

УДК 004.7

А. Я. Куземин, Д. В. Фастова, О. Н. Дяченко

Харьковский национальный университет радиоэлектроники
проспект Ленина, 14, 61166 Харьков, Украина

Моделирование процесса образования лавин

Рассмотрены актуальные вопросы прогнозирования техногенных катастроф, в частности снежных лавин. Первоочередной задачей для получения качественного прогноза является построение модели образования лавин. Проанализированы существующие методы моделирования, выявлены их недостатки и предложен метод группового учета аргументов для моделирования процессов лавинообразования.

Ключевые слова: лавинная опасность, процесс лавинообразования, прогнозирование, математическая модель, задача идентификации, дедуктивные методы, индуктивные методы, принцип самоорганизации модели, внешнее дополнение

Введение

Снежные лавины, как и землетрясения, наводнения, извержения вулканов, относят к стихийным природным явлениям, которые способны вызвать гибель людей и причинить значительные разрушения. Несмотря на то, что лавиноопасные районы занимают 6 % площади суши, проблема исследования подобных явлений является важной для своевременного проведения противолавинных мероприятий. К ним относят следующие мероприятия: выявление лавиноопасных территорий и определение параметров явлений, организация службы прогнозирования времени схода лавин, строительство защитных сооружений, предупредительный спуск лавин [1].

Практическую реализацию перечисленных выше мероприятий осуществляют, опираясь на результаты, полученные с использованием технологии геоинформационных систем (ГИС). ГИС-технологии позволяют моделировать процессы и явления, определяющие условия схода снежных лавин, выявлять зоны аккумуляции, сноса и динамики снега.

Целью данной статьи является анализ существующих методов моделирования и выбора приемлемого метода для моделирования процесса образования лавин.

Процесс моделирования лавинообразования как задача идентификации

Процесс образования лавины является сложной системой. Моделирование является одним из наиболее мощных средств исследования подобных систем на сегодняшний день.

Моделью исходного объекта называется представление объекта в некоторой форме, отличной от формы его реального существования. В настоящее время известны и широко используются в научных исследованиях и инженерной практике различные типы моделей и многочисленные методы моделирования [2]. Если взять за основу степень абстрактности (степень отличия от реального объекта), то можно определить следующие типы моделей:

— физические (натурные) — воспроизводят изучаемый процесс с сохранением его физической природы и являются инструментом физического моделирования;

— аналоговые — заменяют один объект на другой с похожими свойствами;

— математические — абстрактные модели, которые существуют в форме специальных математических конструкций и имеют смысл только для интересующего их человека или машины.

Процесс образования лавин является уникальным, состоящим из большого количества взаимосвязанных между собой параметров, что делает невозможным проведение исследований на реальном объекте. Таким образом, построение процесса лавинообразования может быть выполнено при помощи математического моделирования. Под математическим моделированием понимают способ исследования различных процессов путем изучения явлений, имеющих различное физическое содержание, но описываемых одинаковыми математическими соотношениями.

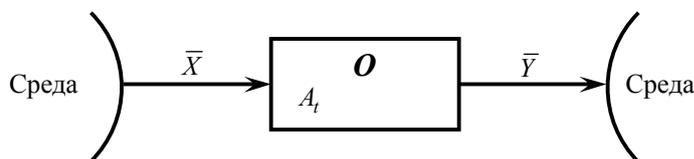
Моделирование процесса образования лавины можно отнести к задачам идентификации (задачам краткосрочного прогнозирования [3]). Задача идентификации формируется следующим образом: по результатам наблюдений за входными и выходными переменными системы строится оптимальная в некотором смысле ее модель [4].

На рисунке проиллюстрировано взаимодействие идентифицируемого объекта O со средой. Это взаимодействие происходит по каналам наблюдения входа $\bar{X} = \{X_1, \dots, X_n\}$ и выхода $\bar{Y} = \{Y_1, \dots, Y_n\}$ в дискретные моменты времени t_1, \dots, t_n . Эти наблюдения связаны неизвестным оператором объекта A_i , т.е. $Y_i = A_i(X_i)$, $i = \overline{1, n}$. По каналу \bar{X} среда взаимодействует на объект, а по каналу \bar{Y} объект воздействует на среду. Задача идентификации сводится к определению оператора модели A_i^* , т.е. в получении оценки A_i по наблюдениям X_i и Y_i в дискретный момент времени t_i .

Таким образом, задачу идентификации можно декомпозировать на следующие подзадачи:

- 1) выбор входных переменных (факторов или аргументов);
- 2) выбор выходных переменных (параметров оптимизации);

- 3) выбор структуры модели;
- 4) исследование алгоритмов использования модели для нахождения экстремумов выходных величин.



