COE CST TENTH ANNUAL TECHNICAL MEETING

ARCTOS: Improved Population Clustering

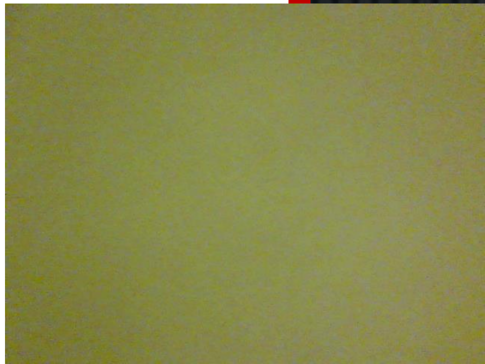
Wenshui Gan, Wije Wathugala

AGENDA

- Team Members
- Task Description
- Schedule
- Goals
- Results
- Conclusions and Future Work

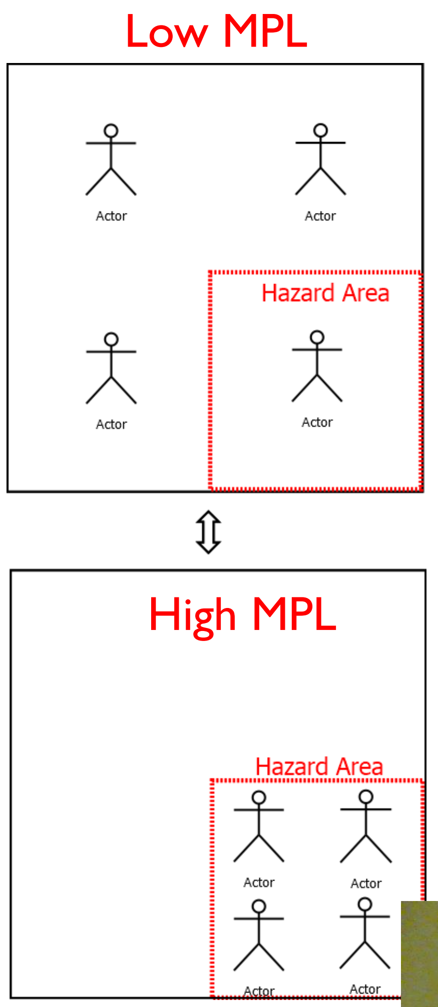
TEAM MEMBERS

- ARCTOS, Risk and Safety Division (www.arctos-us.com)
  - Wenshui Gan (PI)
    - Ph.D. Civil Engineering, Caltech
    - Senior Scientist, over 20 years of experience in risk engineering
  - Wije Wathugala
    - Ph.D., P.E.
    - Program Manager R&D
- Technical Monitors:
  - Dr. Paul Wilde — FAA Technical Lead



BACKGROUND

- FAA/AST seeks to advance the state-of-the-art for safety analyses of launch/re-entry vehicles by improving the modeling of dynamic clustering of people.
- Populations in flight safety analysis are generally represented as the number of people in a particular location.
- While the average population density is often sufficient for collective risk assessments, information about the spatial distributions of the people is often necessary for maximum probable loss (MPL) analysis.
- FAA/AST has identified that improved models of population clustering are critically important for an MPL analysis to rigorously account for transportation modes, recreational areas, and spectator areas.
- MPL analysis is used to determine the insurance requirements and premiums for a launch or reentry event.
- Higher MPL = Higher Premiums to the launch provider; therefore accuracy is key.



TASK DESCRIPTION

- Analyze imagery data on unsheltered populations in recreation, spectator and transportation areas, extract spatial distribution of people
- Based on the extracted data, develop a mathematical model that characterizes the clustering behavior of people
- Implement model into flight safety analysis software
- Evaluate the effects of the new model by comparing old and new models using past mission data

SCHEDULE

- Period of performance is between September 10, 2019 to September 10, 2021.
- As of September 2020
  - Completed population image processing (Oct 2019).
  - Developed a new clustering model based on the acquired data (Feb 2020).
  - Completed implementation of a fast-running model (Aug 2020).
  - Integrated the models in flight safety software (Sep 2020).
- To Do: Compare performance and accuracy of the new model with the current legacy clustering model for selected realistic applications.

GOALS

- Develop a realistic population clustering model to support vigorous maximum probable loss calculations in flight safety analyses.
- Optimize the implementation of the population clustering model to balance the requirements for speed and accuracy under different situations.
- Integrate the model into flight safety analysis tools to improve the calculations of the maximum probable loss and risk profiles for both the government and the commercial space industry.

APPROACH

- Review current models
- Collect observational population data
- Develop metrics model comparison
- Evaluate candidate models based on collected data
- Improve candidate models or develop a new model
- Develop strategies of integrating the model in flight safety analysis software
  - More accurate for both MPL calculations near launch area and downrange
  - Fast, especially for overflight of large areas

- Three models were evaluated by comparing with their capability to predict risk histograms to those computed for experimental population distributions
  - Legacy binomial clustering model
  - Ising based model developed in the last phase of this project
  - New logarithmic clustering model

