Versioning, Stability, Verification, and Validation of *NAADSM*

The *NAADSM* Development Team
Author for correspondence:
Aaron Reeves
Animal Population Health Institute
College of Veterinary Medicine & Biomedical Sciences
Colorado State University
Fort Collins, Colorado 80523
Aaron.Reeves@colostate.edu

NAADSM Technical Paper #1
Version 2009/10/19
http://www.naadsm.org/techpapers/

Document revision history

- 2009/03/29 – Initial public version released
- 2009/10/19 – Published at *http://www.naadsm.org/techpapers*. URLs updated to reflect new file locations at *http://www.naadsm.org*.

Introduction

Epidemiologic simulation models might be judged for their utility in operational settings based on several criteria. These criteria include the availability of well documented releases of the model, stability of the conceptual design and computational implementations of the model, software testing and verification procedures carried out during model development, and efforts to validate the model.

This document describes the approaches used by the *North American Animal Disease Spread Model* (*NAADSM*) Development Team to provide a versioned, stable, verified platform for epidemiologic simulation modeling. This document also describes ongoing efforts of the *NAADSM* Development Team and associated researchers to address the issue of model validation.

The *NAADSM* model specification

All *NAADSM* development is based on a conceptual model, which is thoroughly described in the model specification document. The *NAADSM* model specification is intended to be a plain-language description of the simulation model implemented in *NAADSM*. Its purposes are to facilitate agreement among current members of the international *NAADSM* Development Team on details of the model, to provide a basis for functional testing, and to provide future team members and other users with a complete but accessible description of the model. All versions of the *NAADSM* model specification are published on the *NAADSM* website at *http://www.naadsm.org/documentation*.

When a new conceptual feature for the enhancement of the capabilities of *NAADSM* is proposed, a member of the Development Team submits a Request For Comments (RFC) to the entire team for review. An RFC is a formal document describing exactly what text is to be added to, removed from, or changed in the model specification. An RFC must give a justification for the changes and include any supporting documents, *e.g.*, manuscripts from the scientific literature. An RFC may be revised and clarified through discussions with the entire team. An RFC that is accepted after evaluation becomes part of the documentary history of the project; it is merged into the model specification, and the version number of the model specification is incremented (see below).

NAADSM versioning

The term “versioning”, as it is used with computer software, refers to the assignment of unique identifiers to specific, individual editions or unique releases of software applications or packages. A change in the version identifier indicates that a change has been made to the programming code of the application in question. These changes may or may not affect users of the software: some performance enhancements or changes made to improve the quality and maintainability of the code may go unnoticed by end users.

The *NAADSM* Development Team employs a three-part, sequence-based identification system for the model specification document and the *NAADSM* application. *NAADSM* version numbers are written as “*x.y.z*”, where *x*, *y*, and *z* are numeric values that identify the major version number, minor version number, and release number respectively.

For historical reasons, *NAADSM* version numbers start at 3.0.0. This scheme reflects continuity with *SpreadModel* (Shoenbaum and Disney 2003), the conceptual model on which *NAADSM* was originally based. Version numbers for the *NAADSM* specification document start at 1.0.0: the first version of the *NAADSM* specification document was written for *NAADSM* 3.0.0. As of this writing, the current version of *NAADSM* is 3.1.21, and the most recent publicly released version of the model specification document is 1.1.1. The correspondence between application and specification versions is described below.

Major version numbers (the first of the three parts of a *NAADSM* version number) are used to indicate substantial changes to the conceptual model. When the conceptual model is altered in such a way that it is rendered incompatible with previous versions, the major version number is incremented. For example, the *NAADSM* Development Team is currently implementing features that will make it possible to explicitly model surveillance and detection activities (*NAADSM* Development Team, unpublished data). When completed, these changes will be described in a specification document with a major version number greater than the current major version number of 1, and implemented in an application with a major version number greater 3.

Minor version numbers are incremented when changes are made to the conceptual model that do not affect existing features. These changes are typically new features that are fully backward-compatible with existing *NAADSM* models.