Взаимодействие объекта O со средой

Моделирование процессов с помощью регрессионного анализа

Теория моделирования при определении коэффициентов статических моделей предлагает находить парные и множественные связи переменных входов и выходов. Подобную задачу чаще всего решают с помощью регрессионного анализа, который заключается в минимизации функции ошибки (разность? между моделью и любой экспериментальной точкой) [4]. Как правило, функция ошибки представлена в виде суммы квадратов ошибок.

Применение регрессионного анализа сводится к проведению следующих этапов [5].

1. *Корреляционный анализ.* Коэффициент парной корреляции переменных y_1 и y_2 :

$$r = \frac{\sum_{i=1}^N (y_1 - \bar{y}_1)(y_2 - \bar{y}_2)_i}{\sqrt{\sum_{i=1}^N (y_1 - \bar{y}_1)^2 \sum_{i=1}^N (y_2 - \bar{y}_2)_i^2}},$$

где

$$\bar{y}_1 = \frac{1}{N} \sum_{i=1}^N y_{1i}, \quad \bar{y}_2 = \frac{1}{N} \sum_{i=1}^N y_{2i}.$$

Пороговое значение $r \leq 0,6$. При двух сильно корреляционных переменных из числа аргументов одна исключается (по указанию автора(ов) модели).

2. *Дисперсионный анализ.* Вклад каждой переменной (фактора) оценивается по его влиянию на среднеквадратическую оценку (дисперсию), рассчитанную на всех точках таблицы:

$$\delta^2 = \frac{\sum_{i=1}^N (y - y_m)_i^2}{N - n - 1},$$

где $(N - n - 1)$ — число степеней свободы; N — число точек таблицы; n — число переменных (единица связана с использованием в модели свободного члена a_0).

Переменная, исключение которой вызывает достаточно малое изменение дисперсии, считается не существенной и исключается из рассмотрения. Например, $\Delta\delta^2 \leq 0,1$.

3. *Факторный анализ.* Вклад каждой переменной оценивается по ее влиянию на среднее значение выходной величины.

Рассмотрим некоторый фактор, участвующий в модели:

$$y = \dots + b_i x_i + \dots$$

При наличии и отсутствии данного фактора определяется значение выходной величины для всех точек таблицы. Далее определяется вызванная исключением данного фактора дисперсия:

$$S_{bi}^2 = \frac{1}{N} S_y^2 = \frac{1}{N} \sum_1^N (y_{\text{без } x_i} - y_{\text{с } x_i})^2.$$

Ориентировочно (при $t = 2$) $\Delta b_i = \pm 2S b_i$, т.е. оценка коэффициентов b_i не должна попасть в интервал (лежащий вокруг нуля) от $-\Delta b_i$ до $+\Delta b_i$; члены с малыми коэффициентами исключаются.

4. *Анализ показателей существенности переменных.* Значимость факторов определяется чаще всего по критерию существенности Стьюдента.

Чтобы как-то обосновать (т.е. сделать объективным) выбор пороговых значений всех критериев, используются таблицы доверительных интервалов: рассмотренные выше показатели существенности факторов и структур должны отличаться от нуля меньше, чем на двухсигмовый интервал. Применение доверительных интервалов не уменьшает субъективного характера решения задач.

К основным недостаткам регрессионного анализа, как и к другим детерминированным методам можно отнести следующие:

1) синтез моделей осуществляется на достаточно полной информации о всех компонентах сложного процесса (необходимость достаточно большой выборки данных);

2) модель строится только в области, где число коэффициентов модели равно числу точек таблицы опытных данных или меньше его;

3) для оценивания моделей применяется критерий минимума среднеквадратической ошибки, рассчитанный по всем точкам таблицы опытных данных. Он является внутренним критерием, поскольку его определение основано на тех же эмпирических данных, которые используются и для получения оценок коэффициентов модели. Данный критерий не пригоден для нахождения модели оптимальной сложности из-за слабой помехоустойчивости внутренних критериев [6].

4) субъективный характер моделирования.

Моделирование с помощью методов группового учета аргументов

Процесс образования лавины относится к моделям, в которых неизвестна структура и механизм взаимодействия элементов наблюдаемых параметров. В таком случае подход самоорганизации, предложенный А.Г. Ивахненко, является единственным надежным средством моделирования таких процессов [7].

В основе алгоритмов самоорганизации математических моделей лежат следующие принципы.

1. *Принцип самоорганизации модели* можно сформулировать следующим образом: при постепенном усложнении модели значение внутренних критериев монотонно падает, а все внешние критерии, имеющие свойства «внешнего дополнения», проходят через свой минимум. При переборе большого числа моделей-претендентов находится минимум внешнего критерия, а, следовательно, единственная модель оптимальной сложности.

