# Rapid Analysis of Simulation Outputs for Use in Emergency Preparedness in Animal Disease Outbreaks

| Melissa Schoenbaum, Lindsey Holmstrom | Karla Moreno-Torres, Shankar Yadav | Jonathan Arzt | Carolina Stenfeldt |
|---|---|---|---|
| **USDA–APHIS VS–STAS–CEAH** | **Oak Ridge Institute for Science and Education FADRU-ARS–USDA–PIADC** | **FADRU-ARS USDA–PIADC** | **University of Minnesota FADRU-ARS–USDA PIADC** |
| **Fort Collins, CO** | **Fort Collins, CO** | **Greenport, NY** | **Greenport, NY** |
| Melissa.Schoenbaum@ aphis.usda.gov, Lindsey.K.Holmstrom@ aphis.usda.gov | Karla.Moreno-Torres@ ars.usda.gov, Shankar.Yadav@ ars.usda.gov | Jonathan Arzt@ ars.usda.gov | Carolina.Stenfeldt@ ars.usda.gov |

## ABSTRACT

Stochastic animal disease models produce large quantities of data outputs regarding the epidemiological and economic impacts of disease outbreaks across a range of outbreak possibilities. The amount and complexity of data outputs can increase exponentially as more iterations and scenarios are performed to address the modeling question(s) of interest. Once analyzed, simulated output data can be valuable to decision makers for assessing the effects of different control strategies on outbreak impact and informing policy. In order for animal disease spread models to be useful tools to support emergency preparedness activities, the model outputs must be effectively provided to decision makers in a timely fashion. Typically, the process of obtaining the model output data and its analysis can be the rate limiting step that prevents model output from being readily used to inform decision making, as this process is dependent on the analytical skill sets of the modeler and access to appropriate software. We developed a suite of epidemiologic analyses and interactive model output visualization and scenario comparison tools using the R program (R Core Team, 2017) that link seamlessly to the outputs of the Animal Disease Spread Model (ADSM), a freely available modeling tool that can support emergency preparedness activities. A demonstration of the developed analysis and visualization R tools will be presented, employing a use-case of how they can be applied by a modeler to evaluate control strategies for a hypothetical foot-and-mouth disease outbreak and present a summary and comparison of the model scenarios to a decision maker.

## ABOUT THE AUTHORS

**Melissa Schoenbaum** is a biological scientist with the USDA–APHIS–VS–CEAH Monitoring and Modeling unit. She received her BS in biological sciences from East Texas State University and her MS in computer information technology from Regis University. She serves as the modeling team's data management specialist.

**Karla Moreno-Torres** is a veterinary epidemiologist and Oak Ridge Institute for Science and Education fellow who collaborates with the USDA–APHIS–VS–CEAH Monitoring and Modeling unit and with FADRU–ARS USDA–PIADC. Karla received her Master of Disease Ecology and Wildlife Management degree and her Doctor in Veterinary Medicine (MVZ) from the National Autonomous University of Mexico, and her PhD in epidemiology from The Ohio State University. Dr. Moreno's research involves computational modeling of infectious diseases in livestock and wildlife populations. Her current projects include parameterization of foot-and-mouth disease (FMD) states and modeling the infection dynamics of FMD virus.

**Shankar Yadav** is a postdoctoral researcher in veterinary epidemiology with FADRU–ARS USDA–PIADC at Plum Island Animal Disease Center, New York. Currently, he is collaborating with the USDA–APHIS–VS–CEAH

Monitoring and Modeling team. Dr. Yadav received his DVM from Tribhuvan University of Nepal. As a Fulbright Scholar, he earned a Masters Degree in Public Health (infectious disease and zoonoses) from Kansas State University and received his PhD in veterinary epidemiology from Purdue University. Before coming to the United Sates, Dr. Yadav worked for 6 years as a field veterinarian in Nepal with private and public organizations. His research interests involve mathematical modeling and epidemiologic studies on various highly infectious diseases of animals. Currently, his research focus is on parameterization of FMD states and investigation of the FMD virus persistence infection in cattle.

**Lindsey Holmstrom** is a veterinary epidemiologist with the USDA–APHIS–VS–CEAH Monitoring and Modeling unit. She received her BS in economics from Baylor University, her DVM from the College of Veterinary Medicine at Texas A&M University and did her PhD research at the University of California, Davis, focused on veterinary epidemiologic modeling. As part of the modeling team, she helps advance and apply epidemiologic models to develop scenarios involving disease spread, alternative control strategies, and associated impacts.