Release numbers are incremented as patches are made to the *NAADSM* application. To date, virtually all of these patches have been to the user interface only, and do not affect calculations performed by the model application. When an error is found, either in the user interface or in the model implementation, a description of the error is published on the *NAADSM* website (see *http://www.naadsm.org/bugs*). The error is corrected as quickly as possible, and a new release version is published.

Release numbers are occasionally incremented for new versions of the model specification document as well. These represent clarifications of, rather than changes to, an existing model specification.

There is not necessarily a one-to-one correspondence between model specification versions and application versions. All application versions are written to implement a specific version of the specification, but several application versions may correspond to a single specification version. The correspondence between *NAADSM* application versions and specification versions released to date is shown in Table 1. All released versions of *NAADSM* and the specification document are described on the *NAADSM* website as well: see *http://www.naadsm.documentation* and *http://www.naadsm.org/documentation/versions* for more information.

Table 1. Versions of the *NAADSM* specification and corresponding versions of the *NAADSM* application.

Major application version	***Specification version***	***Most recent corresponding application release***	***Older corresponding application releases***
NAADSM 3.1	Model specification 1.1 Latest revision: 1.1.1, released Sept. 18, 2008	*NAADSM* 3.1.21, released Jan. 13, 2009	3.1.20, 3.1.19, 3.1.18, 3.1.17, 3.1.16, 3.1.15
NAADSM 3.0	Model specification 1.0 Latest revision: 1.0.7, released Apr. 24, 2007	*NAADSM* 3.0.84, released Oct. 30, 2006	3.0.83, 3.0.82, 3.0.81, 3.0.80, 3.0.79

Adapted from the *NAADSM* website, *http://www.naadsm.org*

Verification of *NAADSM*

Verification of a simulation model refers to the process of determining whether the model, as implemented in software, conforms to the desired and specified conceptual model (Sargent 1999). In other words, verification provides an assessment of whether the software is working correctly and as intended. Verification of *NAADSM* has relied on three major steps: the creation of the model specification (described above), extensive automated testing, and detailed manual testing by an independent analyst.

As part of the build process for the *NAADSM* application, all code is subject to considerable automated testing. Simple model scenarios have been developed to test virtually every aspect of the *NAADSM* application. These tests are developed based on information contained in the model specification document. The suite of tests currently contains over 1000 individual tests, and new tests are continually developed and incorporated into the test suite. During compilation of the code, every test is automatically performed with the aid of a widely used, freely distributed framework for software testing (Savoye 2004). Every test in the suite must be passed before an updated version of *NAADSM* is released. Whenever a programming error is identified, at least one test is added to the test suite: the test or tests must fail in the presence of the error, and pass once the error has been fixed. Every test is published along with the complete source code for the *NAADSM* application: see *http://www.naadsm.org/source*.

Prior to the initial public release of *NAADSM*, every aspect of the model framework was examined in detail by an analyst who worked essentially independently of the *NAADSM* Development Team. More sophisticated tests than those in the automated test suite were developed, again based on information contained in the model specification. Any questions raised or inconsistencies identified during this process had to be explained or addressed before release of the model. New components of *NAADSM* are subjected to similar manual testing on an ongoing basis.

Stability of *NAADSM*

"Stability" might refer to either the relative permanence of the conceptual *NAADSM* model or the reliability of the *NAADSM* application. Both aspects are addressed below.

Stability of the *NAADSM* conceptual model

For operational purposes, it may not practical to issue multiple substantially different versions of a model over a short duration. It frequently takes time for modelers and analysts to become familiar with the nuances of complex simulation models. If dramatic changes are made frequently, it can be quite difficult for modelers to maintain a detailed understanding of all of the mechanisms that a model attempts to represent.

Version 1.0 of the *NAADSM* model specification document was initially published in 2006. Although significant additional capabilities have been introduced since that time, all fundamental aspects of the model have remained unchanged. Special care has been taken to ensure that every new release of *NAADSM* in the version 3 series maintains full backward compatibility with prior versions: *NAADSM* 3.1.21 (released in 2009) can be used to run scenarios developed with the very first public version of *NAADSM* (released in 2006) without modification, and will produce virtually identical output values.

