Sandia National Laboratories

GAMVT: **G**enerative **A**lgorithm for **M**ulti**V**ariate **T**imeseries

Joint Statistical Meetings (JSM)

August 8-12, 2021

Presenting Author: Jamie Thorpe, jthorpe@sandia.gov

U.S. DEPARTMENT OF ENERGY • NNSA
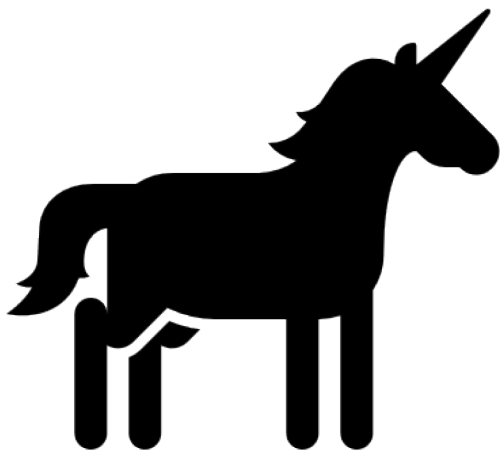
SAND number

A Familiar Problem

The Dataset You Want VS

◦ Large

◦ Robust
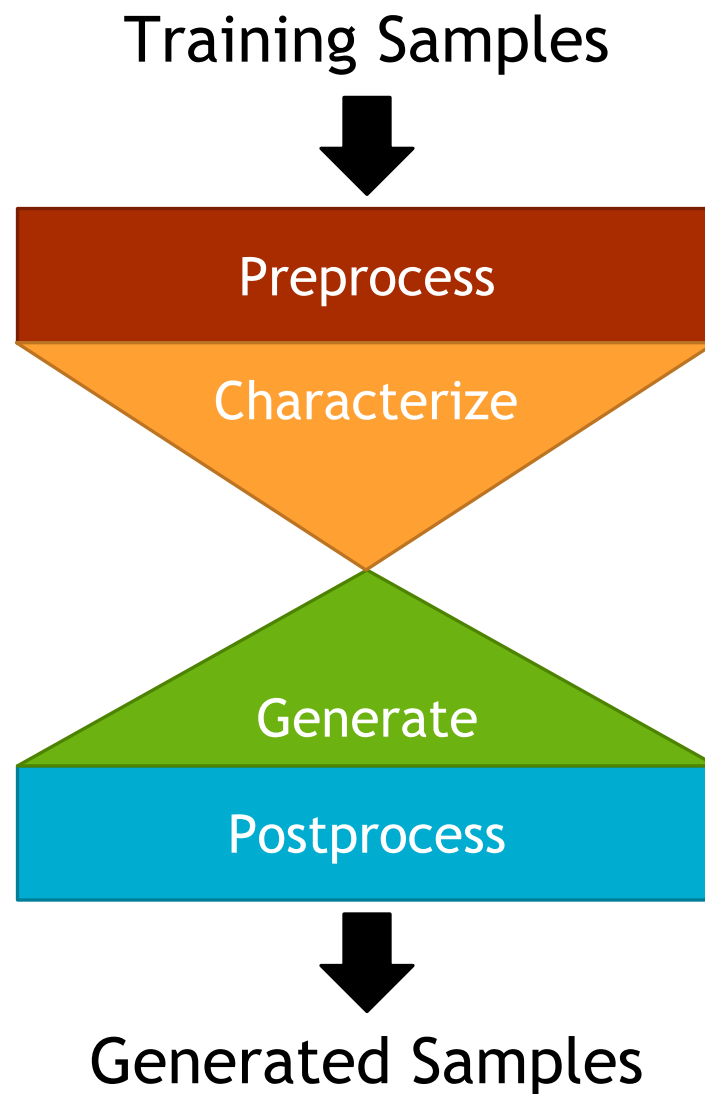
◦ Varied
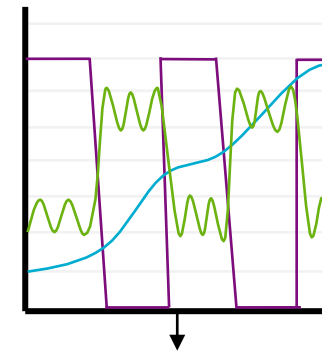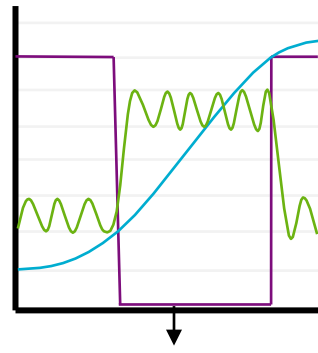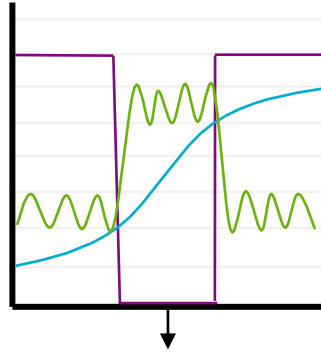
◦ Representative

◦ Immediately Applicable

The Dataset You Have

◦ Small (if it Exists at all)

◦ Sparse or Missing Values

◦ Limited Scope

◦ Underrepresented Classes

◦ Needs Work

GAMVT: Generative Algorithm for MultiVariate Timeseries

Training Samples

Preprocess

Characterize

Generate

Postprocess

Generated Samples

Preprocessing

Characterization:
- Patterns
- Statistics
- Relationships

Pattern = [A [B A]$^{1, 2}$]
Covariance (A)
Covariance (B)

	Section Statistics	Value Statistics
A	…	…
B	…	…

Generation

Postprocessing

Results
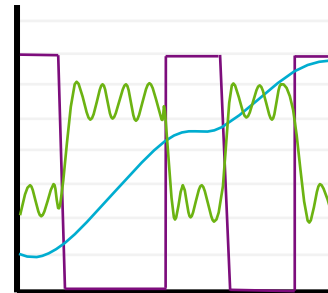
Space-Cyber System Use Case
- Small Satellite Emulation
- GPS and Camera data
- Series of Attack Classes

Compared to Existing Data Generation Techniques
- Generative Adversarial Network
- Variational Autoencoder

GAMVT Outperforms Existing Techniques on a Multivariate Timeseries Dataset

Preview of Full Talk

Motivation – Why Multivariate Timeseries Data?

GAMVT – How Does it Work?

Use Case – What Does the Data Look Like?

Results – How Does GAMVT Perform on the Use Case?