**Jonathan Arzt** is a veterinary medical officer at the Foreign Animal Disease Research Unit (FADRU), ARS–USDA at the Plum Island Animal Disease Center. He received a DVM and an MPVM from the University of California and received a PhD in pathology from Colorado State University.

**Carolina Stenfeldt** is a research associate at the University of Minnesota currently stationed at the Foreign Animal Disease Research Unit (FADRU), ARS–USDA at the Plum Island Animal Disease Center. She received her DVM from the University of Copenhagen, and a PhD in Veterinary Virology from the Technical University of Denmark. Her research focuses mainly on investigating aspects of pathogenesis and transmission of FMD in ruminants and pigs.

# Rapid Analysis of Simulation Outputs for Use in Emergency Preparedness in Animal Disease Outbreaks

| Melissa Schoenbaum, Lindsey Holmstrom | Karla Moreno-Torres, Shankar Yadav | Jonathan Arzt | Carolina Stenfeldt |
|---|---|---|---|
| **USDA–APHIS VS–STAS–CEAH** | **Oak Ridge Institute for Science and Education FADRU-ARS–USDA–PIADC** | **FADRU-ARS USDA–PIADC** | **University of Minnesota FADRU-ARS–USDA PIADC** |
| **Fort Collins, CO** | **Fort Collins, CO** | **Greenport, NY** | **Greenport, NY** |
| **Melissa.Schoenbaum@ aphis.usda.gov, Lindsey.K.Holmstrom@ aphis.usda.gov** | **Karla.Moreno-Torres@ ars.usda.gov, Shankar.Yadav@ ars.usda.gov** | **Jonathan Arzt@ ars.usda.gov** | **Carolina.Stenfeldt@ ars.usda.gov** |

## INTRODUCTION

The U. S. Department of Agriculture (USDA) has prioritized contagious animal diseases in various tiers. Foreign animal diseases (FADs), or diseases not currently in the United States, are a top priority due to their devastating impacts on animal health, animal well-being, and economy of the livestock industry if introduced into U.S. agriculture. Foot-and-mouth disease (FMD) is considered the most contagious and economically costly disease of cloven-hoofed animals. Early detection and characterization of an FMD outbreak is critical in order to implement effective control measures and reduce outbreak size and economic costs.

FMD outbreaks in previously disease-free countries have resulted in significant economic losses due to costs of disease control efforts and international trade restrictions on livestock commodities. For example, the 2001 FMD outbreak in the United Kingdom cost an estimated $13 billion (Thompson et al., 2002). It is anticipated that any incursion of FMD into the United States might lead to unprecedented economic losses for the livestock industry due to current livestock production and marketing practices. The U.S. livestock industry exports a significant percentage (one-third) of its production (meat and dairy products) to other countries, and all such exports would be banned for an extended period of time after detection of an FMD outbreak. Therefore, USDA has developed stringent control strategies for FMD which will be active once an outbreak is confirmed. These control strategies include depopulation of infected, detected livestock herds, quarantine of neighboring suspected farms within a 10-km ring of infected farms, and movement controls. As seen from FMD outbreaks in other countries, the impact of disease control strategies such as movement controls resulted in animal welfare problems that resulted in noninfected animals being culled (Thompson et al., 2002). As such, implementing alternative control strategies such as vaccination are now being considered. Depopulating/disposing and/or vaccinating a large number of animals during an outbreak are human resource intensive activities. U.S. animal health authorities need to be prepared before an FMD outbreak to know how they will estimate and manage such resources urgently during an outbreak.

For the selection and effective implementation of FMD control strategies before and during an outbreak emergency, decisions are guided mostly by the initial assessment and predictions of the nature of an outbreak (such as outbreak scope and size). Disease models, such as the Animal Disease Spread Model (ADSM), are tools decision makers can use to estimate consequences from an FMD outbreak and evaluate control strategies. ADSM (https://github.com/NAVADMC/ADSM/releases/latest) is a freely available tool that can be used for research and decision support for disease control programs. However, the model generates large amounts of data outputs, which might not be intuitive and easily understandable for decision makers. Once analyzed, simulated output data can be valuable to decision makers for assessing the effects of different control strategies on outbreak impact and informing

policy. In order for animal disease spread models to be useful tools to support emergency preparedness activities, the model outputs must be effectively provided to decision makers in a timely fashion. Typically, the process between obtaining the model output data and their analysis can be the rate-limiting step that prevents model output from being readily used to inform decision making, as this process is dependent on the analytical skill sets of the modeler and access to appropriate software. We have developed a suite of common epidemiologic analyses and interactive model output visualization and scenario comparison tools using the R program (R Core Team, 2017) that link seamlessly to the outputs of ADSM. Such a tool can be a powerful asset to animal health authorities for effectively incorporating models into their emergency preparedness activities.

