more consistent with insurance claims data; and (4) The FPHLM model exploits more abundant hurricane events, company sets, and so on in the validation phase. Detailed discussions of those aspects listed above are presented in the following sections.

The remainder of this paper is organized as follows. Section II discusses the proprietary models that have been also accepted by the Florida Commission on Hurricane Loss Projection Methodology, and compares the validation process of the FPHLM model with them. Section III details the overall process of the FPHLM model, including pre-processing, model simulation, and validation process. Section IV presents the evaluation of the validation results of the FPHLM model. Finally, section V presents conclusions and lists some potential future work.

II. Related Work on Validation

Model validation is considered as one of the most important steps in the FPHLM model design. The validation process helps evaluate how accurately, within acceptable bounds, the model predicts the losses [12]–[14]. In this section, different validation processes from other hurricane models are first discussed. Then, innovations in the validation process in our FPHLM model are examined.

A. Different Validation Models

Model validation has been one of the hot topics in the studies of Cat models. It aims to evaluate the models based on the differences between the modeled results and what happened in the real world. Paper [15] presents a model validation methodology with a hybrid dynamic simulation. The authors use two application instances on generator and load model validation to show the validity of their methodology in the field of power engineering. Similarly, [16] shows a power plant model validation to meet the reliability standard requirements. Utilizing the distinguishable events can better conduct the validation process in evaluating the model.

In terms of hurricane modeling, validation is also an important step. In the HAZUS-MH hurricane model methodology [17], the authors validate losses from modeling components with the insurance claims data. They validate each component (hurricane hazard model, empirical pressure coefficient model, and gravel debris model) through comparisons with field observations and wind tunnel data. Although the validation is through comparisons of data derived from both the historical data and the model simulation results, it is more like testing runs since they do not consider different parameters and factors, and do not have a systematic validation plan. Another paper [18] focuses on the construction of the validation dataset. When compared with the method in [17], the improvement is that they considered different sizes of loss distributions by county, state, and event. To construct a dataset, they select what they considered being the best available estimate of the industry aggregate insured losses for each event. Along with a normalization process, they use the validation data generated to perform the test runs. Albeit with a well-built validation dataset, more work needs to be done to design a validation system.

B. Innovations in FPHLM Validation

Created by the Florida Legislature in 1995, the Florida Commission on Hurricane Loss Projection Methodology (FCHLPM) is an independent body of experts in charge of developing standards and reviewing hurricane loss models for estimating probable maximum loss (PML) levels AAL for residential properties. The AAL and PML are used to determine insurance premiums and the reinsurance requirements. In this section, we briefly introduce the innovations in our model that was submitted to the Commission in 2015 under the 2013 standards, and compare the validation process of the FPHLM with some of the other models that are proprietary. Since the model is relatively transparent and public, a more detailed description is available. The FPHLM presents models for both personal residential and commercial residential properties. For example, the details of Monte Carlo simulation used to obtain the average annual loss costs and output ranges are mentioned in the submission document. As another example, the FPHLM shows the exact number when sampling error is negligible.

Secondly, the FPHLM model also has a series of functions which can be executed to check and validate the data and then prepare it for processing. There is a checklist to outline the initial tests that are performed. In addition, the mitigation attributes are checked for valid, numeric entries, and are mapped to the code description. Our validation is automatic including data pre-processing (cleans the data to fit it to the model), model analysis (generates wind speed correction data and simulates losses accordingly), and data analysis (compares the data generated by model with the actual history data and generated results).

While in some other hurricane models, the primary storm (e.g., radius, forward speed, and filling rate) and site (e.g., friction, gust factor) parameters are all random variables, the FPHLM uses 57,000 years of simulation with more than 40-thousand storms. The 57,000 is a carefully selected number based on mathematics foundation which decreases the uncertainty of the validation results. To demonstrate that the building vulnerability function relationships are consistent with the insurance claims data in the FPHLM model, the building loss consists of external and internal losses, appurtenant structure losses are derived independently, and all the losses are based on a combination of engineering principles, empirical equations, and engineering judgment.