Substantial changes to the conceptual model are now planned for versions of *NAADSM* tentatively scheduled for release in summer 2009 (Dubé *et al.* 2008). These changes represent a marked improvement over existing *NAADSM* capabilities, but they will necessitate the discontinuation of backward compatibility with *NAADSM* 3. Although new development is now concentrated on these forthcoming major versions, there are currently no plans to discontinue support for *NAADSM* 3. It is anticipated that version 3 will continue to be used in operational settings alongside newer versions for some time to come.

Stability of the *NAADSM* application

NAADSM has been released under an open source software license (Free Software Foundation 2007): the application is freely available to any interested user. This characteristic has fostered the establishment of a considerable international community of users (Table 2). Consequently, stability and reliability of the application are high development priorities.

NAADSM has been subjected to considerable testing for stability as well as verification. An additional objective of the independent manual testing described above was to provide a thorough test of the *NAADSM* user interface under operational conditions.

NAADSM has been used extensively in training courses by users with a wide variety of backgrounds: 196 participants have attended eight such courses since 2007. These courses are *de facto* testing sessions for the *NAADSM* application. Although minor problems with the application (nearly always in the user interface) are occasionally identified, the program has proven to be quite stable and largely free of programming errors. Errors that do affect model accuracy and results are documented on the *NAADSM* website so that modelers may assess whether their results may have been affected. Users of the model are also automatically notified through the *NAADSM* user interface about such errors.

Further evidence of the stability of the *NAADSM* application is its daily use in an operational setting by analysts from USDA Center for Emerging Issues, USDA National Surveillance Unit, and the Canadian Food Inspection Agency, as well as personnel from various other government agencies and universities. All *NAADSM* users are actively encouraged to report any problems that they encounter with the application to members of the *NAADSM* Development Team. Contact information for the developers is openly published on the *NAADSM* website and in the *NAADSM* user's guide (Hill and Reeves 2006).

Table 2. Registered and/or trained *NAADSM* users[1] since 2007, by country

Country	*Number of users*
USA	168
Canada	48
Brazil	11
Mexico	7
Chile	6
Uruguay	4
Australia	3
China	3
Ireland	3
Argentina	2
Colombia	2
Finland	2
New Zealand	2
Nigeria	2
Panama	2
Russian Federation	2
Uganda	2
Others	13
Total	*282*

[1] Neither user registration nor training is required. These values may under-represent the total number of *NAADSM* users.

Validation of *NAADSM*

In the context of simulation modeling, the process of "validation" has been defined as "determining whether the simulation model is an acceptable representation of the real system – given the purpose of the simulation model" (Sargent 1999). Although variety of criteria have been suggested to assess the validity of epidemiologic models (Goodall 1972, Miller 1976, Rykiel 1996, Taylor 2003), many of these criteria are problematic, especially when models are to be used in countries without recent experience with the diseases of interest and without relevant data (Reeves 2008).

The issue of validation is further complicated when the variety of models that can be constructed with an application like *NAADSM* is considered. It should be emphasized that *NAADSM* is not a single model: it is a flexible framework for the development of epidemiologic simulation models. It is possible to use *NAADSM* to construct models that are wholly inadequate and demonstrably false. Thus, the validity of a *NAADSM*-based model depends far more on the user of the framework than on the framework itself. Nonetheless, the *NAADSM* Development Team has undertaken several major initiatives to establish confidence in the *NAADSM* framework when it is used appropriately.

It should also be emphasized that model validation is an ongoing, iterative process. As we strive to improve our capabilities for the future, it is important to critically review the assumptions made in existing versions of *NAADSM*. Members of the *NAADSM* Development Team are continually seeking out new collaborations and sources of information with which to evaluate our work (*e.g.*, Dubé *et al*. 2004, Reeves *et al*. 2006, Dubé *et al*. 2008, Dubé *et al*. submitted for publication).