2. *Принцип внешнего дополнения.* Под внешним критерием будем понимать критерий, вычисляемый на информации, неиспользованной при оценке параметров. Этот принцип был сформулирован С. Биром [3], для преодоления следствия теоремы неполноты Геделя, которая утверждает, что для любой системы всегда можно задать такую теорему, для доказательства которой недостаточно данной системы аксиом, и требуется новая аксиома — внешнее дополнение. В применении к самоорганизации моделей идеи Геделя можно интерпретировать так: по минимуму заданного внешнего критерия можно решить все вопросы о выборе опорных функций, структуре и параметрах модели.

3. *Принцип сохранения свободы выбора.* Свобода выбора обеспечивается тем, что на каждый следующий ряд селекции передается не одно решение, а несколько лучших, отобранных в предыдущем ряду.

К методам, основанным на принципе самоорганизации, относится метод группового учета аргументов (МГУА), с помощью которого будет реализовано моделирование процесса образования лавин.

МГУА как воплощение индуктивного подхода — это оригинальный метод построения моделей по экспериментальным данным в условиях неопределенности. Полученные по этому методу модели оптимальной сложности отображают неизвестные закономерности функционирования исследуемого объекта (процесса), информация о которых неявно содержится в выборке данных. В МГУА для построения моделей применяются принципы автоматической генерации вариантов, неокончательных решений и последовательной селекции лучших моделей по внешним критериям. Такие критерии основаны на делении выборки на части, при этом оценивание параметров и проверка качества моделей выполняются на разных частях исходной выборки. Это разрешает обойтись без отягощающих априорных предположений, поскольку деление выборки разрешает автоматически (неявно) учесть разные виды априорной неопределенности.

Основная структура алгоритма МГУА состоит из следующих блоков [8]:

- 1) предварительной обработки наблюдений (выборка данных);
- 2) выбора вида опорных функций;
- 3) генерации множества моделей-претендентов;
- 4) вычисления критериев селекции, являющихся внешним дополнением;

5) отбора моделей оптимальной сложности.

В алгоритмах МГУА применяются следующие опорные функции [5].

1. Степенные полиномы

$$\varphi_i(t) = t^i$$

для построения модели

$$y(t) = \sum_{i=0}^n c_i t^i .$$

2. Тригонометрические функции:

$$\varphi_{2i}(t) = \cos \omega_i t , \quad i = 0, 1, \dots, m ,$$

$$\varphi_{2i-1}(t) = \sin \omega_i t$$

для образования модели

$$y(t) = \sum_{i=0}^m c_i \cos \omega_i t + \sum_{i=0}^m d_i \sin \omega_i t .$$

3. Экспоненциальные функции

$$\varphi_i(t) = e^{\beta_i t}$$

для построения модели

$$y(t) = \sum_{i=0}^m c_i e^{\beta_i t} .$$

Графическое представление процесса позволяет сделать правильный выбор опорных функций. Для стационарных колебаний со средним значением, тождественно равным нулю, рекомендуется применять тригонометрические функции, для монотонного процесса — полиномиальную функцию. (или множ. число?)

В качестве целесообразных внешних дополнений в теории самоорганизации моделей используются следующие критерии.

1. *Критерий регулярности* — среднеквадратичная ошибка, рассчитанная на новых точках, неиспользованная для получения оценок коэффициентов модели:

$$\Delta^2(B) = \frac{\sum_{i=1}^{N_B} (q_{табл} - q_m)_i^2}{\sum_{i=1}^{N_B} q_{табл\ i}^2} \rightarrow \min ,$$

где N_B — число точек отдельной проверочной последовательности данных $N = N_A + N_B$; $q_{табл}$ — табличное значение выходной переменной; q_m — значения, рассчитанные по данной модели.

Критерий регулярности основан на разделении имеющихся экспериментальных данных на две части: обучающую N_A и проверочную N_B последовательности точек.

2. *Критерий минимума смещения* требует максимального совпадения значений выходной величины двух моделей, полученных на двух различных частях таблицы исходных данных:

$$n_{см}^2 = \frac{\sum_{i=1}^{\alpha N} (q_A - q_B)^2}{\alpha \sum_{i=1}^N q_{i табл}^2} \rightarrow \min ,$$

где A — точки с большим значением дисперсии выходной величины; B — точки с меньшим ее значением; α — коэффициент экстраполяции, $\alpha = 1,5 \div 3,0$; N — все точки таблицы исходных данных.