III. Overall Process

The overall FPHLM validation process is composed of data pre-processing (e.g., data separation and data cleaning), step by step simulation of the routine, and the validation process of the combined outputs. The original insurance portfolios that are obtained from the selected companies are stored and maintained in the database. In order to meet the individual requirements of the input data that vary from component to component, complex SQL functions are created to fulfill the pre-processing step. Since all inputs are formatted before the environment deployment, the model simulation together with the outputs of the validation process can be executed by a flexible and automatic computer program.

A. Pre-processing

Data pre-processing is the process of transferring raw, noisy, and unstructured data into the clean and structured data for the further analysis. As the pre-processing step is domain-specific, different applications apply different pre-processing techniques. The first step of the FPHLM validation process involves several data pre-processing routines, such as data cleaning, data integration, and data reduction, to name a few. For this purpose, both premium data and claim data are pre-processed as described below.

1) Data separation: To conduct the FPHLM validation study, nine hurricanes happened in Florida were used, including Hurricane Andrew in 1992, major storms in 2004 (namely Charley, Frances, Jeanne, and Ivan), and major storms in 2005 (namely Dennis, Katrina, Rita, and Wilma) [4]. To pre-process the insurance portfolio, the hurricanes are separated based on the effective date and expiration date. For instance, Hurricane Charley made landfall in Florida on August 13, 2004 [19], and therefore the policies that have an effective date before that date and an expiration date after that date are included in the dataset for Charley. We also check if the effective data falls within the range of the hurricane to make sure that it is still included in the dataset for that hurricane. It is decided that if there is a claim in the data after the hurricane hit, the effective date was the day when the hurricane hit. Then, we separate data in two parts: (1) Premium data, the policies that included all the information and were active during the hurricane period, and (2) Claim data, the policies that were actually claiming for losses and the amounts of losses. Thus, two tables are created in the database, one for the claim data and one for then premium data. Finally, the premium data is extracted for insurance data processing; whereas the claim data is extracted for final validation, as explained in Sections III-B and III-C, respectively. In summary, the formatted claim data includes policy information, such as the county, zip code, construction type, and year built, as well as the actual loss information for building structure, content, appurtenant structure, Additional Living Expense (ALE), and total losses. On the other hand, the formatted premium data includes general policy information provided by the insurance company (e.g., construction type, exposures, deductibles, zip code, year built, and street address).

The proposed FPHLM validation model will cover all residential property types including personal residential (PR) and Commercial Residential (CR) for both low-rise (LR) and mid-high-rise (MHR) buildings. Therefore, the policies are divided into these three groups based on the number of stories of each building. In addition, building properties, such as roof cover, roof shape, roof to wall, and opening protection are assigned randomly for both PR and LR properties to evaluate the model capabilities even when the insurance companies do not provide such critical information.

2) Data cleaning: As raw insurance portfolios are noisy and sometimes incomplete, data cleaning is essential to resolve inconsistencies, assign missing values, and update flawed values. After the original data is divided based on different companies and hurricanes, both claim and premium datasets are automatically pre-processed using a large and complex set of SQL functions. The summary of data cleaning is explained as follows.

- Misspelled county names are updated using the county table which includes all the counties in Florida.
- Year built is set to zero when it is unknown or less than or equal to 1600. A year built is valid only if it is between 1600 and the year that the corresponding hurricane made landfall.
- County names and zip codes are updated based on the geographic information. There is a mapping table in the database which checks if counties and zip codes match.
- A region is assigned to each policy based on its county. There is a mapping table in the database which checks if counties, regions, and zip codes match.
- Construction types are updated. Generally, the construction types include frame, masonry, mobile home, and unknown. Unknown or invalid construction types are set to "Other".
- All the invalid (non-numeric) exposures are set to zero.
- All the unknown deductibles are updated to the average of the deductibles per storm policy file.
- For each policy, coordinates are assigned based on the zip code. There is a mapping table in the database which maps zip codes to the corresponding latitudes and

longitudes. Accordingly, two columns are added to each premium table as the policy coordinates which will be used later for winds analysis (please see Section III-B1).
- Finally, the order of columns for each dataset (the combination of a company and a hurricane) is corrected to be consistence for the further processes.