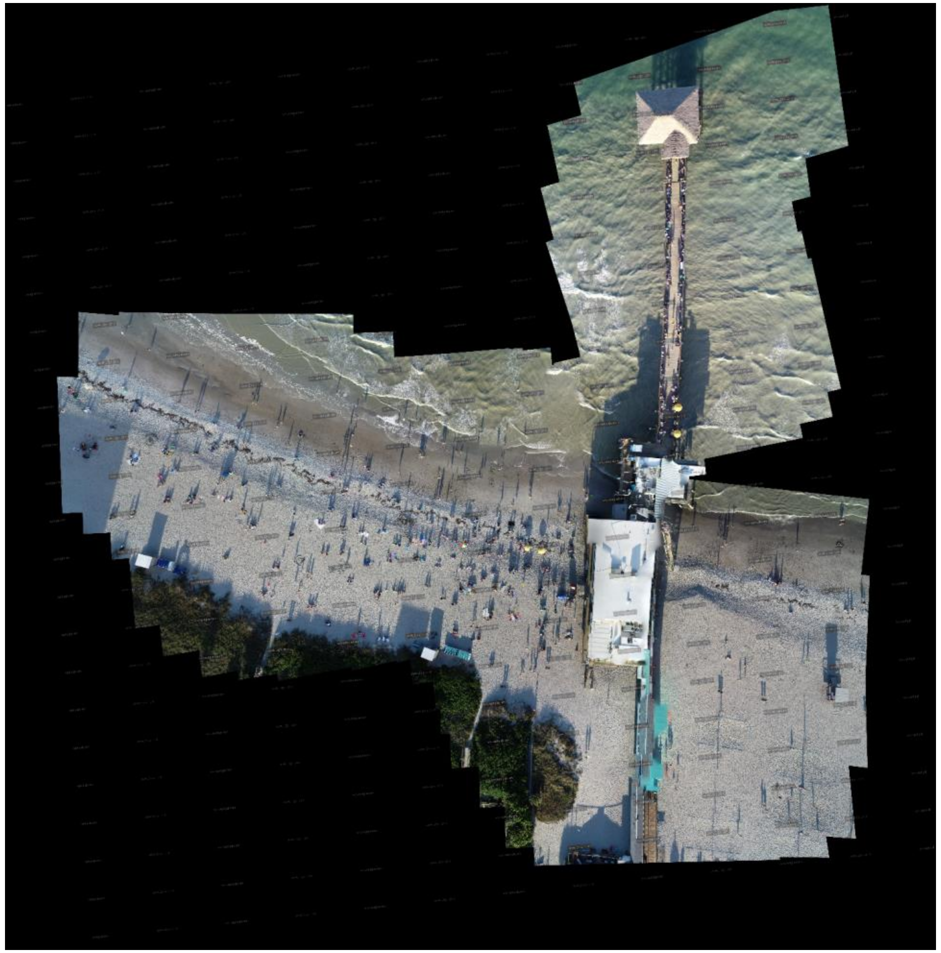
RESULTS

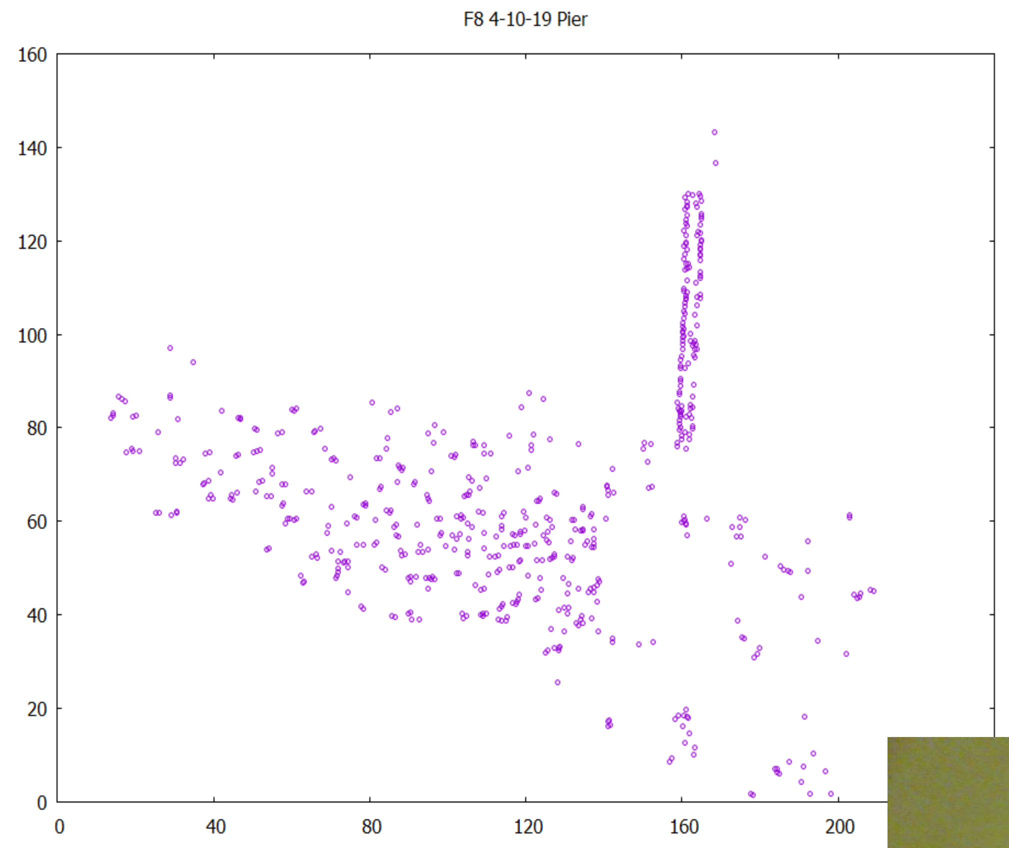
POPULATION DATA SETS ACQUIRED

- Categories: beaches, parks, outdoor markets, etc.

