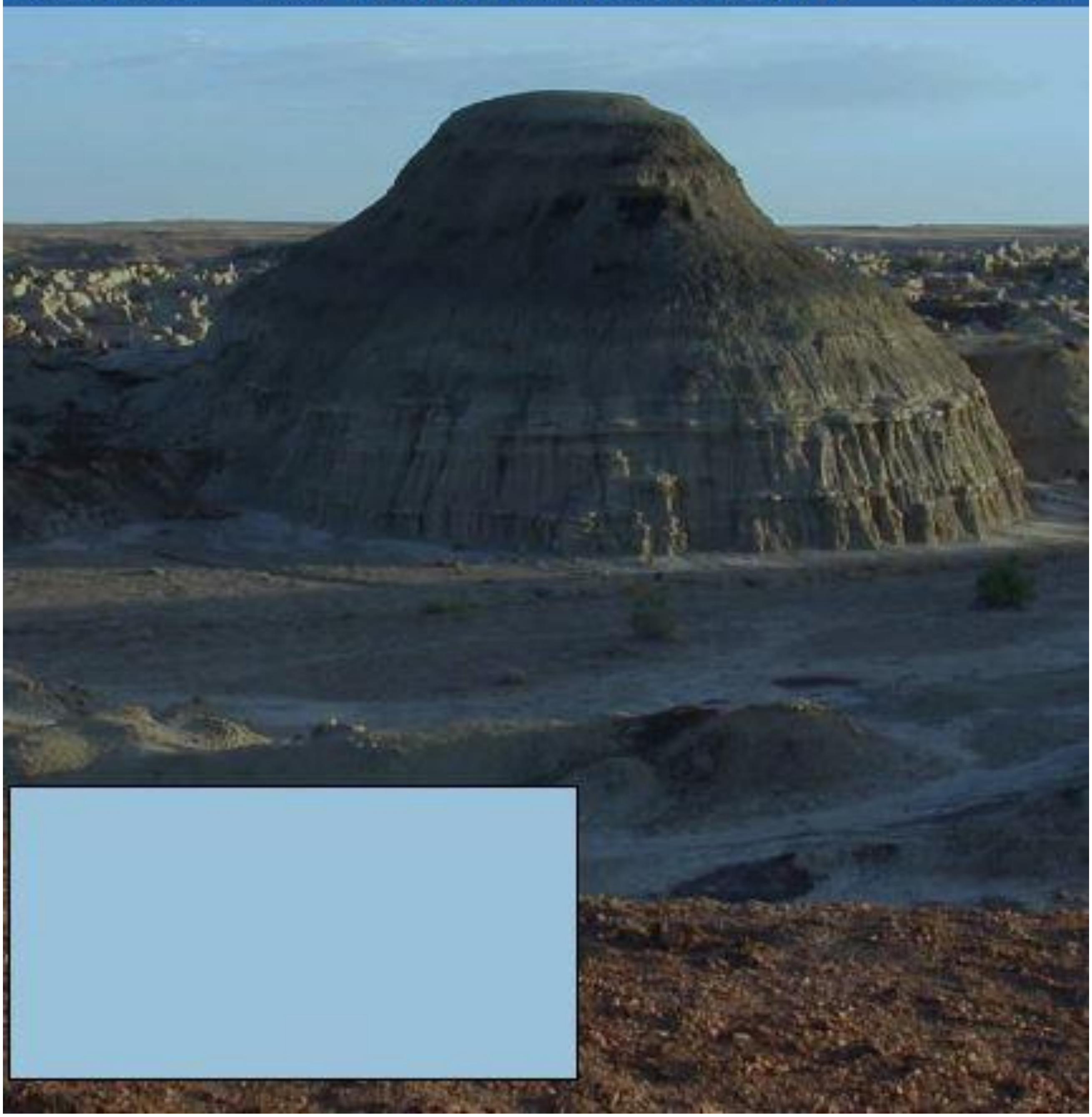
TPG

Volume 53, Number 1 THE PROFESSIONAL GEOLOGIST Jan.Feb.Mar 2016



TPG

Volume 53 Number 1 Jan.Feb.Mar 2016

INSIDE THIS STUDENT ISSUE

DEPARTMENTS

Bisti/De-Na-Zin Wilderness located in San Juan County, New Mexico. The Wilderness lies in the southwest portion of the San Juan Basin and consists of sandstones, mudstones, shales, and coal of the Upper Cretaceous Fruitland Formation and Kirtland Shale. Photo by Richard Renn, CPG-06229.

