

21/11/2016

La major part de les anàlisis de sensibilitat publicades són errònies



Les anàlisis de sensibilitat, que relacionen la variabilitat dels valors de sortida amb la dels valors d'entrada en un procés, s'apliquen a un ventall molt ampli de camps de recerca. L'investigador Andrea Saltelli ha observat, però, que el 96% de les anàlisis d'aquest tipus publicades en revistes científiques són errònies. Saltelli defensa que molts investigadors construeixen models d'anàlisi equivocats perquè passen per alt condicions que serien trivials per als estadístics.

Imagina el model matemàtic d'un procés - com per exemple el rendiment d'una reacció química o el flux que passa per un aqüífer- que produeix una sortida diferent -com pot ser la quantitat d'una substància química o el flux d'aigua, per seguir amb els exemples -que depèn de cada una de les combinacions de les variables d'entrada. Aquestes variables podrien ser les característiques dels reactius químics o de la hidrologia de la conca en els nostres exemples.

Imagina també que les variables d'entrada són incertes, és a dir, que poden prendre diferents valors o bé perquè són desconeguts (el fet que la ràtio de la reacció química no es conegui amb precisió, a l'exemple) o bé perquè solen variar a la natura (com pot ser la intensitat de la precipitació que alimenta l'aqüífer). Hi ha, aleshores, un número infinit de possibles combinacions de variables d'entrada que donen cada valor diferent de sortida.

Mitjançant l'aplicació del model matemàtic un número suficient de vegades, es pot tenir una idea de com es distribueixen els valors de sortida, és a dir, quina és la seva mitjana, la

seva variància, etc. Aquest procediment s'anomena habitualment "anàlisi d'incertesa", ja que investiga la incertesa dels valors a la sortida. En aquest sentit, un pot preguntar-se si totes les variables d'entrada tenen el mateix pes sobre el grau d'incertesa de les sortides. De fet, això no passa gairebé mai, de manera que sovint moltes de les variables d'entrada poden modificar completament el valor de sortida amb un petit canvi en els valors de les entrades, al mateix temps que unes poques variables dominen el caràcter del procés. Aquesta recerca s'anomena "anàlisi de sensibilitat", ja que investiga com les sortides són sensibles a les variables d'entrada en un determinat procés. Un anàlisi de sensibilitat pot conduir a declaracions com "el 70% de la variància del valor de sortida és causada pel factor X_4 , el factor X_2 causa el 25% de la variància, i el restant 5% és causat per altres factors".

La major part de les anàlisis de sensibilitat (pel què sembla 96 de cada 100 l'any 2014, segons el nostre estudi) sorprenentment pateixen un gran defecte, fins al punt de ser inútils. Aquestes no combinen tots els valors d'entrada, i deixen inexplorada una llarga fracció –o bé gairebé tota- de les possibles combinacions.

Quin és el procediment que condueix cap a aquest resultat tant decebedor? El procediment –només adequat en l'aparença - és el següent. En primer lloc, calcular el valor de sortida en correspondència a un punt central on totes les variables tenen la seva "estimació més probable" o "més adequada". I, a continuació, allunyar cada variable d'aquest estàndard central, i calcular el valor de sortida per fer servir aquesta diferència com a una mesura de sensibilitat.

Tot això sembla lògic però els seus resultats només són bons si el model és lineal. Quan el model no ho és, poden succeir coses interessants quan més d'una variable es mou lluny del seu valor central. Per posar un exemple, una reacció química pot esdevenir explosiva per a una combinació particular d'entrades, cosa que no pot veure el procediment que acabem de descriure. Així, una anàlisi de sensibilitat feta d'aquesta manera produiria un rànquing invàlid sobre els factors d'entrada. Es podrien, per exemple, obviar factors que són cabdals. Això és ben conegut i trivial pels estadístics o practicants de l'anàlisi de sensibilitat, al mateix temps que segueix sent passat per alt per la majoria d'agents dedicats a la construcció de models d'anàlisi.

Fent servir un argument geomètric, el nostre article de recerca mostra que la típica anàlisi de sensibilitat de 10 factors deixa inexplorat el 99,75% de les possibles combinacions de valors. Això equival a un estat total de ceguesa quan s'equipara amb el potencial del model si se'l permet moure amb llibertat, en lloc d'estar lligat al centre dels seus valors de factor d'entrada.

Per a un procediment correcte s'hauria de distribuir els punts tant uniformement com fos possible dintre de l'espai on totes les combinacions del valor dels paràmetres fossin possibles.

Andrea Saltelli

Institut de Ciència i Tecnologia Ambientals (ICTA-UAB)

andrea.saltelli@uab.cat

Referències

[View low-bandwidth version](#)