## INTERACTIVE MODEL OUTPUT VISUALIZATION AND SCENARIO COMPARISON TOOLS

### The reality of comparing model scenario outputs

ADSM is effectively able to visualize a single scenario within the application, such as providing a summary map on completion of all iterations that illustrates the range of outcomes in a single scenario. Median statistics for critical variables are also displayed prominently for the selected scenario. Additional variables can be accessed using the Results tabs, organized by category, and each of the 274 possible variables can be visualized individually if parameter inputs were provided by the user. However, for emergency preparedness it is necessary to compare many possible scenarios. In emergencies, decision makers would like to be able to have rapid analytics to consider options as outbreak events are happening. ADSM currently does not provide a quick method for comparisons across multiple scenarios.

In reality, many outbreak scenarios have to be considered to determine optimal disease control strategies. Examples of questions that can be asked of disease models to assess outbreak consequences and optimal control strategies include:

- How long did the outbreak last and can we reduce the duration of the event?
- How many animals had to be destroyed and disposed of and can we reduce that number?
- How many doses of vaccine were given?
- How long will vaccinated animals stay in the food production chain, affecting national and international trade restrictions?
- How many farms/animals were located within quarantine and surveillance zones?

In this project, we created a base of R code that provides interactive visualization and summary tools for comparing different model scenarios and evaluating model questions such as those above. The tools are designed to be easily editable so different users can modify the tool outputs as needed based on their model questions of interest.

The developed tools can also be used for training exercises. Federal and State government agencies regularly stage training events to practice emergency preparedness activities that require complex coordination. A set of scenario outputs from disease models can provide a valuable component in these drills to incorporate characteristics of a disease outbreak.

This suite of tools is in three forms: an R Shiny application, an R Markdown document, and simple R scripts. Each of these can be further customized by a user requiring only basic R skills. Users can select relevant scenarios for comparison and modify variables as needed. These tools also provide a template for users who may want to include more model output variables relevant to their questions of interest.

### Shiny application in R

A Shiny application (Chang et al., 2017) and related packages (R Special Interest Group on Databases, 2017) used the R program to create an interactive interface. Users are able to select specific inputs and visualize analytically or descriptively the outputs of ADSM. It also allows users to download reports of their selected visualizations. The Shiny application can be deployed on the Internet for multiple-user access, although there may be a cost to the hosting entity, or deployed locally on a user's PC for free. The ADSM setup default is the local PC for implementation so no cost is initiated.

**Markdown document in R**

An R markdown (Allaire et al., 2017) and related packages (Baptiste A, 2017; Yihui X, 2014, 2015, 2017), is a tool with a different format. R Markdown provides a report in document form, such as a Microsoft Word document, PDF document, or HTML. The benefit of the R markdown is that it allows a link between the data and statistics behind the analysis with the documentation. The analytical code is embedded within the document format, so a user does not have to keep up with a data source, statistical method, or visualization attributes. This provides an organized, easily used package that can be re-executed as data are added to create an updated report with minimal work for users to compile data sources. It also allows for transparency and repeatability of outcomes, which are critical in complex data analyses.

**R scripts**

R scripts are simple code blocks that can provide analysis or visualization within the R program. These blocks of code can drill into a specific question from a large dataset. R scripts are the simplest form of analysis and may be easiest for new users to become experienced with manipulating ADSM data. Users are able to access the SQLite database of ADSM using a method that is commonly accepted in the scientific community and freely available to users around the world (Müller et al., 2017). Users can experiment with concepts in simple blocks of R code and then incorporate those into R markdown or R Shiny.

The time to implement a working prototype of all three of these R tools took only a few days for a user with some basic R skills.