Data Set	Area Dimensions		Number of People	Square Feet Per Person	Data Set	Area Dimensions		Number of People	Square Feet Per Person
	Width	Height				Width	Height		
F9 4-10-19 early set1	260	60	126	1333	Bellflower set1	120	70	14	6458
F9 4-10-19 early set2	220	140	171	1939	Bellflower set2	15	45	8	908
F9 4-10-19 late set1	120	15	82	236	Cypress 3-1-19	380	220	1571	573
F9 4-10-19 late set2	150	50	116	696	Cypress 3-9-19	400	220	958	989
F9 4-10-19-pier set1 (pier)	6	55	124	29	Don Knabe 3-1-19	180	160	17	18235
F9 4-10-19-pier set2 (beach)	200	70	445	339	Don Knabe Halloween Early	120	180	50	4650
F9 4-11-19 early set1	90	5	134	36	Don Knabe Halloween Late	90	200	246	788
F9 4-11-19 early set2	250	100	232	1160	Heritage Park 2018-10-26	240	80	56	3690
F9 4-11-19 late set1	680	160	486	2410	Heritage Park mosaic	120	110	57	2493
F9 4-11-19 late set2	130	8	190	59	Lompoc	60	50	15	2153
F9 4-11-19 late p2	160	90	221	701	Wilson Park, Reduced1	200	70	165	[illegible]
Alamitos Beach	22	95	24	937	Wilson Park, Images3	24	6	12	[illegible]
Artesia 3-1-19	70	15	7	1615	Wilson Park, Images2	190	50	140	[illegible]
Artesia Mission 1	95	50	12	4261					