B. Model Simulation

FPHLM is intended to store, retrieve, and process an enormous amount of hurricane historical and simulated data. It is a large-scale system, which has the reinforced computing capability that supports hurricane damage valuation and insured loss projection.

As described in Section I, there are three major components in FPHLM, which consist of dozens of sub-components that simulate a hurricane from its occurrence to landfall and predict the total and insured losses for residential structures in Florida. The major components are developed independently before being integrated into a very complex set of computer programs. The computer platform is designed to accommodate future hookups of additional sub-components or enhancements.

1) Atmospheric Science Model: The atmospheric science component generates storm tracks and intensities up to close of land for simulated storms based on the premier conditions procured from the historical record of Atlantic tropical cyclone basin [4]. The output based on the stochastic algorithm contains 3-second terrain gust wind speeds for each of the hurricane affected zip codes. It is further corrected by considering the roughness.

As explained in Section III-A2, the input portfolio consists of a set of building properties. In order to be accepted by the meteorology component, street addresses need to be geocoded to obtain the geographic locations (i.e., latitude and longitude coordinates).

The meteorology component is further divided into storm forecast, wind field, and Wind Speed Correction (WSC) models. The storm forecast model is responsible for generating thousands of years of stochastic storm tracks based on the initial conditions followed by the intensity changes sampled from probability distribution functions created by trial and error. Once a simulated storm is moving close to within a threshold distance to Florida, the Wind Field model is activated to create the 1-hour marine surface wind swath for each storm that covers the entire Florida by using a fixed grid at every time step. The wind swaths over the fixed gird are stored into several look up tables that represent rectangular geographic areas of Florida. Since they are independent of the input property portfolio, these two models will be only executed once per version of FPHLM [10].

The final wind speeds will be used to design the vulnerability matrices, which will be described in the following section. The WSC component uses the roughness information and the marine surface winds to generate the terrain-corrected 3-second gust winds at the street level, which are followed by the actuarial component of FPHLM to estimate the expected insured losses for each property.

2) Vulnerability Model: By introducing Monte Carlo simulation, the vulnerability component determines the external risk of each type of residential constructions at various wind speeds that are generated by the meteorology component (introduced in the previous section). The internal, utilities, and content damages to the building are then predicted by the external vulnerability. The final formulation, which estimates the total building damage, represents as a set of vulnerability matrices for each building type that is significant in Florida, including manufactured homes. The vulnerability matrices are created for every combination of building characteristics (i.e., structural type, region, sub-region, and roof cover type).

The damage model is complemented with the prediction of contents, appurtenant structure damage, and ALE estimated by the damage ratio for each structure type.

3) Insured Loss Model: For each individual policy in the portfolio, the expected losses are predicted for each type of damages that were classified in Section III-B2 (e.g., building structure, appurtenant structure, contents, and ALE) based on the exposures and respective vulnerability matrices. For a given wind speed, the ground up loss is computed, deductibles and limits are applied, the loss net of deductible is calculated as well. The wind probability weighted loss is calculated to produce the expected loss for each property, which can be adjusted by the appropriate expected demand surge factor.

As briefly mentioned in Section III-A1, the Insured Loss Model (ILM) can be further divided into three sub-models: ILM-PR, ILM-LR, and ILM-MHR, in order to better associate with the other two components (as described above in several). The respective group of vulnerability and damage matrices is assigned to each sub-model. In addition to the regular output of ILM, the scenario based actuarial evaluation provides the expected losses for specific hurricane events that can be used to validate the model by comparing them with the actual hurricane losses.

As domain experts model each component individually, the overall framework of FPHLM is extremely complex for both validation and evaluation. The procedure designed for validating the entire model that fully covers all components included in FPHLM is discussed in the next section.

