

## **IMPROVED GROUNDWATER SYSTEM CHARACTERIZATION AND MAPPING USING HYDROGEOPHYSICAL DATA AND MACHINE-LEARNING WORKFLOWS**

*Michael J. Friedel<sup>1,2,\*</sup>, Neil Symington<sup>3</sup>, Larysa Halas<sup>3</sup>, Kokpiang Tan<sup>3</sup>, Ken Lawrie<sup>3</sup> & David Gibson<sup>4</sup>*

<sup>1</sup>*GNS Science, [m.friedel@gns.cri.nz](mailto:m.friedel@gns.cri.nz)*

<sup>2</sup>*University of Colorado, [Michael.J.Friedel@gmail.com](mailto:Michael.J.Friedel@gmail.com)*

<sup>3</sup>*Geoscience Australia, [Neil.Symington@ga.gov.au](mailto:Neil.Symington@ga.gov.au); [Larysa.Halas@ga.gov.au](mailto:Larysa.Halas@ga.gov.au); [ken.lawrie@ga.gov.au](mailto:ken.lawrie@ga.gov.au); [kokpiang.tan@ga.gov.au](mailto:kokpiang.tan@ga.gov.au);*

<sup>4</sup>*David Gibson Geological Services, [davidgibsongoe@gmail.com](mailto:davidgibsongoe@gmail.com)*

The Australian machine-learning workflows apply fusion, clustering, and estimation operations to hydrogeophysical data for deriving hydrostratigraphic units (HSUs). Data fusion is performed by training a self-organizing map (SOM) with these data. The application of Davies-Bouldin criteria to K-means clustering of SOM nodes determines the number and location of HSUs. Estimation is handled by iterative least-squares minimization of the SOM quantization and topographical errors. Two workflows provide 3D characterization of HSUs (and related attributes) from different hydrogeophysical data (measured, derived, interpolated, and estimated values) sets.

In Workflow 1, the SOM learns to recognize relationships among a subset of borehole geophysical and hydrogeologic data. Using the data-fusion approach described above, the missing hydrological data are estimated using these learned relationships and HSUs determined at borehole sample locations resulting in a low lateral density and high vertical density spatial distribution. Variogram modeling of the regional field data and HSU estimates is undertaken to evaluate the spatial statistical structure of selected attributes.

In workflow 2, the learned relationships between borehole data and the more spatially extensive AEM conductivity model are used to estimate the key attributes and HSUs at a number of locations away from the borehole. The AEM conductivity profile at a number of random locations are mapped to the SOM network and estimation performed to arrive at a set of continuous HSUs with high lateral density and medium vertical density (based on m-layer modelled structure). Performance metrics and validation are used to test each step of both workflows.