Conceptual validity

A particularly useful criterion for the validation of a model is the answer to the question "does the structure of a model make logical and biological sense?" This issue has been referred to as "conceptual validity" or "face validity". Review by independent subject matter experts can be

used to assess whether the structure of a model sufficiently captures the relationships and interactions among components of the epidemiologic system being modeled. *NAADSM* has been subject to formal review by subject matter experts on three separate occasions (Table 3; USDA Centers for Epidemiology and Animal Health 2002, 2004, Dubé *et al.* 2008). To the best of our knowledge, no other simulation model for animal diseases has undergone this rigorous conceptual validation process.

Table 3. Subject matter expert meetings conducted for review and evaluation of *NAADSM* and its precursors

Dates	Location	Number of participants	Countries represented
July 9 – 11, 2002[1]	Fort Collins, Colorado	41	Australia, Canada, Mexico, The Netherlands, the United States
June 15 – 17, 2004	Fort Collins, Colorado	22	Australia, Canada, New Zealand, the United States, Uruguay
March 30 – April 4, 2008	Rio de Janiero, Brazil	21	Argentina, Brazil, Canada, Chile, Colombia, Paraguay, the United States, Uruguay

[1] This meeting was held prior to the organization of the international *NAADSM* Development Team. One of the models reviewed was *SpreadModel* (Schoenbaum and Disney 2003), the conceptual model on which initial *NAADSM* development was based.

Suggestions from subject matter expert meetings are used by the *NAADSM* Development Team to guide future development: suggestions from the 2002 meeting (USDA Centers for Epidemiology and Animal Health 2002) led directly to capabilities now available in *NAADSM* 3, and information collected during the 2008 meeting (Dubé *et al.* 2008) is now being used to guide development of future versions of *NAADSM*.

Furthermore, a complete description of the mechanisms included in *NAADSM* has been published for review by the entire scientific and regulatory communities (Harvey *et al*. 2007). The *NAADSM* Development Team invites comment and criticism from these communities in order to improve our future capabilities.

Relative validity

Comparison of several independently developed models may be used to improve the level of confidence in the models tested: general agreement among tested models may lend credibility to conclusions or predictions from model-based studies. This process has been called “relative validation” (Dubé *et al*. 2007). *NAADSM* is one of three simulation models recently evaluated in a model comparison of relatively simple disease scenarios. Among the findings of this comparison was that, although statistically significant differences were observed among model outputs, results from all three models supported the same or very similar conclusions regarding approaches for disease control. This finding could be used to increase the confidence of end users and decision makers in modeling results (Dubé *et al*. 2007). Efforts are now underway to compare model output from more complex scenarios (N. Harvey, unpublished data).

Validation of components of *NAADSM*

Although it is quite difficult to demonstrate the validity of an entire complex simulation model, especially in the absence of relevant data, it may be possible to assess the validity of some individual components of a complex model. An example is a recently completed validation of the process used in *NAADSM* to simulate animal movements and contacts among farm premises (Dubé *et al*. submitted for publication).

Briefly, the objective of Dubé *et al.* (submitted for publication) was to validate this *NAADSM* contact model by comparing simulated movements to real-world, farm-to-farm movements that had been recorded for adult milking cows in Ontario, Canada. Simulations were performed using

two approaches to represent the real-world monthly and yearly networks of adult dairy cow movements in Ontario. The study concluded that the approach used in *NAADSM* performed reasonably well in simulating average network characteristics observed in real-world movement data, but did not perform as well in simulating extreme upper percentiles of movement network components, involving rare but observed farms with excessively high shipment frequencies. The results of this study will be used to inform future development, with the objective of providing better representations of actual events and greater confidence in the results of modeling studies.

Ongoing efforts to validate *NAADSM* against historical data

When available, data from historic outbreaks can be used to develop parameters for and to evaluate epidemiologic models. As these historic outbreaks represent unique occurrences, such efforts are not definitive and should be undertaken with care (Reeves 2008), but the exercise can still be of considerable value. Retrospective analysis of past outbreaks is critical to understanding them, and modeling can be a very useful tool in this pursuit (Kitching *et al*. 2006, Garner *et al.* 2007).

