**Manifold Surface**: Colors show values on the vertical axis. In manifold learning, the shape of the surface is learned based on known points. Black dots show a few given and a few predicted points. Blue arrows show a special feature of our manifold learning algorithm: A user can adjust any predicted value and the algorithm adjusts the surface near the adjusted point. It simultaneously spreads the adjustment farther or not depending on the surface slope and smoothes it to avoid over-fitting. Dotted circles show the new position of the dot. Thinner gray arrows on the grid show that it is also possible to adjust a whole section of the manifold surface.

---

**Geophysical Data Mining Using Interactive Manifold Learning**

**The Challenge**

Geothermal resources are theoretically more than adequate to supply human energy needs. According to a report published by the U.S. Geothermal Energy Association, there has been a 52% increase of countries with projects under development or active consideration since 2007. However, only a very small fraction of potential geothermal targets are exploited. The cost associated with drilling and completing wells is a major factor; the cost increases non-linearly with depth. In enhanced geothermal system power plants, drilling costs account for 42%-95% of total power plant cost. Geothermal data has yet to be effectively used to characterize known targets and explore new targets. The diversity of sensors, in addition to the continuing increase in data volume and diversified modalities, causes the data to be sparse in the location and composition of critical features. Lockheed Martin Advanced Technology Laboratories (ATL) collaborates with National Information and Communications Technology Australia Ltd. (NICTA) on geothermal target exploration and characterization. The objective is to apply machine learning and data fusion methods to assimilate and characterize massive quantities of high dimensional data and demonstrate that they can characterize known targets and explore new targets.

**Technical Approach**

Learning and estimating multiple correlating values remains an important and challenging problem. In recent years, Gaussian Processes (GP) have become one of the mainstream leading tools for machine learning. Usually formulated for a single output, multiple output GPs have been developed by NICTA. ATL developed a manifold learning-based approach to learn and estimate multiple correlating values. Experiments show that our manifold learning-based methods achieve similar results as multi-kernel GPs. A unique feature of the method is that it facilitates incremental and interactive learning, which allows gradual model adaptation and user feedback. ATL's manifold learning method can incrementally incorporate newly discovered data into the learned model without complete retraining of the model. It also allows the operators to enhance the model through manual adjustments of a few samples.

**Applications**

This research focuses on ill-posed inversion problems, such as the estimation of mineral sources and rock types deep in the ground from gravity, magnetic susceptibility, and seismic measurements. The technology developed can be applied to many pattern-discovery and anomaly-detection problems to aid decision making.

---