EXAMPLE – NEAR CAPE CANAVERAL, FLORIDA



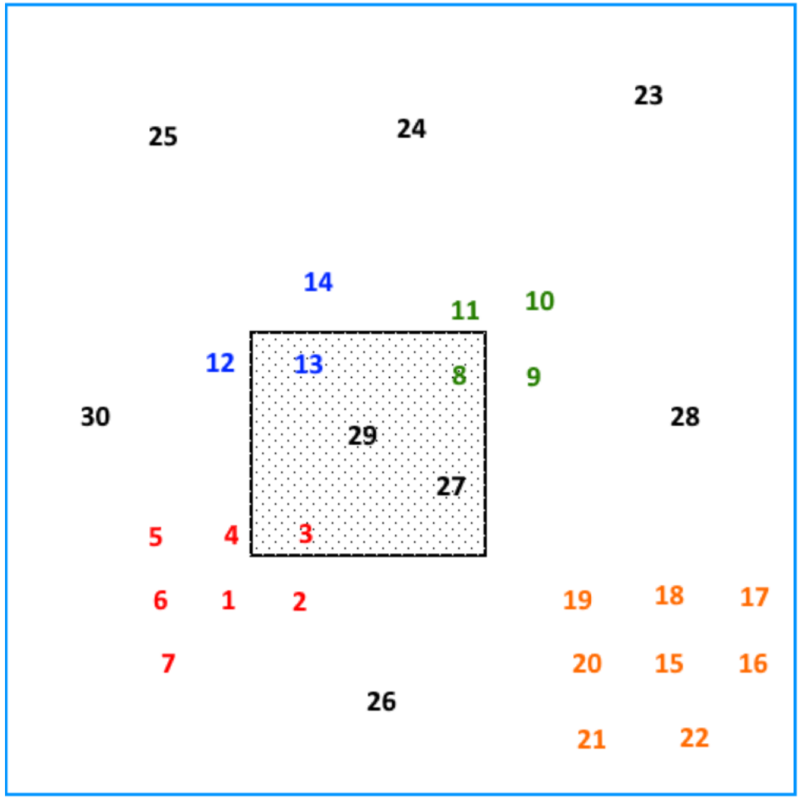


POPULATION CLUSTERING MODELS

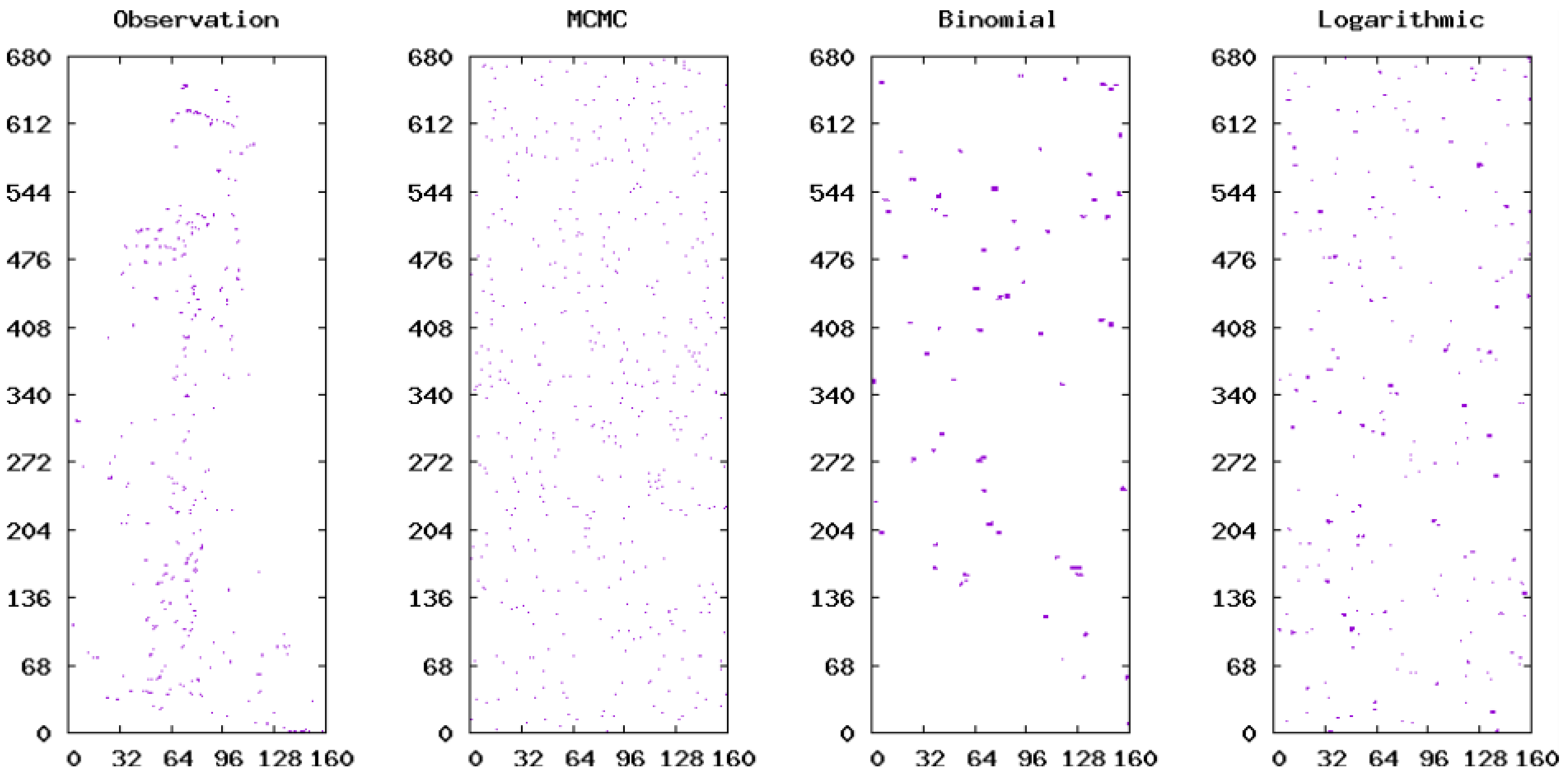
- Implicit Clustering Model
  - Based on the classic Ising model borrowed from statistical mechanics
  - Intriguing, but difficult to relate model parameters to observed population data
- Explicit Clustering Models – with stochastic parameters that can be estimated from observed population data
  - Cluster size:
    - Binomial (existing model)
    - Logarithmic (selected model)
  - Person to person distance
  - Cluster to cluster distance

LOGARITHMIC CLUSTERING MODEL

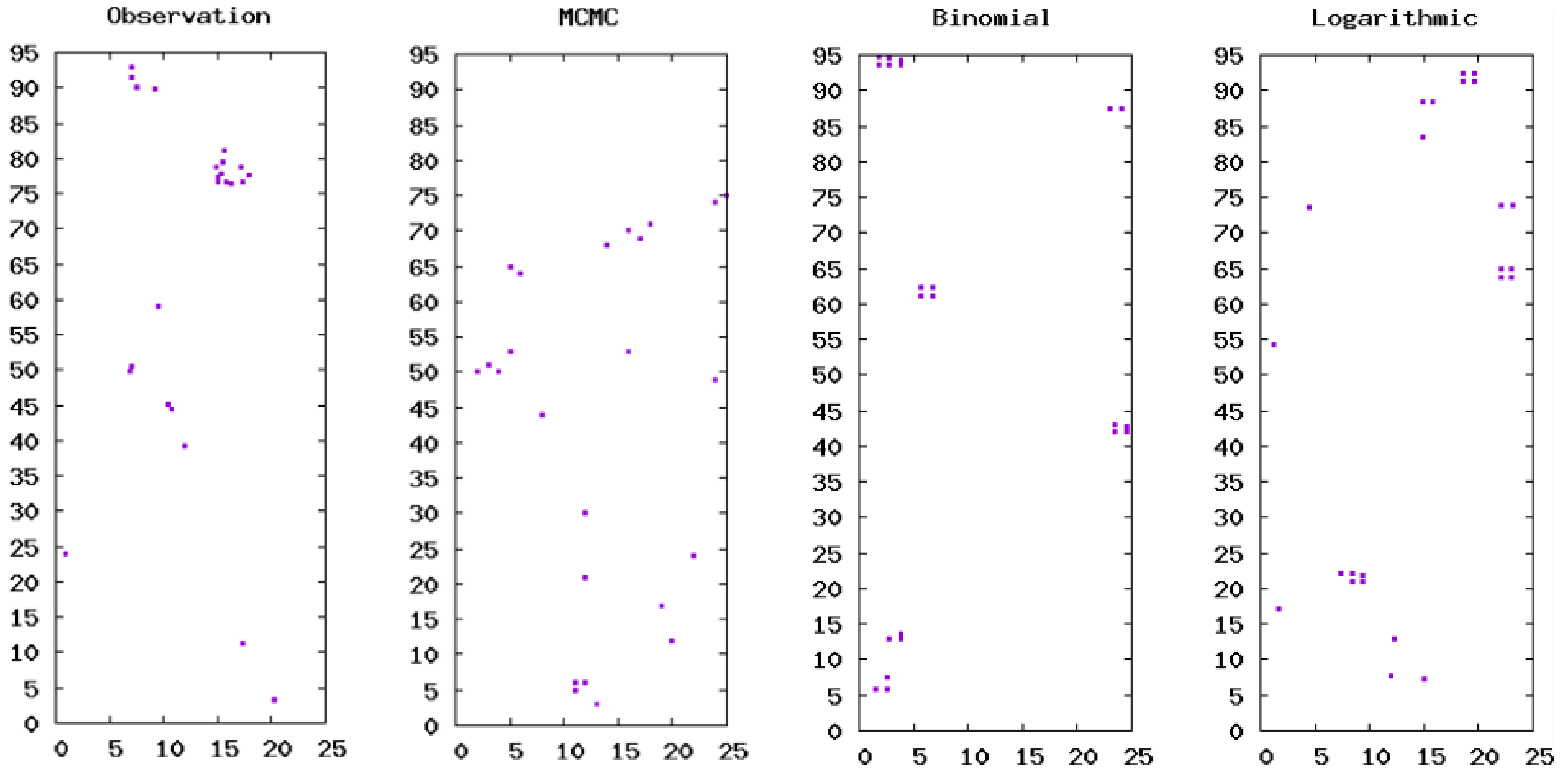
- Cluster size distribution
  - Logarithmic distribution $W(N) \propto N^{-1}(1 - 1/N_P)^N$
  - $W(N)$ is the fraction of clusters of size N.
  - N_P is a function of the mean cluster size $\overline{N}$
- Personal to person distance model
  - Log normal distribution fitted to observation data
- Uniform distribution for cluster locations in the domain.
- Algorithm to populate a domain
  - Sample a cluster size *n*
  - Sample a valid starting location
  - Place the *n* individuals at and around the starting location in an outward spiral pattern, with random person-to-person distances



