

Noise complaints: how statistical innovations are cleaning up noisy data

Dr Catia Scricciolo

© Triff/Shutterstock.com



www.futurumcareers.com

Inspiring the
next generation

Noise complaints: how statistical innovations are cleaning up noisy data

Scientific data are rarely clean and clear. Often, they are blurred or distorted by 'noise' – random disturbances that get mixed in with real data and make them difficult to interpret. Thankfully, statisticians like **Dr Catia Scricciolo** from the **University of Verona** in Italy are developing methods to counteract the problem, drawing on mathematical optimisation and probability to remove this noise and reveal the signals it was hiding.



Dr Catia Scricciolo

Associate Professor, Department of Economics,
University of Verona, Italy

Field of research

Statistics

Research paper

Adaptive minimax-optimal Wasserstein deconvolution with unknown error distributions. Scricciolo, C., (2026)
doi: [10.1016/j.spl.2025.110589](https://doi.org/10.1016/j.spl.2025.110589)

doi: [10.33424/FUTURUM700](https://doi.org/10.33424/FUTURUM700)

Talk like a ... **statistician**

Deconvolution – a mathematical method used to recover the original signal from data degraded by noise

Distribution – a function that describes the relative frequency of different outcomes in a dataset

Noise – in data analysis, the random fluctuations, measurement errors, or irrelevant information in a dataset that are not part of the underlying true signal and can obscure meaningful patterns

Optimisation – the process of finding the optimal solution to a mathematical problem

Regularisation term – a component in a model that prevents the model from trying to fit every fluctuation in the data

Wasserstein metric – a way to define the distance between two probability distributions

Imagine taking a photo with an unsteady hand. Your picture comes out blurry due to the shaking of the camera. “In theory, if you knew (or could guess) how the camera shook, you could digitally process the photo to remove the blur and get a sharp image,” says Dr Catia Scricciolo, a statistician at the University of Verona. “The same principle applies to removing the ‘blur’ from data, via a process called deconvolution.”

This is an important principle for virtually every scientific field. Blurry data can hide meaningful trends, which undermines

scientific progress. So, statisticians have been working out how to counteract blurry data using the mathematical process of deconvolution. But doing so reliably, especially since randomness is involved, takes careful and innovative thought. Catia works on exactly this challenge, exploring how statistical methods can minimise the effects of noise and reveal the meaningful information within a dataset.

Blurry data

When scientists record data – whether astronomers are measuring stars, medical personnel are taking MRI scans, or

physicists are measuring subatomic particles – their results are rarely squeaky clean. This makes it difficult to see the true signal through the blur. “In science, ‘blur’ comes from equipment limitations, physical effects or interference, as well as random noise,” says Catia. “To counter this, scientists use deconvolution – a process that mathematically recovers the original clear signal from the blurred measurement.”

When astronomers look up at the night sky using telescopes here on Earth, the atmosphere bends light randomly, which is why stars appear to twinkle.



© Gorodenkoff/Shutterstock.com

When biologists look at cells under the microscope, all the parts of the three-dimensional cell that are not in the targeted two-dimensional plane show up as out-of-focus haze. “We can use deconvolution to turn blurry, hard-to-interpret data like these into sharper, more useful information,” says Catia. “Mathematically speaking, this involves estimating the original signal by undoing the effects of an unknown or partially known blurring process.”

Noise: the enemy of deconvolution

But deconvolution becomes trickier when the blurring is caused by random, unpredictable effects. “This type of unwanted, random disturbance is called noise,” explains Catia. “Noise makes deconvolution much harder, and sometimes impossible.” The more noise there is, the harder it is to decide what is signal and what is just random, limiting what can be recovered. Ideally, noise is minimised when the measurements are taken, but this is not always possible. “Because noise is random, you can never be completely sure which parts of your observation are the original signal and which are just noise,” says Catia. “So, deconvolution can give you an estimate of the truth, but not the exact truth.”

Thankfully, statisticians have some nifty tricks up their sleeves to help with deconvolution for noisy data. “The Wasserstein metric defines a distance between probability distributions,” says Catia. “In plainer language, it’s a way to focus on the overall shape of a distribution

“

All together, these methods turn messy, noisy data into a clean estimate of the original distribution.

”

of data, rather than getting caught up in the ‘bumps’ caused by noise.” Catia studies Wasserstein deconvolution through a mix of mathematics and probability. “I use mathematical optimisation to find the original distribution that best explains the noisy data,” she says. “The Wasserstein metric then acts as a score that measures how far my guess is from the truth.” Catia also adds in a regularisation term that forces a recovered distribution to smooth out any wiggles that are likely to be due to noise rather than actual quirks in the data. “All together, these methods turn messy, noisy data into a clean estimate of the original distribution,” she says.

Helping out scientists

Catia’s research is improving the reliability of Wasserstein deconvolution methods, which other scientists can then use to clean up their own datasets. “Scientists in other fields can apply my deconvolution method without needing to be experts in the mathematics behind it,” she says. “For

example, doctors can use it to sharpen MRI scans, helping them see tumours more clearly.” Going back to astronomers, they can use Catia’s method to clean up blurry images of distant galaxies, uncovering details that were previously hidden by the noise of the atmosphere. Space telescopes can also benefit – when the Hubble Telescope was first launched, there was an error in its mirror array which introduced blur into the data it recorded. Using deconvolution, astronomers recovered stunning images of space from the blurry data.

Catia hopes that her research can help accelerate scientific discovery by taking the burden off scientists to develop their own data cleaning methods. “Instead of struggling to invent their own mathematical solutions, they can use the Wasserstein deconvolution techniques I helped develop,” she says. “In short, my work removes a barrier so other scientists can focus on their own research – whether that is curing a disease, discovering a new planet or understanding human behaviour.”



© Triff/Shutterstock.com

+44 117 909 9150
info@futurumcareers.com
www.futurumcareers.com

futurum)