C. Validation Process

For model validation purpose, the sub-models, which are responsible for major individual components, are included in the completed simulation process, considering the actual and modeled losses for different hurricanes/companies for residential coverage. The logical flowchart for the whole process is presented in Figure 1. As illustrated in the figure, the required input data and running environment of each component are set up as designed to process the premium data of every company for some specified hurricanes in the data processing module. Subsequently, the validation component runs as an additional data analysis process to finalize the output of ILM as needed and calculates the differences between actual and modeled losses, followed by the generation of well-formatted summarization reports. In order to speed up the

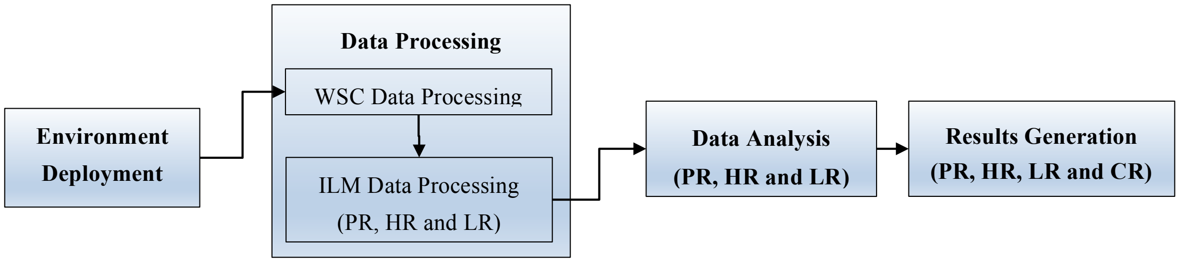

Fig. 1: Validation Program Logical Flowchart

whole validation process with less manual intervention, an automatic validation program is designed, which also reduces the chance of errors that may happen within the complex procedure while dealing with the complex data flow through multiple component models.

1) Automatic Data Validation Program: The Automatic Data Validation (ADV) program is a Java program, which is exported as a runnable JAR file that receives necessary input parameters from tiny configuration files, which meets the requirements of running all the models for every component. The program manipulates all the input and output data until the very end of the process, and gradually generates a set of results for all the company/hurricane portfolios.

2) Environment Deployment: The environment deployment integrates several configuration files and templates that will be used in each single run in one-off. Furthermore, relative data directories, like WSC output results – respective wind speeds of property in one single premium file, are defined in advance, in order to be automatically read by the next component (i.e., ILM) as the input.

For each company run, the templates are modified basing on the main configuration files with specified parameters and control the data generation of every single component. The data flow can be described as Figure 2.

3) Data Processing and Data Analysis: Starting with providing a list of companies that qualified for evaluation, the data validation environments are deployed and be ready to receive the well-prepared inputs (premium data and claim data), and then produce the output for later steps. The WSC model generates specified historical hurricane wind speeds according to the location of each property that feeds into the model by reading the premium files. To highly automate the validation process, three ILM models (ILM-PR, ILM-LR, and ILM-MHR), which are responsible for different types of residential buildings, are selected based on the premium data labeled at the very beginning and create the environment for each individual company/hurricane without conflicts. The vulnerability matrices are required on behalf of validating the engineering component. Since all the running environments for different components are well-created and defined with certain related inputs (i.e., swath files, rough tiles, and vulnerability matrices) and directories, interrupting the procedure to specify intermediate results that are going to devote to the next step becomes unnecessary.

By means of the ADV program, the process simulates regular FPHLM and accepts changes of each major component. Moreover, it is compatible to work with different versions of the model. Fewer modifications are needed when the model upgrades to a new version. The normal process runs the program from the very beginning to the end once and generates the results altogether. Besides, it is able to select the point of model re-running, if there is any changes that should be made in the middle, without reproducing everything from the initial step.

The welcome interface of the program is demonstrated in Figure 3. As mentioned above, the ADV program provides the functionality, which not only automatically arranges the configuration files while setting up the whole environment, and runs the whole model completely to generate the final results, but also can be partially used to reproduce the results by selecting the step number as the starting point.