**USE CASE: COMPARING OUTPUTS OF MODEL SCENARIOS OF FMD SIMULATED OUTBREAKS**

**Overview: Simulating disease spread with ADSM**

ADSM is a stochastic, spatially explicit software application that simulates the spread of highly contagious animal diseases between herds or flocks. It can be considered a compartmental epidemiological model, as herds start in a susceptible disease state and may transition into an infected and recovered state or compartment if exposed to the selected disease. The stochastic nature of the model allows for the creation of a range of outcomes that mimic the inherent variability in complex natural systems.

ADSM may be used to simulate a number of highly contagious animal diseases. The model takes input parameters that determine the characteristics of the disease. It also includes parameters that dictate the transmission of disease by contact or air. Users are able to implement measures to control the disease, such as movement restrictions, likelihood of detection, vaccination, and depopulation of diseased animals. The application is able to calculate basic economic costs such as cost of vaccination per animal vaccinated.

Herds/flocks are described by both their species and by their management practice which is called a production type. For example, cattle may be designated to go into dairy production or beef production. This difference in management practice may change how some of the parameters are applied in the model by the user. The management practice also changes how the animals come in contact with other animals to spread disease.

**Method**

*Model parameters*

We developed an FMD model using ADSM (version 3.3.8.5) to simulate an outbreak that starts in Texas with potential for disease spread in six surrounding states: Arkansas, Colorado, Kansas, Louisiana, New Mexico, and Oklahoma. For this model, ADSM used synthesized locations of 12 different production types for the study area: company feedlots (>50,000 head), stockholder feedlot (20,000 to 50,000 head), custom feedlot (5,000 to 20,000 head), yearling-pasture feedlot (<5,000 head), large cow-calf (≥100 head), small cow-calf (<100 head), stockers (<35,000 head), large dairy (≥2,000 head), small dairy (<2,000 head), large swine (≥200 head), small swine (<200 head), and small ruminants (all sizes). Parameters for the duration (in days) of the latent, subclinical infectious, and clinical infectious

periods for each production type were based on animal-level disease data from studies published on FMD experimental infections (USDA, 2013). Disease transmission via direct contact, indirect contact, and local-area spread between infected and susceptible herds was modeled.

The detection and reporting of infected herds was modeled based on the likelihood that livestock owners observe sick animals and report disease. We varied this likelihood of disease detection as discussed below. Once a herd is detected, contact herds are identified through trace-forward and track-back investigations which were modeled based on literature review and subject matter expert opinion. Movement restrictions within 10-km quarantine and surveillance zones around detected herds were implemented 1 day after initial detection. All detected herds were depopulated. When implemented, 10-km circular vaccination rings were created around detected herds and only cattle and swine species on production types within the ring were vaccinated.

*Model scenarios*

To demonstrate the R suite of tools for model scenario comparison, we modeled four different scenarios that evaluated the effect of time to FMD detection and vaccination as follows:
- Detection of disease 10 days after FMD introduction with baseline control measures (quarantine + depopulation), denoted as "Early" in the figures below
- Detection of disease 20 days after FMD introduction with baseline control measures, denoted as "Late" in the figures below
- Detection of disease 10 days after FMD introduction with baseline + vaccination control measures, denoted as "Vx Early" in the figures below
- Detection of disease 20 days after FMD introduction with baseline + vaccination control measures, denoted as "Vx Late" in the figures below.

For all scenarios, a large cow-calf operation in Moore County, Texas, was used as the start location. We ran 500 iterations of each scenario.

In addition to the standard ADSM outputs that are generated into the SQLite database, supplemental outputs were selected. The supplemental outputs provide herd-level detail of each daily event and each daily disease exposure. As can be seen in this example, the model results for all four scenarios created an abundance of output for analysis, as shown in Table 1.

**Table 1. ADSM outputs for four FMD model scenarios in Texas**

| Level of variable | Output | Count |
|---|---|---|
| Daily variables | Columns | 274 * 4 |
| Herd variable | Columns | 24 * 500 |
| Daily records | Rows | 3,642,375 |
| Herd record | Rows | 769,931 |
| Herd events | Files | 2,000 |

**Results of scenario comparisons**

Using the abundant data created from the four outbreak scenarios, we used the developed suite of R tools for scenario comparison. The sections below highlight a portion of the analysis and show median results to give an example of how to use the tools and the outputs that are available.

