#### Multivariate Analysis in Vislt

#### **Robert Sisneros**

Oak Ridge National Lab

April 27, 2011

Robert Sisneros Multivariate Analysis in Vislt

<ロト <回 > < 回 > < 回 > .



- Algorithmically visualize and understand multivariate data
  - Attributes
  - Interactions among attributes
- System for specifying high-dimensional models
  - General, unconstrained
  - Implemented as an operator in VisIt
- Address basic visualization problem: "What/how much can we show in visualizations?"

(日) (圖) (E) (E) (E)

## High-dimensional Rendering (Parallel Coordinates)



Robert Sisneros Multivariate Analysis in Vislt

・ロト ・ 日 ・ ・ ヨ ・ ・ ヨ ・



Robert Sisneros Multivariate Analysis in Vislt

◆□ > ◆□ > ◆臣 > ◆臣 > ○ 里 ○ の Q @

## How Many High-Dimensional Relationships?



#### Robert Sisneros Multivariate Analysis in Vislt

< 注)→ 注



Robert Sisneros Multivariate Analysis in Vislt

◆□ > ◆□ > ◆臣 > ◆臣 > ○ 里 ○ の Q @

The Operator Demonstration



◆□ > ◆□ > ◆臣 > ◆臣 > ○ 里 ○ の Q @

The Operator Demonstration

- Specifying models
  - Specified concurrently
  - Sets of high-dimensional points
  - From space created by selected variables
  - Each point classified by its distance to candidate models
- Calculating distances
  - Distances calculated from data point to nearest model specification point
  - Different metrics
  - Different spaces

(日) (圖) (E) (E) (E)

The Operator Demonstration

## GUI

Specify Models							
Variables	hydroxyl_ra	stoichiomet					
Timesteps	Current	Current					
Point 1	0	.42					
Add	d Model	Add Varia	ible 🔻	Add	Point		
Delete Model Delete Point Add Threshold				reshold			
Select Calcula	Select Calculation Space:						
🔘 Variat	⊖ Variable		1 🔾 log		O Probability		
Distance Sele	ctions:						
● Euclidean () Manhattan () Maximum () Minkowski () Canberra							
Normalizing Scores:							
○ None ○ Individual Models							
Make defa	ault		Load	Save	Reset		
Apply	$\supset$			Post	Dismiss		

Robert Sisneros Multivariate Analysis in Vislt

◆□ > ◆□ > ◆ □ > ◆ □ > □ = のへで

The Operator Demonstration

## Output

- At each spatial location:
  - Model ID with minimal distance
  - Distance
- Operators creating new variables in Vislt?

The Operator Demonstration

#### Yes!



The Operator Demonstration

## **Multiple Attributes**



・ロト ・ 四 ト ・ 回 ト ・ 回 ト

The Operator Demonstration

## Simple Models



Robert Sisneros Multivariate Analysis in Vislt

The Operator Demonstration

## The Specification

Specify Mode	ls	
Variables	vortex_ts0	vortex_ts1
Timesteps	Current	Current
Point 1	0	0
Point 2	1	1
Point 3	2	2
Point 4	3	3
Point 5	4	4
Point 6	5	5
Point 7	6	6
Point 8	7	7
Point 9	8	8
Point 10	9	9
Point 11	10	10
Point 12	11	11

Robert Sisneros Multivariate Analysis in Vislt

< 注)→ 注

The Operator Demonstration

## **Using Scores**



Robert Sisneros Multivariate Analysis in Vislt

・ロト ・聞 ト ・ ヨ ト ・ ヨ ト

The Operator Demonstration

## **Complex Models: The Specification**

pecify Mode	ls				
	hast selected	huden al er	and an altest		
Variables	neat_release	nydroxyI_ra	scalar_dissi	stoicniomet	vorticity_m
Timesteps	Current	Current	Current	Current	Current
Point 1	a	a	a	a	a
Variables	heat_release	hydroxyl_ra	scalar_dissi	stoichiomet	vorticity_m
Timesteps	Current	Current	Current	Current	Current
Point 1	m	a	a	a	a
Variables	heat_release	hydroxyl_ra	scalar_dissi	stoichiomet	vorticity_m
Timesteps	Current	Current	Current	Current	Current
Point 1	м	a	a	a	a
Variables	heat_release	hydroxyl_ra	scalar_dissi	stoichiomet	vorticity_m
Timesteps	Current	Current	Current	Current	Current
Point 1	a	m	a	a	a
Variables	heat_release	hydroxyl_ra	scalar_dissi	stoichiomet	vorticity_m
Timesteps	Current	Current	Current	Current	Current
Point 1	a	м	a	a	a
Variables	heat_release	hydroxyl_ra	scalar_dissi	stoichiomet	vorticity_m

Robert Sisneros Multivariate Analysis in Vislt

・ロ ・ ・ 一 ・ ・ 日 ・ ・ 日 ・ ・

The Operator Demonstration

#### **Complex Models**



Robert Sisneros Multivariate Analysis in Vislt

・ロト ・ 日 ・ ・ ヨ ・ ・ ヨ ・

The Operator Demonstration

## Complex Models: Probability Space



Life Zones The Data Visualizing Life Zones

# Climate Change via Holdridge Life Zones

- System for classification of land areas based on atmospheric conditions
- Used to analyze vegetation pattern alterations due to global warming
- Specification as 3-dimensional models ideal for visualization approach

Life Zones The Data Visualizing Life Zones

## Canonical Representation of Life Zones



Robert Sisneros Multivariate Analysis in Vislt

・ロト ・ 四 ト ・ 回 ト ・ 回 ト

Introduction	Life Zones
The Approach	The Data
Application	Visualizing Life Zones

- Two different emission scenarios of CCSM 3 climate simulation data (2000-2099)
- ► A2
  - High global emissions
  - Continuously increasing population
  - Slow technological change
- ► B1
  - Low global emissions
  - Population peaks at 2050 then declines
  - Introduction of clean and resource-efficient technologies
- Data averaged into first (2000-2009) and last (2090-2099) decade

▲御 ▶ ▲ 陸 ▶ ▲ 陸 ▶ 二 陸

Life Zones The Data Visualizing Life Zones

## The First Decade



Robert Sisneros Multivariate Analysis in Vislt

Life Zones The Data Visualizing Life Zones

#### The Last Decade



Robert Sisneros Multivariate Analysis in Vislt

Life Zones The Data Visualizing Life Zones

## Should I Move?



Robert Sisneros Multivariate Analysis in Vislt

・ロト ・聞 ト ・ ヨ ト ・ ヨ ト

Life Zones The Data Visualizing Life Zones

## Future Work

- System upgrades
  - Temporal analysis
  - Save subspaces vs. rendering subspaces
  - "Pick" mode
- Application to Astrophysics simulation data

・ロ ・ ・ 一 ・ ・ 日 ・ ・ 日 ・ ・

Life Zones The Data Visualizing Life Zones

## To Models: Definitions

- Interactions formally identifiable as a set where collections of predicates are true
  - E.g. x|P(x), predicate may be x > 4
- Build hypothetical predictive models
  - Representation for any data point
  - Use models for classification
- Approach similar to model-based clustering

▲□ → ▲ □ → ▲ □ → □

Life Zones The Data Visualizing Life Zones

#### **Persistence Metric**



Life Zones The Data Visualizing Life Zones

# Magnitude of Change Metric



<ロト < 回ト < 回ト