4) Result Generation: After generating the hurricane specified modeled losses for different companies by ILM, the final step of simulating FPHLM, the historical actual losses are calculated based on the respective claim files. The eventual step of the validation process is formatting the modeled and actual losses, considering the demand surge factor for each part of residential coverage and comparing the differences through different aspects and combinations. For instance, the losses validation of commercial residential models can be considered as a whole or make comparisons separately within LR and MHR model. All the intermediate results translating and conveying the model losses, like county-level aggregations, are the calculations for the final validation output. They are very useful while creating deeper investigations to evaluate the model or to dig out the potential errors.

By representing the accumulated losses as tables for different purposes, the validation results can be further evaluated as described in the following section.

IV. Evaluation

In this section, FPHLM is compared with four private hurricane loss models, which are not able to be fully disclosed. The similar standard requirements were set by the Florida

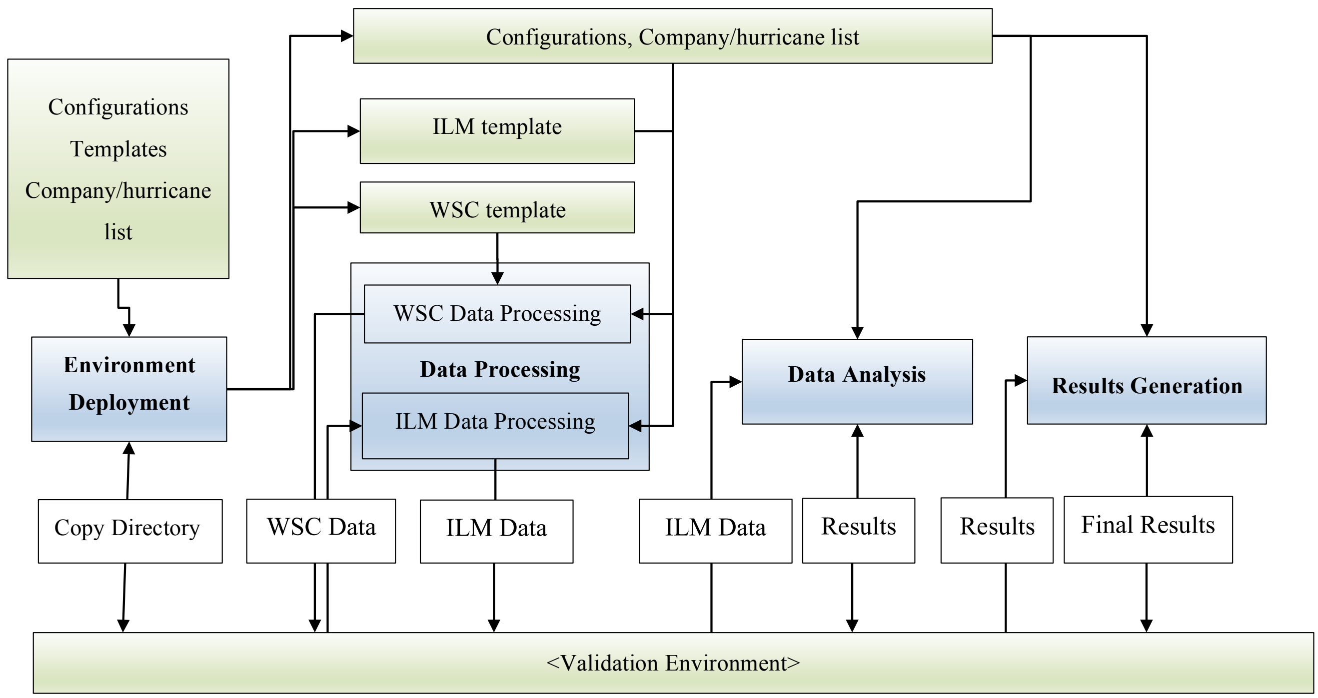

Fig. 2: Validation Environment Data Flowchart

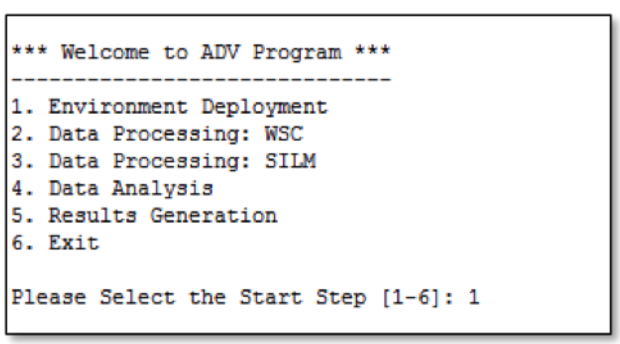

Fig. 3: ADV Program welcome interface

commission to validate Hurricane Loss Projection Methodology for each model. Thus, the validation results of each model are used and a series of evaluations is performed from different perspectives to justify FPHLM's validation process.

A. Validation Data Sets

Table I lists the information of the claim portfolios, which are used to validate each model, and to indicate the different levels of varieties these portfolios are. The number of companies represents the total number of insurance companies

TABLE I: Insurance company claim data sets used in different hurricane loss models

	Number of Companies	Number of Hurricanes	Total Number of Combinations
Modeler A	6	7	10
Modeler B	9	4	15
Modeler C	9	9	22
Modeler D	5	6	13
FPHLM	19	7	66

