Transforming Geotechnical Data into Smart Data

The UK Institution of Civil Engineers' State of the Nation Report 2017 on "Digital Transformation" highlighted the importance of integrating physical and data assets:

"We often think of infrastructure as fixed networks and assets, but the reality of this is changing. We must think about not only the physical asset, but also its digital twin – all the associated data and the information that this can reveal. If we truly consider infrastructure as a service, then making this mental shift is essential. Delivering infrastructure based on outcomes for users drives us toward whole life decisions and recognising the value of the entire data estate. This approach makes best use of the endless flow of data, information and knowledge we can use to improve the services we deliver."

The geotechnical engineering profession is steeped in empiricism and experience. An example is the concept of presumptive bearing values or allowable soil pressures prescribed in old building codes. At present, data are primarily used as inputs to physical models at the design stage. The Burland Triangle describes the relationship between ground profile, soil behaviour, and modelling within a broader design context that must account for precedent, experience, and risk management.

We should view data as a primary asset that can directly improve our decisions and our service to society, rather than as a secondary asset in physical modelling and monitoring. To take the next leap in transforming our geotechnical data into smart data, it is necessary to go beyond classical univariate statistics to address the actual characteristics of geotechnical data that can be succinctly described as MUSIC – Multivariate, Uncertain and unique, Sparse, and InComplete. The conventional random variable model, although familiar and readily understood by many engineers, is not sufficiently powerful to model actual geotechnical data. More importantly, there is limited emphasis on automatic "learning" that will allow decision making to improve as more data accumulate. Our algorithms must meet the ongoing revolution in smart technologies.

For example, although it is commonly accepted that each site is unique to some degree, there is no method of characterizing this "uniqueness" that can lead to an automatic selection of "similar" sites. It is evident that generic correlation models that are widely used in the absence of sufficient site-specific data can be refined when the supporting database is drawn from similar sites only. Recent research explores how Bayesian learning can address this "site challenge" under the constraint of MUSIC.

The world is digitally connected and being revolutionised by new and powerful ways of collecting, analysing, and monetizing data. Clearly, there is a pressing need for the geotechnical engineering community to engage in this digital transformation. The ongoing compilation and analyses of large generic databases is an important first step in a broader digitalization agenda to connect geotechnical engineering to Industry 4.0. It is possible to envisage realizing "precision construction," where characterization of "site-specific" model factors and "site-specific" soil parameters based on both site-specific and generic data can lead to further customization of design to a particular site and even a particular location in a site.

If we do not start now, we will never arrive.