Критерий минимума смещения позволяет выбрать модель, наименее чувствительную к изменению множества опытных точек, по которым она получена. Этот критерий позволяет решать задачу восстановления закона, скрытого в зашумленных экспериментальных данных, а потому рекомендуется для решения задачи идентификации.

3. *Критерий баланса переменных*. Восстановление функции по ряду заданных опытных точек сводится к решению двух различных между собой задач: аппроксимации и экстраполяции. В задаче аппроксимации требуется восстановить функцию по опытным точкам на интервале интерполяции, т.е. там, где заданы точки. В задаче экстраполяции необходимо найти продолжение функции в области, где уже нет опытных точек.

Выбор метода группового учета аргументов для моделирования процесса образования лавин

Задача идентификации процесса образования лавин состоит в получении модели с наилучшими прогнозирующими свойствами. В связи с этим из всех подходов структурной идентификации наибольший интерес представляет подход «перекрестного подтверждения». Применение МГУА для автоматического поиска лучшей модели является более эффективным по сравнению с другими методами.

Достоинствами данного метода являются следующие.

1. Объективный характер моделирования алгоритмов самоорганизации. Эффективность достигается тем, что диалог человека (автора(ов) модели) и машины ведется не на языке детальных указаний, как при имитационных методах моделирования, а на языке постановки критериев общего вида: критерия регулярности, минимума смещения, баланса переменных и многих других критериев. В спорных

случаях стороны договариваются о критериях и больше в процесс моделирования не вмешиваются, что и обеспечивает объективность выбора модели [4].

2. Сходимость к модели оптимальной сложности. Алгоритмы МГУА, применяемые для системного анализа взаимодействий в системе, позволяют найти систему уравнений оптимальной сложности. Алгоритмы для идентификации дают единственную модель оптимальной сложности.

3. Небольшой объем эмпирической информации. Модели сложных компонентов системы могут быть получены по весьма малым выборкам данных наблюдений, таким образом, МГУА направлен на всемерное уменьшение необходимой априорной информации.

4. МГУА позволяют строить модели в области, где число коэффициентов модели намного больше числа точек таблицы опытных данных.

5. Относительная скорость вычислений. В практическом использовании МГУА расчет прогнозов по полученной модели в реальном масштабе времени может быть получен относительно быстро.

6. Стоимость построения модели. Издержки при реализации МГУА значительно меньше, чем при имитационном моделировании.

Выводы

В статье изложены основные этапы моделирования сложных процессов, проведен сравнительный анализ методов, выявлены их недостатки. Задача моделирования процесса образования лавин отнесена к задаче идентификации. Рассмотренные дедуктивные методы моделирования непригодны для решения данной задачи. Выявленные недостатки дедуктивных методов можно устранить с помощью применения индуктивных методов самоорганизации систем.

В статье предложен метод группового учета аргументов для моделирования процесса лавинообразования. МГУА является комбинированным методом, т.к. используется априорная достоверная информация об объекте, осуществляются индуктивный перебор и оценивание моделей претендентов. Оценки моделей находятся по небольшой части исходных опытных данных. Остальные данные используются для получения оценок коэффициентов и определения оптимальной структуры модели.

Построение модели лавинообразования необходимо для оценки и прогнозирования лавинной опасности, что делает возможным проведение организационно-хозяйственных мероприятий для минимизации возможного ущерба и предотвращения человеческих жертв.

1. *Войтковский К.Ф.* Лавиноведение: учебное пособие. — М.: МГУ, 1989. — 158 с.

2. *Скурихин В.И., Шифрин В.Б., Дубровский В.В.* Математическое моделирование. — К.: Техніка, 1983. — 270 с.

3. *Ивахненко А.Г., Мюллер Й.А.* Самоорганизация прогнозирующих моделей. — К.: Техніка, 1985; Берлин: ФЭБ Ферлаг Техник, 1984. — 223 с.

4. *Молчанов А.А.* Моделирование и проектирование сложных систем. — К.: Выща школа, 1988. — 359 с.
5. *Ивахненко А.Г.* Индуктивный метод самоорганизации моделей сложных систем. — К.: Наук. думка, 1982. — 296 с.
6. *Ивахненко А.Г.* Долгосрочное прогнозирование и управление сложными системами. — К.: Техніка, 1975. — 312 с.
7. *Ивахненко А.Г., Юрачковский Ю.П.* Моделирование сложных систем по экспериментальным данным. — М.: Радио и связь, 1987. — 120 с.
8. *Ивахненко А.Г., Степашко В.С.* Помехоустойчивое моделирование. — К.: Наук. думка, 1985. — 300 с.

Поступила в редакцию 27.11.2006