The 2001 outbreak of Foot-and-Mouth Disease (FMD) in the United Kingdom provides an exceptional source of data that can enable modeling groups to retrospectively examine a real FMD outbreak. Data collected during the outbreak was obtained from the Department for the Environment, Food, and Rural Affairs in the UK, with the intention of validating *NAADSM*. The goal of this validation was not to determine whether *NAADSM* could have predicted the spatial spread of the epidemic, nor to determine whether it could have been used during the outbreak to make predictions on the best control measures to use. The purpose instead was to determine whether *NAADSM* could recreate the epidemic, or some parts of the epidemic, with what is now known about the spread mechanisms that existed and the control measures that were applied. A specific region of the UK, North Cumbria, was identified as the location of an important cluster of diseased premises, and was selected for detailed study. Results of the investigation showed that *NAADSM* well represented the actual outbreak that occurred in North Cumbria (Dubé *et al*. 2004).

Similar investigations are now under way for an outbreak of FMD that occurred in Uruguay in 2001 (preliminary results were reported by Reeves *et al*. 2006) and for an outbreak of Exotic Newcastle Disease that occurred in 2002-2003 in California (E. Monaco, K. Forde-Folle, and A. Reeves, unpublished data).

Conclusions

The *NAADSM* Development Team has gone to great lengths to ensure that we have provided a sound, stable model for use by epidemiologists, veterinarians, researchers, and other workers in the fields of animal health and preparedness planning. *NAADSM* is now being used in operational settings to address specific policy and research questions (*e.g.*, Pendell *et al.* 2007, Portacci *et al.* 2007, Paarlberg *et al.* 2008, Sanderson *et al.* 2009, USDA Centers for Epidemiology and Animal Health 2009a,b). *NAADSM* has been subjected to detailed testing for the purpose of model verification. As development of this tool continues, verification and validation efforts will continue, as will efforts to improve the quality and credibility of the modeling framework.

References

Dubé, C., Corso, B., Cartwright, C. 2004. Validation of SpreadModel/SharcSpread using the 2001 UK FMD outbreak data. Ottawa, Ontario, Canada: Canadian Food Inspection Agency.

Dubé, C., Geale, D., Sanchez, J., 2008. *NAADSM* orientation workshop and project plan for pilot studies, sofware development, and oversight of *NAADSM* application in South America. Ottawa, Ontario, Canada: Canadian Food Inspection Agency.

Dubé, C., Ribble, C., McNab, B., submitted for publication. An assessment of the assumptions used in the *North American Animal Disease Spread Model* (*NAADSM*) to represent real-world livestock movements. *Proceedings of the 12th International Symposium on Veterinary Epidemiology and Economics.*

Dubé, C., Stevenson, M.A., Garner, M.G., Sanson, R.L., Corso, B.A., Harvey, N., Griffin, J., Wilesmith, J.W., Estrada, C., 2007. A comparison of predictions made by three simulation models of foot-and-mouth disease. *New Zealand Veterinary Journal* 55: 280-288.

Free Software Foundation, 2007. GNU General Public License. Web page *http://www.fsf.org/licensing/licenses/gpl.html*. Last accessed March 28, 2009.

Garner, M.G., Dubé, C., Stevenson, M.A., Sanson, R.L., Estrada, C., Griffin, J., 2007. Evaluating alternative approaches to managing animal disease outbreaks - the role of modelling in policy formulation. *Veterinaria Italiana* 43: 285-298.

Goodall, D.W., 1972. Building and testing ecosystem models. In: Jeffers, J.N.J. (Ed.), Mathematical Models in Ecology. Oxford: Blackwell, pp. 173-194.

Harvey, N., Reeves, A., Schoenbaum, M.A., Zagmutt-Vergara, F.J., Dubé, C., Hill, A.E., Corso, B.A., McNab, W.B., Cartwright, C.I., Salman, M.D., 2007. The *North American Animal Disease Spread Model*: A simulation model to assist decision making in evaluating animal disease incursions. *Preventive Veterinary Medicine* 82: 176-197.

Hill, A., and Reeves, A. 2006. User's Guide for the *North American Animal Disease Spread Model*, 2nd ed. Fort Collins, Colorado: Animal Population Health Institute, Colorado State University. Available online at *http://wwww.naadsm.org*. Last accessed March 23, 2009.