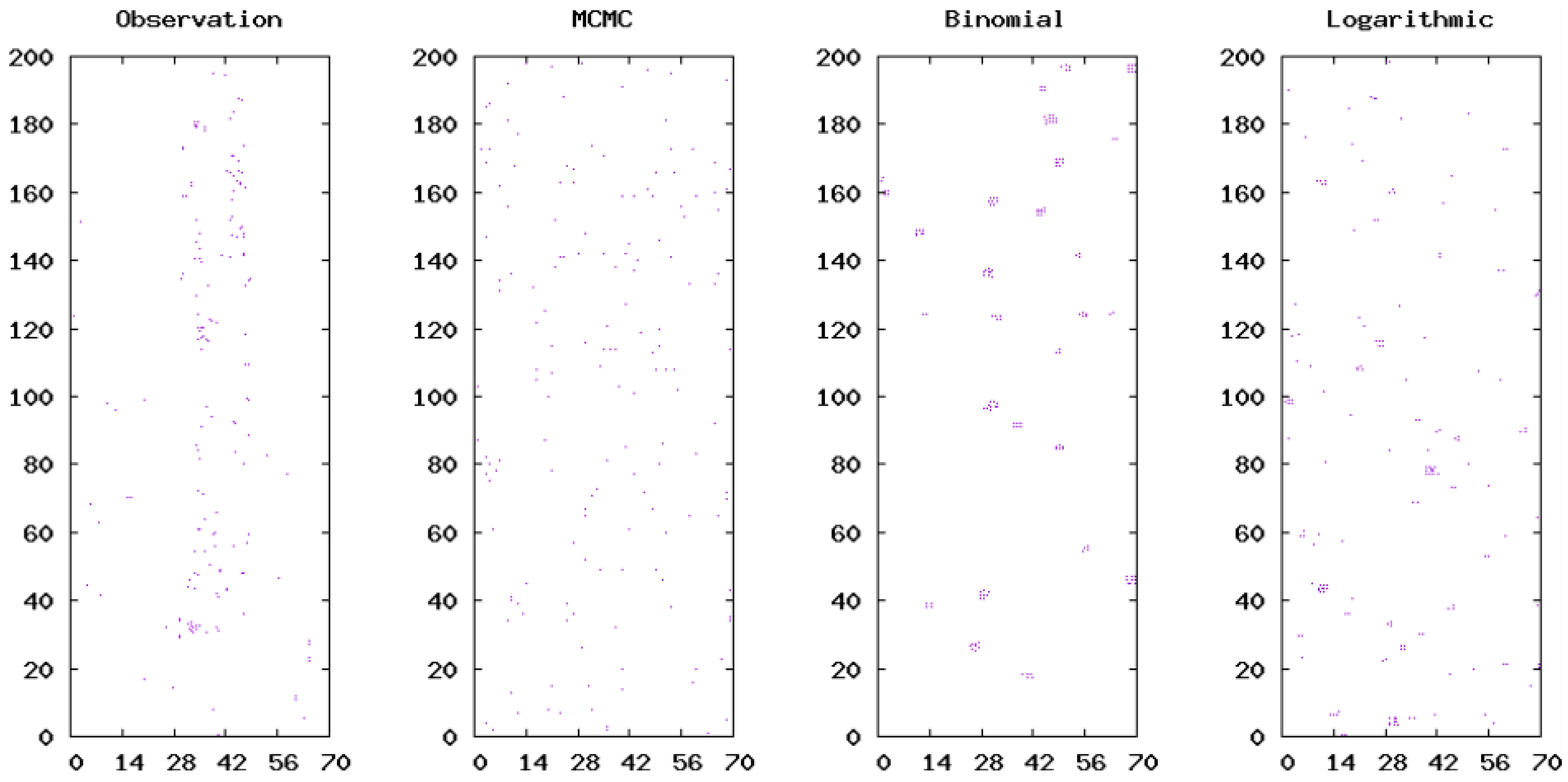
OBSERVED AND SIMULATED CLUSTERS (SITE 1)