who provide both the insurance premium policies and the claim files to each model. The number of hurricanes indicates the number of hurricane events which caused the losses in each claim portfolio. The total number of combinations is not the product of the number of companies and the number of hurricanes in each model because one company might only provide the claim portfolio caused by one hurricane. It is distinguishable that FPHLM adopted the largest set of claim portfolios, when comparing to the other four private models.

B. Validation Results

The way of conducting the validation process is to compare the actual losses and the modeled losses from all possible perspectives. For example, presenting the results by personal

TABLE II: Validation comparison of different hurricane loss models

Model Name	PR + CR	# of Claim Portfolios >20	Present the results by		
			Coverage	Construction Type	County
Modeler A	✓				✓
Modeler B	✓		✓	✓	
Modeler C	✓	✓	✓	✓	✓
Modeler D	✓		✓	✓	✓
FPHLM	✓	✓	✓	✓	✓

residential or commercial residential, breaking down the results into county levels, presenting the results by construction types, etc.

Figure 4 presents a comparison of total actual losses vs. total modeled losses for different hurricanes of the FPHLM personal residential model. It shows a reasonable agreement between the actual and modeled losses. The correlation between actual and modeled losses is found to be 0.970, which shows a strong positive linear relationship between the actual and modeled losses.

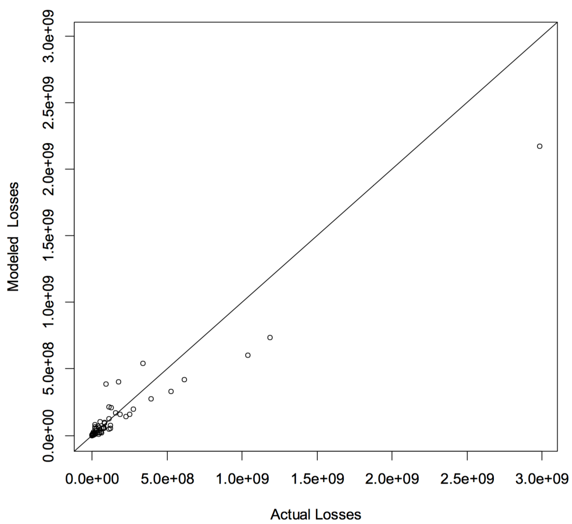

Fig. 4: Scatter plot between total actual vs. total modeled losses - Personal Residential

Table II presents the features that are included in the validation results per hurricane loss model, where the first column "PR+CR" meaning that the models can produce the results for both personal residential policies (PR) and commercial residential policies (CR). "# of Claim Portfolios" can refer to Table I. The last three columns indicate whether the model presents the validation results in different granularities. All the models are missing one of two features except the Modeler C, but the claim portfolios used in our model have essentially large losses compared to Modeler C's.