Kitching, R.P., Thrusfield, M.V., Taylor, N.M., 2006. Use and abuse of mathematical models: an illustration from the 2001 foot and mouth disease epidemic in the United Kingdom. *Revue Scientifique et Technique - Office International des Epizooties* 25: 293-311.

Miller, W.M., 1976. A state-transition model of epidemic foot-and-mouth disease. *New techniques in veterinary epidemiology and economics (ISVEE I)* 1: 56-72.

Paarlberg, P.L., Seitzinger, A.H., Lee, J.G., Mathews, K., 2008. Economic Impacts of Foreign Animal Disease. US Department of Agriculture Economic Research Service, publication number ERR-57.

Pendell, D.L., Leatherman, J., Schroeder, T.C., Alward, G.S., 2007. The economic impacts of a Foot-and-Mouth Disease outbreak: A regional analysis. *Journal of Agricultural and Resource Economics* 39: 19-33.

Portacci, K.A., Reeves, A., Corso, B.A., Salman, M.D., 2007. Recommendations for the National Veterinary Stockpile: Pseudorabies virus in commercial swine. Fort Collins, Colorado: US Department of Agriculture Centers for Epidemiology and Animal Health.

Reeves, A. 2008. Verification and validation of epidemiologic models. *World Organization for Animal Health (OIE) Epidemiological Modeling Workshop*, August 11 – 13 2008, Fort Collins, Colorado.

Reeves, A., Gil, A.D., Zagmutt-Vergara, F., Hill, A.E., Corso, B.A., Salman, M.D., 2006. Validation of the *North American Animal Disease Spread Model* using data from the 2001 outbreak of Foot-and-Mouth Disease in Uruguay. In: Ellis, R.P. (Ed.), Proceedings of the 87th Annual Meeting of the Conference of Research Workers in Animal Diseases. Ames, Iowa: Blackwell Publishing.

Rykiel, E.J., 1996. Testing ecological models: the meaning of validation. *Ecological Modelling* 90: 229-244.

Sanderson, M.W., Forde-Folle, K., Reeves, A., 2009. Effect of movement controls and biosecurity on transmission of disease by indirect contact and the role of vaccination in the control of foot and mouth disease in livestock production systems in the central United States. *Proceedings of the 12th International Symposium on Veterinary Epidemiology and Economics*. Accepted for publication.

Sargent, R.G., 1999. Validation and verfication of simulation models. *Proceedings of the 1999 Winter Simulation Conference* 39-48.

Savoye, R., 2004. DejaGnu: The GNU Testing Framework. Free Software Foundation. Available online at *http://www.gnu.org/software/dejagnu/manual/*. Last accessed March 29, 2009.

Schoenbaum, M.A., Disney, W.T., 2003. Modeling alternative mitigation strategies for a hypothetical outbreak of foot-and-mouth disease in the United States. *Preventive Veterinary Medicine* 58: 25-52.

Taylor, N., 2003. Review of the use of models in informing disease control policy development and adjustment. UK Department of Environment, Food, and Rural Affairs. Available online at *http://www.defra.gov.uk/science/documents/publications/2003/UseofModelsinDiseaseControlPolicy.pdf*. Last accessed March 29, 2009.

USDA Centers for Epidemiology and Animal Health, 2002. North American Animal Health Committee Conference. Fort Collins, Colorado: US Department of Agriculture Centers for Epidemiology and Animal Health.

USDA Centers for Epidemiology and Animal Health, 2004. Validation of *SpreadModel*: Minutes of the subject matter expert team meeting. Fort Collins, Colorado: US Department of Agriculture Centers for Epidemiology and Animal Health.

USDA Centers for Epidemiology and Animal Health, 2009a. Preliminary estimation of biological consequences of a Classical Swine Fever outbreak in the state of Iowa using the *North American Animal Disease Spread Model*. Fort Collins, Colorado: USDA Centers for Epidemiology and Animal Health.

USDA Centers for Epidemiology and Animal Health, 2009b. Vaccine bank requirements for Foot-and-Mouth Disease in southwest Kansas: A preliminary report. Fort Collins, Colorado: USDA Centers for Epidemiology and Animal Health.