**Results in the Shiny application in R**

The Shiny application deploys in a browser window. Three sections of analysis are presented on different tabs: Outbreak Characteristics, Epidemic Spread, and Herd Level Analyses (Figure 2). R Shiny allows the user to interact with the application using a set of check boxes delineating the scenarios to be compared. The user can navigate between tabs and be able to compare various epidemiological metrics across the selected scenarios. In Figure 2, a scenario has been selected and the median summary of infected farms can be visualized per week (upper map) and day (bottom map), using the respective slider. Furthermore, a frequency plot provides the veterinary action occurring during the specific week of the simulated outbreak. A click over option opens a window with information about the farm. Finally, the maps have the function of zoom in and out, thus improving the user interaction to evaluate direction of the spread and geographic attributes of the farm surroundings. This example was purposefully located in the ocean due to sample data. In this case, the user could optionally visualize maps for each of the four scenarios.
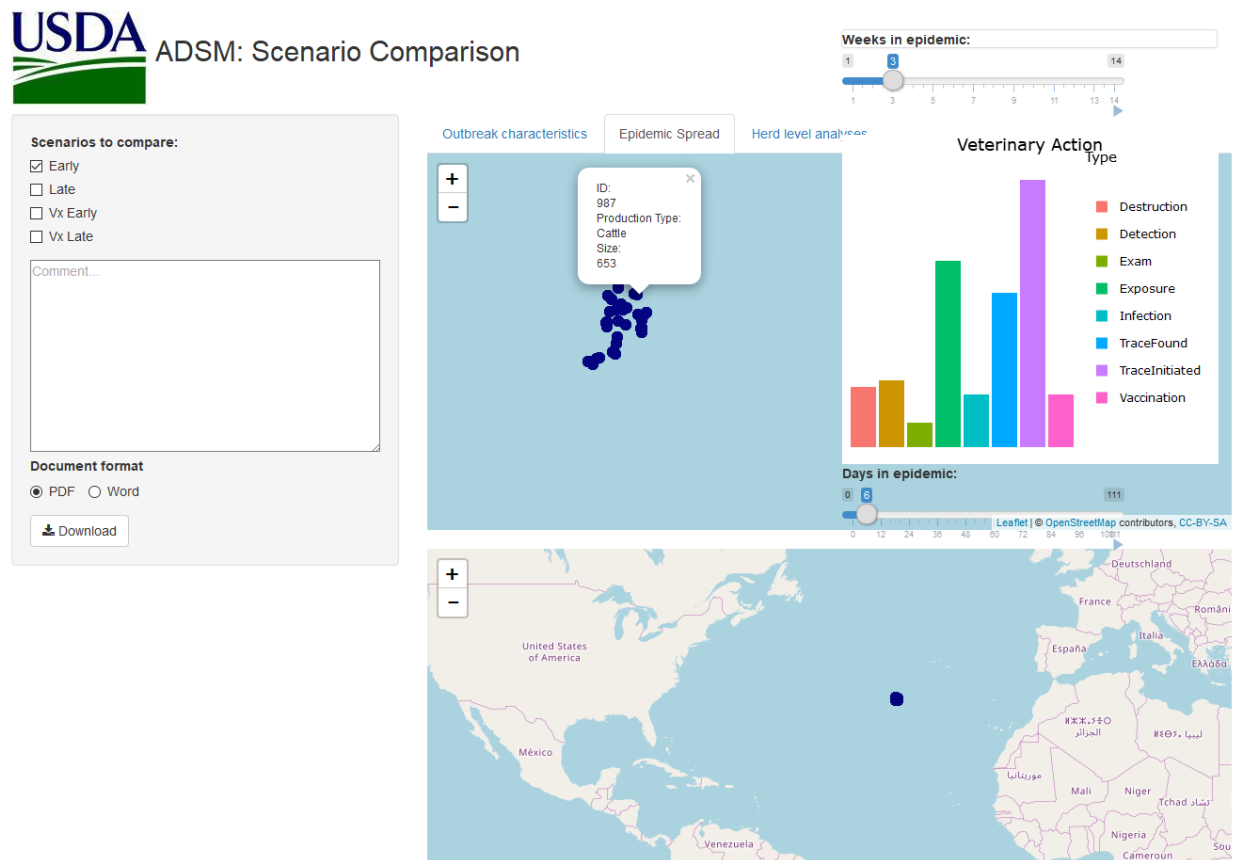