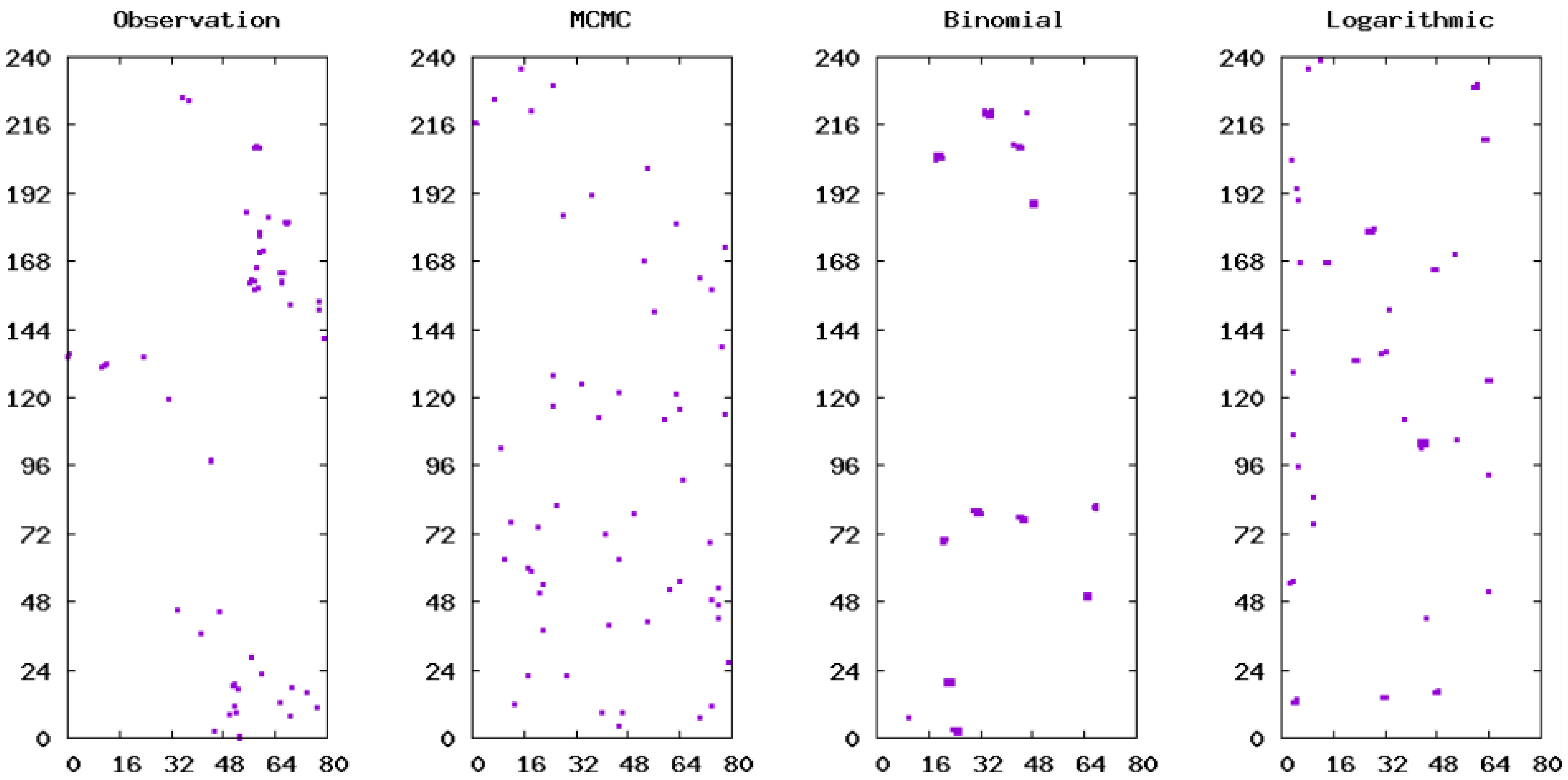
OBSERVED AND SIMULATED CLUSTERS (SITE 2)



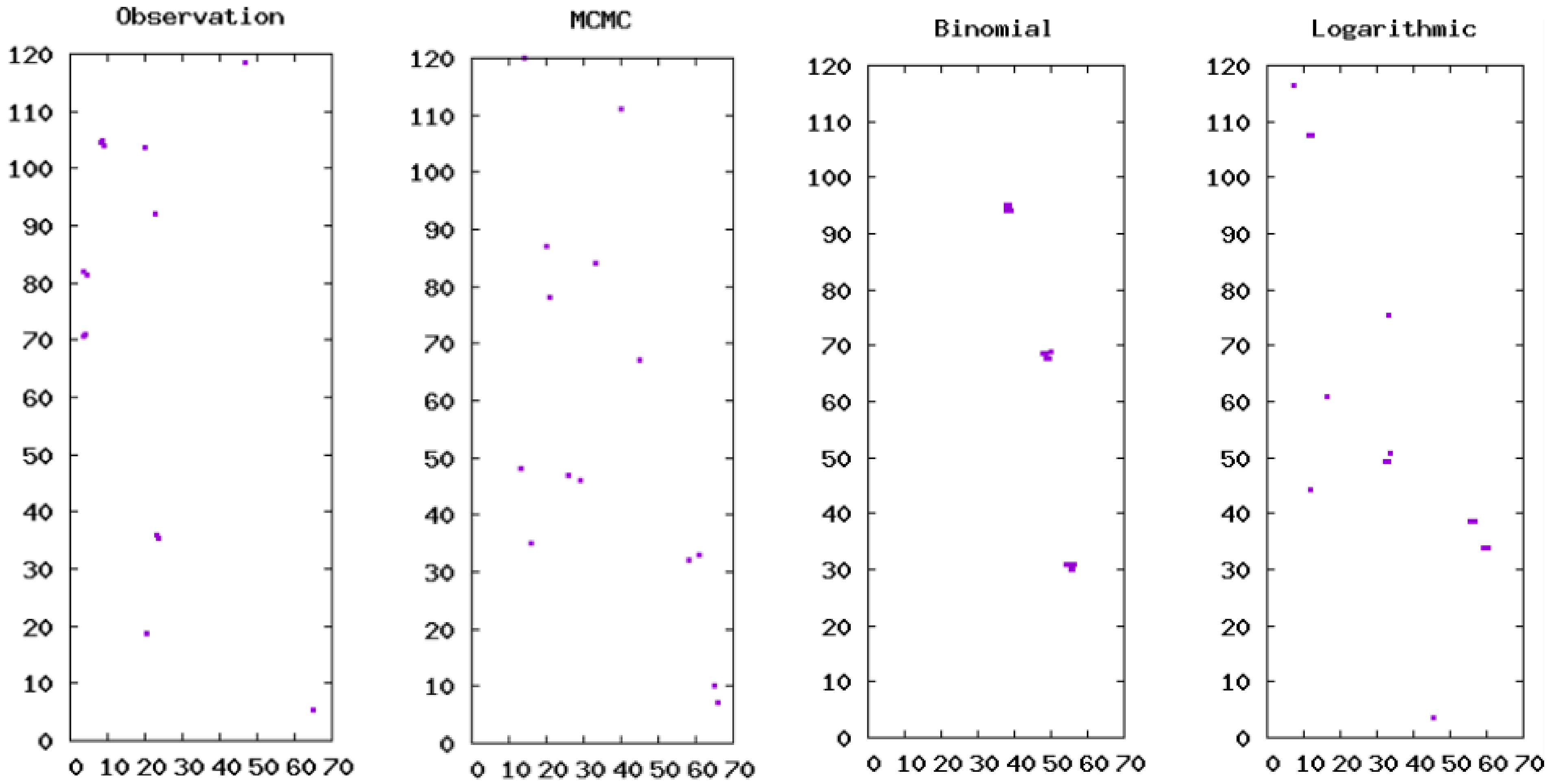
OBSERVED AND SIMULATED CLUSTERS (SITE 3)