FEATURES

Development of a Multifunction Best Fit Computer Program to Model Sediment Transport Data in Rivers

T. V. Hromadka II[1], A. N. Johnson[2], M. Phillips[3]

(1) Professor, Department of Mathematical Sciences, United States Military Academy, West Point, New York

(2) Academy Professor, Department of Mathematical Sciences, United States Military Academy, West Point, New York

(3) Professor and Department Head, Department of Mathematical Sciences, United States Military Academy, West Point, New York

INTRODUCTION

Analysis of sediment load in large rivers is an important problem in the management of the river hydraulics as well as environmental concerns. Increases in sediment load can result in sediment deposition in the river uniformly or at high concentrations at particular locations. Such sediment deposition can interfere with river flow carrying capacity and also may impact or improve environmental opportunities such as fish and wildlife populations. The usual approach in the analysis of sediment load and changes in the sediment load is to perform a statistical analysis of the sediment load data collected over time at monitoring stations. Several computer models and programs are currently available that operate upon such sediment load data and prepare the usual fitting of interpolation functions for subsequent use in statistical inferences. The USGS LOADEST program for estimating constituent loads in streams and rivers is one of these available programs. In the current Note, we describe the ongoing research effort to develop an interpolation computer program based upon Mathematica. In the documentation, 11 regression models are tabulated and are available in the LOADEST program to be fitted to the data. In the current study, more than 100 regression models are available with more models being incorporated. The expanded Mathematica computer program is directly developed from the program presented in Hromadka et al, 2007.

The typical problem in statistical analysis of data is to consider fitting (regression) functions that minimize some pre-defined measure of goodness-of-fit, such as the well-known least squares residual minimization technique, among others. Generally, a selection of basis functions are chosen, such as polynomials, or other basis functions, and then all of the basis functions are minimized together as a group according to the selected measure. In this note, we develop a Mathematica program that accomplishes the task of fitting prescribed functions to data, but we build the program to investigate combinations of the basis functions within the family of basis functions as an alternative to fitting the entire family of basis simultaneously. Because the procedure is automated, we have the advantage of doing such an effort without a significant investment of time or effort. The Mathematica program is expandable to include other families of basis functions or other measures of fit. The program measures the goodness of each set of basis functions used and then lists the measures achieved for each attempt.

The use of a set of basis functions to form a linear combination with coefficients to be determined by minimizing some best fit measure function to data according to some measure of fit is a process that is well-known and so will not be repeated here (Gallant 1987, among many others.) A set of basis functions are selected, such as a set of monomials, and then this entire set of basis functions is fitted to the target set of data by a technique such as a weighted Gramm-

Schmidt approach or other approach. Computer program Mathematica is particularly well-suited to applying to this task and provides automated determination of the best fit function to the prescribed data set. The program can be extended by simply adding additional bases to the list of functions being used as fitting functions.

THE MATHEMATICA PROGRAM

The Mathematica program is currently based upon use of the Mathematica FindFit operation. The various combinations of the basis functions are built from the selected basis function family up to the dimension programmed into the notebook. The dimension of this family, can be extended by adding additional basis functions. For each combination of these basis functions, a new trial function is defined, and the target data set is fitted using the Mathematica FindFit operation and the resulting measure of fit computed. Once all the combinations of basis functions are examined, they are ranked according to goodness of fit by sorting the resulting error residual measures.

Families of basis functions of typical interest include trigonometric basis functions such as: a Sin (b x + c) + d which results in several variations just with these four parameters. For example, a Sin(x), Sin(b x), Sin(x+c), Sin(x)+d, a Sin(b x) and so forth. With a competing goal in fitting functions to data being the possibility that simpler fitting functions may be better models of the underpinnings of these data, it is instructive to examine all combinations of basis functions from a prescribed family.

The Mathematica notebook is easily expandable. The proposed extended program software for loading into a web-site is currently under construction for publication on the web.

CONCLUSIONS

During the summer, cadets at the United States Military Academy at West Point can choose to participate in an AIAD program. The Advanced Individual Academic Development (AIAD) program provides cadets with an opportunity to observe and implement concepts from their course work in computational engineering mathematics with an immersion experience during a summer internship. The AIAD program enables cadets to work side by side with leaders in government and industry both stateside and abroad. The described research and development effort will be completed during the summer 2016 session.

REFERENCES

1. Gallant, A. Ronald, *Nonlinear Statistical Models*, John Wiley and Sons, New York, 1987.
2. Hromadka II, T. V., Glen, A.G., and Miguel, F., 2007, An Expandable Mathematica Program for Fitting Families of Basis Functions to Data, 6th Annual Hawaii International Conference on Statistics, Mathematics and Related Fields, 2007 Conference Proceedings, 6 pgs.
3. Runkel, R.L., Crawford, C.G., and Cohn, T.A., 2004, Load Estimator (LOADEST): A FORTRAN Program for Estimating Constituent Loads in Streams and Rivers**:** U.S. Geological Survey Techniques and Methods Book 4, Chapter A5, 69 p.