**Figure 2. Shiny application showing the Epidemic Spread scenario comparison**

Switching to the Outbreak Characteristics tab, the user can view the Epidemic Curve and Outbreak Duration visualizations as time-series and boxplot graphs (Figure 3). All scenarios are included in each image as shown in Figure 3. We can quickly see that early detection results in better outbreak outcomes for all control strategies implemented.
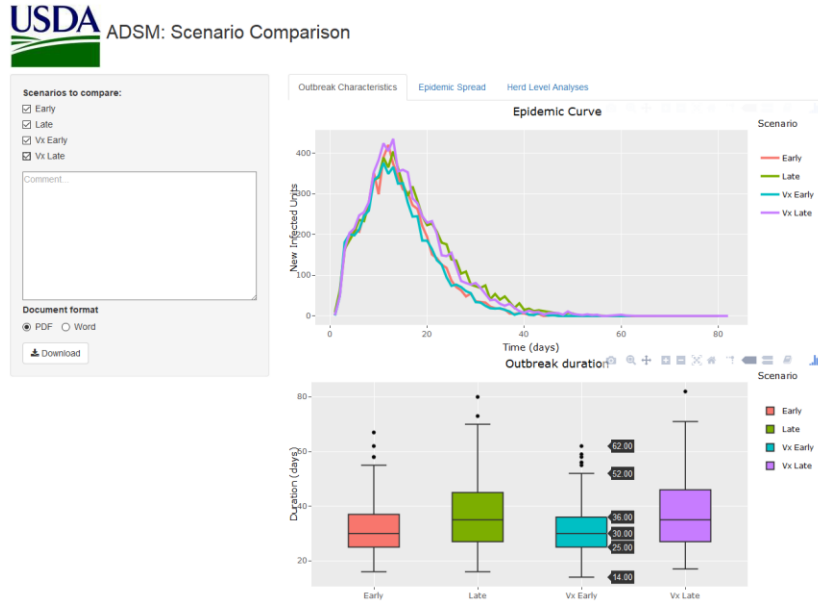
**Figure 3: Shiny application showing the Outbreak Characteristics scenario comparison**
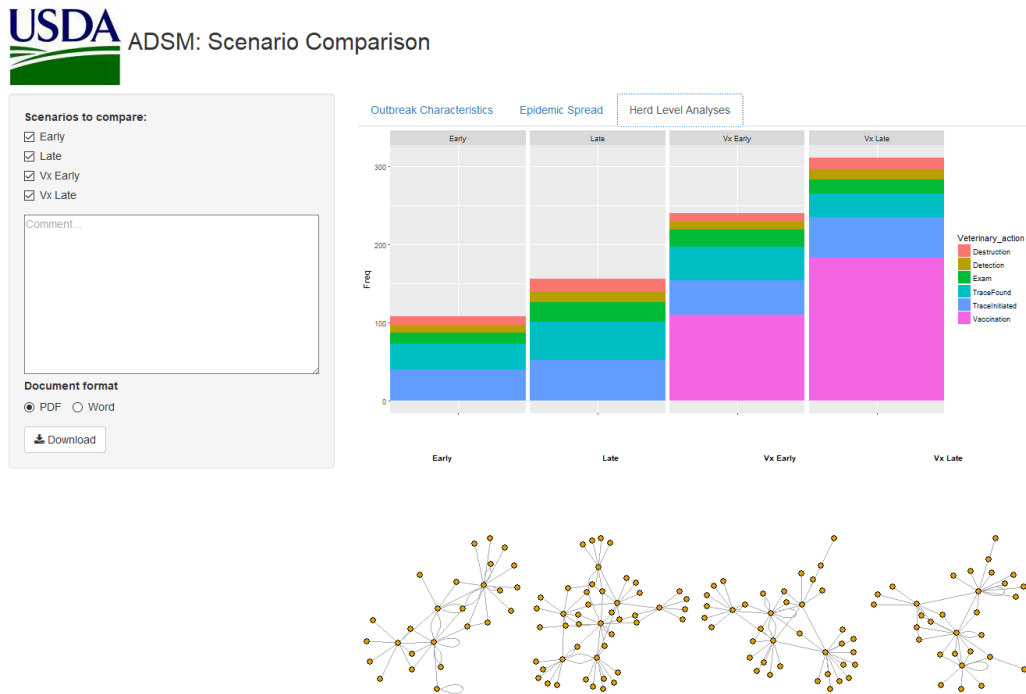


**Figure 4: Shiny application showing the Herd Level Analysis scenario comparison**

The Shiny application also provides several visualization options in the Herd Level Analyses tab. The first visualization in Figure 4 allows the comparison of the ADSM daily events output for the median results of each scenario. As can be seen, delayed detection results in more trace-back requirements to identify infected farms and more farms being depopulated compared with early detection. Delayed detection also results in more farms that use vaccination. The second visualization shows the disease transmission network, which allows users to evaluate disease spread between farms and identify farm-to-farm interactions that result in increased risk of disease spread.