OBSERVED AND SIMULATED CLUSTERS

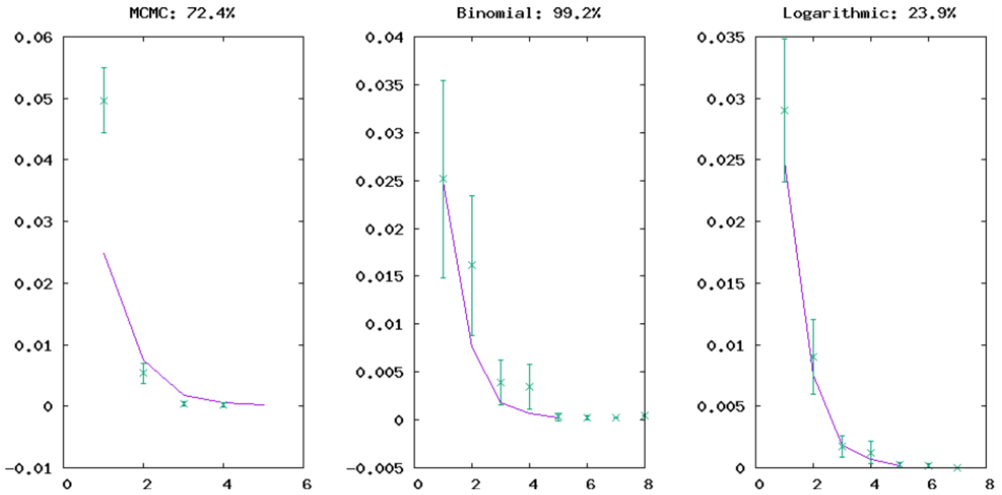


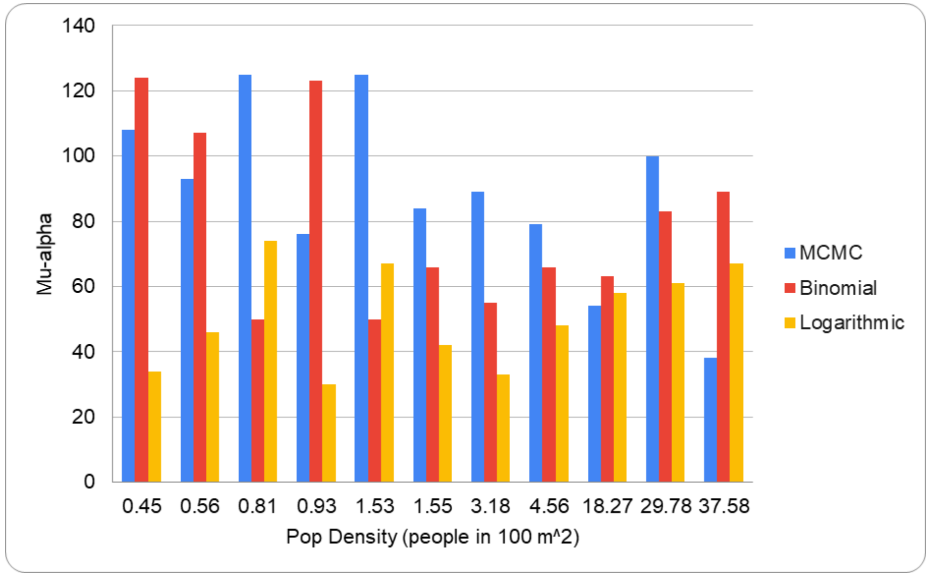
OBSERVED AND SIMULATED CLUSTERS (SITE 4)