V. Conclusions and Future Work

FPHLM is a huge and complex model, which integrates several expert components from different domains to estimate the residential property losses caused by different events. However, the complexity increases the difficulties and challenges for model validation. In addition to sequentially imitate the regular running steps of FPHLM on the historical hurricane events for all the listed companies, the automatic program makes the requisite files well-organized and less modification is needed during the processing steps. Compared to the validation process of the other models, the ADV program successfully handles a greater amount of portfolios with respect to the combinations of companies and hurricanes, and is able to provide reasonably diversified comparison results.

Currently, the ADV program gradually processes the model losses by following the list of companies/hurricanes. Since the results of the items among the list are uncorrelated with each other, it is possible to set up the program to run in parallel. Consequently, it is expected to accelerate the process to a higher degree.

Acknowledgment

This work is partially supported by the Florida Office of Insurance Regulation 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] 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.

[2] S. Hamid, B. G. Kibria, S. Gulati, M. Powell, B. Annane, S. Cocke, J.-P. Pinelli, K. Gurley, and S.-C. Chen, "Authors responses to the discussion on predicting losses of residential structures in the state of florida by the public hurricane loss evaluation model," *Statistical Methodology*, vol. 7, no. 5, pp. 596 – 600, 2010. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1572312710000729

[3] F. C. Fleites, S. Cocke, S. C. Chen, and S. Hamid, "Efficiently integrating mapreduce-based computing into a hurricane loss projection model," in *2013 IEEE 14th International Conference on Information Reuse and Integration (IRI)*, Aug 2013, pp. 402–407.

[4] 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.

[5] 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.

[6] 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.

[7] Florida Public Hurricane Loss Model, "Florida public hurricane loss model 6.1," https://www.sbafla.com/method/Portals/ Methodology/ModelSubmissions/2015/20150509_FIU_2013Standards_SubmissionDocument.pdf, accessed June 2016.

[8] 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*. Digital Government Society of North America, 2008, pp. 99–106.

[9] 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 *2006 IEEE International Conference on Information Reuse and Integration*. IEEE, 2006, pp. 57–62.

[10] Y. Yang, D. Lopez, H. Tian, S. Pouyanfar, F. C. Fleites, S.-C. Chen, and S. Hamid, "Integrated execution framework for catastrophe modeling," in *2015 IEEE International Conference on Semantic Computing (ICSC)*. IEEE, 2015, pp. 201–207.

[11] H. Willoughby and M. Rahn, "Parametric representation of the primary hurricane vortex. part i: Observations and evaluation of the holland (1980) model," *Monthly Weather Review*, vol. 132, no. 12, pp. 3033–3048, 2004.

[12] 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 *2004 IEEE International Conference on Systems, Man and Cybernetics*, vol. 6. IEEE, 2004, pp. 5215–5220.

[13] 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 *IEEE International Conference on Information Reuse and Integration, 2003. IRI 2003.*, Oct 2003, pp. 535–542.

[14] S.-C. Chen, S. Gulati, 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 International Conference on Information Technology: Research and Education, 2003. ITRE2003.* IEEE, 2003, pp. 113–117.

[15] Z. Huang, M. Kosterev, R. T. Guttromson, and T. B. Nguyen, "Model validation with hybrid dynamic simulation," in *Power Engineering Society, IEEE General Meeting*, 2006.

[16] P. Pourbeik, C. Pink, and R. Bisbee, "Power plant model validation for achieving reliability standard requirements based on recorded on-line disturbance data," in *Power Systems Conference and Exposition (PSCE), 2011 IEEE/PES*, March 2011, pp. 1–9.

[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] D. J. Collins and S. P. Lowe, "A macro validation dataset for u.s. hurricane models," *Casualty Actuarial Society*, pp. 217–252, 2001.

[19] Flordia Division of Emergency Management, "Florida disaster," http://floridadisaster.org/hurricanes/2004/, 2011, retrieved at: 2016-06-08.