**Markdown document in R**

The R markdown creates a summary document of the results in a traditional and familiar format for decision makers that can be shared electronically (e.g., via email). The process is repeatable and transparent, as the user does not have to change the programming code or worry about documenting the steps that were involved in the data analysis process, as this is automatically done by R markdown. The steps for all analyses are available for review, if needed.

The R markdown document was designed based on documents used by the USDA–APHIS–VS–STAS–CEAH modeling team to report modeling results to USDA leadership during previous animal disease outbreaks that have occurred in the United States within the last 2 years. Figure 5 shows the first two pages generated by the R markdown into MS Word.



**Figure 5: R markdown document**

**R scripts**

R scripts are stand-alone code blocks that can create a simpler visualization. Users can use the scripts as components to determine which visualizations would be most important to add to an R markdown or Shiny application. The inclusion of R scripts permits the ability for users to share code with other users. R scripts also give the user a starting point to expand upon versus having to start data analysis from scratch. Figure 6 shows an example of scripts in R Studio (RStudio Team, 2016) showing the code with the package ggplot2 (Wickham H, 2009), the run console, the environment with the ADSM data, and the visualization example.
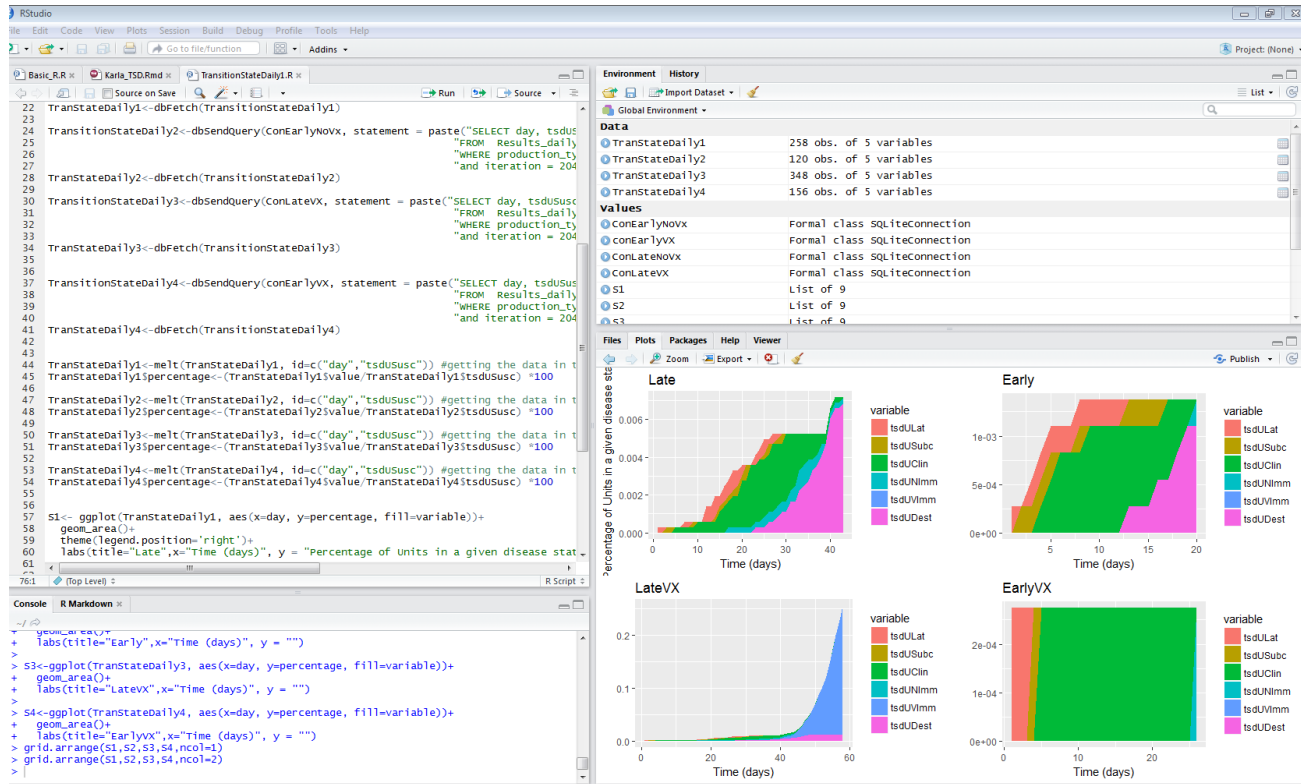
**Figure 6: Example of R script showing Transition State Daily variables**