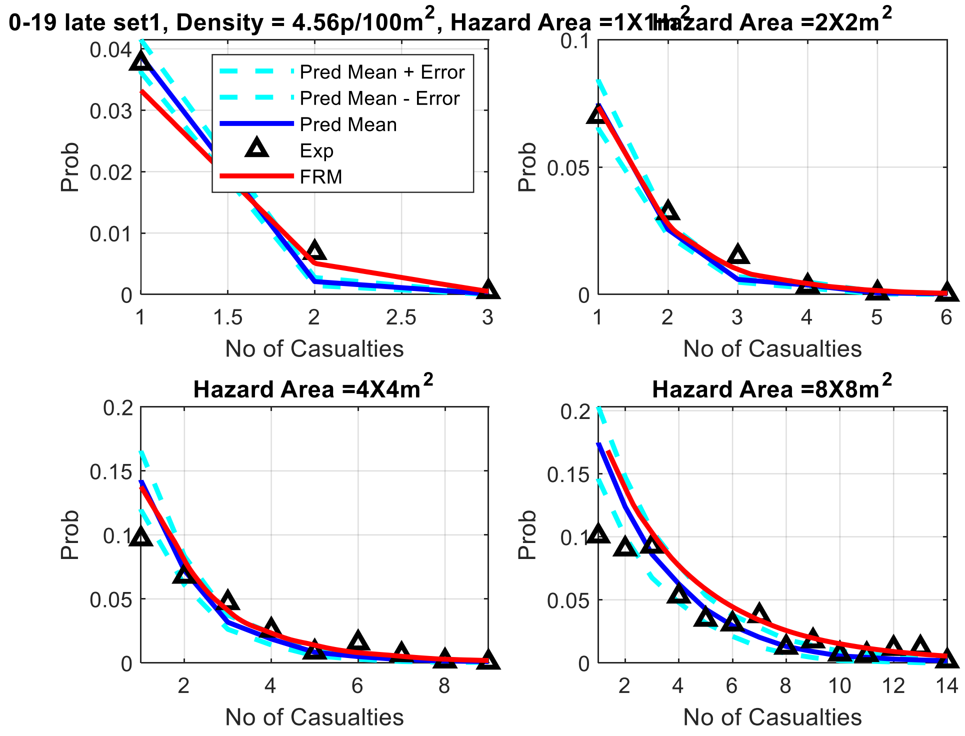
MODEL COMPARISON

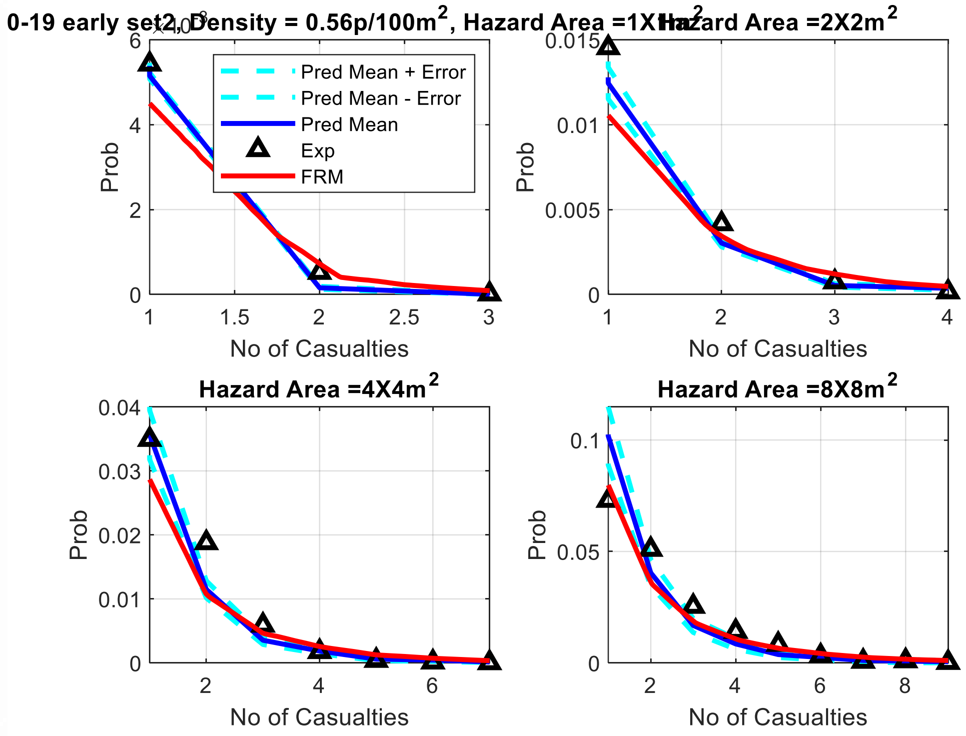
- How to compare?
  - Risk histogram
- $\varepsilon(\alpha) = \sum_{k=1}^{n} |P_k - Q_k| k^{\alpha}$, where P_k and Q_k are the probabilities of k casualties as predicted by a model and calculated from an observed population, respectively. Exponent $\alpha > 1$ is to give more weight to the probability of large number of fatalities.





FAST-RUNNING MODEL





CONCLUSIONS AND FUTURE WORK

- A new population clustering model was developed to provide improved risk profile and maximum probable loss calculations in flight safety analysis.
- A new FRM that can predict risk histograms without MC simulations was developed for fast run time for down range over flight cases.
- The new logarithmic model was found to predict risk histograms better than the previous models.
- The model has been integrated with an existing flight safety analysis tool developed by Arctos.
- Path Forward
  - Compare the new and legacy models for maximum probable loss calculations using past mission data.
  - Further calibrations of the model based on the above comparisons.