## Implementation of R tools in ADSM

The most notable value of these R tools is that they can be seamlessly implemented with ADSM. Upon installation, ADSM creates a file system called the ADSM Workspace that is used by the application to manage scenarios, input and output files. This file system is separate from the program file installation and provides a flexible method for delivering additional code to users.

Currently, the ADSM Workspace contains a directory called Example Database Queries with examples of code for SQL queries within text files which can be cut and pasted into any SQL browser once a database connection has been created. Query scripts allow the user to access the SQLite database behind ADSM. The query scripts can be easily modified by users to provide custom outputs. The distribution of the suite of R tools will work in a similar fashion and be located in a directory named Example R Code that is built as the ADSM application installs on a user's computer. It will require that users install the free R program that is widely used in the scientific community. As more R tools are developed, they will be packaged with ADSM and available to users.

## CONCLUSION

We developed a suite of common epidemiologic analyses and interactive model output visualization and scenario comparison tools using the R program that seamlessly link to the outputs of ADSM. Using the model results from four scenarios of a simulated outbreak of FMD starting in Texas, we demonstrated how large datasets can be compared effectively in a timely fashion. In this example, the analyses and visualizations done using the R tools showed the importance of early detection in reducing outbreak impact.

Implementing this suite of R tools within the ADSM package will add value to modelers and decision makers by allowing rapid analysis of model results that can then be used to support emergency preparedness activities.

## REFERENCES

Allaire JJ, Joe Cheng, Yihui Xie, Jonathan McPherson, Winston Chang, Jeff Allen, Hadley Wickham, Aron Atkins, Rob Hyndman and Ruben Arslan. (2017). rmarkdown: Dynamic Documents for R. R package version 1.6. https://CRAN.R-project.org/package=rmarkdown

Baptiste Auguie (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra

Chang Winston, Joe Cheng, JJ Allaire, Yihui Xie and Jonathan McPherson (2017). shiny: Web Application Framework for R. R package version 1.0.5. https://CRAN.R-project.org/package=shiny

Müller Kirill, Hadley Wickham, David A. James and Seth Falcon (2017). RSQLite: 'SQLite' Interface for R. R package version 2.0.  https://CRAN.R-project.org/package=RSQLite

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL   https://www.R-project.org/.

RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL http://www.rstudio.com/.

R Special Interest Group on Databases (R-SIG-DB), Hadley Wickham and Kirill Müller (2017). DBI: R Database Interface. R package version 0.7.  https://CRAN.R-project.org/package=DBI

Thompson, D., Muriel, P., Russell, D., Osborne P., Bromley A., Rowland M., Creigh-Tyte S., and Brown C. 2002. Economic costs of the foot and mouth disease outbreak in the United Kingdom in 2001. Rev. sci. tech. Off. int. Epiz. 21 (3): 675-697.

USDA 2013. Parameters used to simulate the spread of FMD in Texas using the North American Animal Disease Spread Model (NAADSM) for use in FMD Response Workforce Requirement Estimates. April 2013. Available at: https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/sa-epidemiology-animalhealth-ceah/sa_monitoring_and_modelling/ct_monitoring_and_modelling

Wickham H.. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009.

Yihui Xie (2017). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.17.

Yihui Xie (2015) Dynamic Documents with R and knitr. 2nd edition. Chapman and Hall/CRC. ISBN 978-1498716963

Yihui Xie (2014) knitr: A Comprehensive Tool for Reproducible Research in R. In Victoria Stodden, Friedrich Leisch and Roger D. Peng, editors, Implementing Reproducible Computational Research. Chapman and Hall/CRC. ISBN 978-1466561595