To achieve accurate classification in such a high-dimensional space, Willett and colleagues' classification algorithm used current machine-learning methods, along with a type of artificial neural network called a recurrent neural network (RNN), which is especially good at predicting sequential data. Harnessing the power of RNNs requires ample training data, but such data are limited in neural interfaces, because few users want to imagine writing for hours on end. The authors solved this problem using an approach known as data augmentation, in which neural activity patterns previously generated by the participant are used to produce artificial sentences on which to train the RNN. They also expanded their training data by introducing artificial variability into the patterns of neural activity, to mimic changes that occur naturally in the human brain. Such variability can make RNN BCIs more robust[9].

Thanks to these methods, Willett and colleagues' algorithm provided impressively accurate classification, picking the correct character 94.1% of the time. By including predictive-language models (similar to those that drive auto-correct functions on a smartphone), they further improved accuracy to 99.1%. The participant was able to type accurately at a speed of 90 characters per minute — a twofold improvement on his performance with past iBCIs.

This study's achievements, however, stem from more than machine learning. A decoder's performance is ultimately only as good as the data that are fed into it. The researchers found that neural data associated with attempted handwriting are particularly well-suited for typing tasks and classification. In fact, handwriting could be classified quite well even with simpler, linear algorithms, suggesting that the neural data themselves played a large part in the success of the authors' approach.

By simulating how the classification algorithm performed when tested with different types of neural activity, Willett *et al.* made a key insight — neural activity during handwriting has more temporal variability between characters than does neural activity when users attempt to draw straight lines, and this variablility actually makes classification easier. This knowledge should inform future BCIs. Perhaps counter-intuitively, it might be advantageous to decode complex behaviours rather than simple ones, particularly for classification tasks.

Willett and co-workers' study begins to deliver on the promise of BCI technologies. iBCIs will need to provide tremendous performance and usability benefits to justify the expense and risks associated with implanting electrodes into the brain. Importantly, typing speed is not the only factor that will determine whether the technology is adopted — the longevity and robustness of the approach also require analysis. The authors present promising evidence that their algorithms will perform well with limited training data, but further research will probably be required to enable the device to maintain performance over its lifetime as neural activity patterns change. It will also be crucial to conduct studies to test whether the approach can be generalized for other users, and for settings outside the laboratory.

Another question is how the approach will scale and translate to other languages. Willett and colleagues' simulations highlight that several characters of the Latin alphabet are written similarly (r, v and u, for instance), and so are harder to classify than are others. One of us (P.R.) speaks Tamil, which has 247, often very closely related, characters, and so might be much harder to classify. And the question of translation is particularly pertinent for languages that are not yet well represented in machine-learning predictive-language models.

Although much work remains to be done, Willett and co-workers' study is a milestone that broadens the horizon of iBCI applications. Because it uses machine-learning methods that are rapidly improving, plugging in the latest models offers a promising path for future improvements. The team is also making its data set publicly available, which will accelerate advances. The authors' approach has brought neural interfaces that allow rapid communication much closer to a practical reality.

Pavithra Rajeswaran and **Amy L. Orsborn** are in the Department of Bioengineering, University of Washington, Seattle, Washington 98195, USA. **A.L.O.** is also in the Department of Electrical and Computer Engineering, and at the Washington National Primate Research Center, University of Washington.
e-mail: aorsborn@uw.edu

1. Willett, F. R., Avansino, D. T., Hochberg, L. R., Henderson, J. M. & Shenoy, K. V. *Nature* **593**, 249–254 (2021).
2. Mott, M. E., Williams, S., Wobbrock, J. O. & Morris, M. R. in *Proc. 2017 CHI Conf. Human Factors in Computing Systems* 2558–2570 (ACM, 2017).
3. Hochberg, L. R. *et al. Nature* **442**, 164–171 (2006).
4. Hochberg, L. R. *et al. Nature* **485**, 372–375 (2012).
5. Collinger, J. L. *et al. Lancet* **381**, 557–564 (2013).
6. Pandarinath, C. *et al. eLife* **6**, e18554 (2017).
7. Ajiboye, A. B. *et al. Lancet* **389**, 1821–1830 (2017).
8. Rezeika, A. *et al. Brain Sci.* **8**, 57 (2018).
9. Sussillo, D., Stavisky, S. D., Kao, J. C., Ryu, S. I. & Shenoy, K. V. *Nature Commun.* **7**, 13749 (2016).

Atmospheric chemistry

How ant acid forms in the atmosphere

Joost de Gouw & Delphine Farmer

Known sources of formic acid could not explain the observed atmospheric concentrations of this compound. The discovery of a previously unknown pathway that generates formic acid in the atmosphere resolves this discrepancy. **See p.233**

Formic acid is one of the simplest and most abundant organic molecules in Earth's atmosphere, but its sources have been poorly understood for many years. Laboratory and field studies[1–3] have shown that most formic acid is not emitted directly from sources, but is produced by chemical reactions in the atmosphere. However, the chemistry responsible has been a mystery. On page 233, Franco *et al.*[4] report that formic acid could be formed by a mechanism that starts with formaldehyde reacting with water in cloud droplets.

The word 'formic' derives from *formica*, the Latin word for ant, and the compound is indeed released from ant hills[5]. Other, and larger, emission sources include vegetation, biomass burning[6] and fossil-fuel combustion[7]. However, the combined emissions from known sources are too small to explain the concentrations of formic acid in the atmosphere, and several studies have concluded that formation in the atmosphere is a much bigger contributor (see ref. 1, for example).

Levels of formic acid can be measured by mass spectrometry and optical spectroscopy, and from satellite instruments, so there is excellent information about the distribution of this compound in the atmosphere. Observations have shown that atmospheric concentrations of formic acid increase rapidly in urban[8] and forest[9] air during the day. However, researchers have been unable to identify the chemical reactions responsible for this increase. Detailed studies that considered all of the possible known chemical pathways could explain only a fraction of formic acid produced, both in polluted and remote regions[3,9], and so the search for alternative

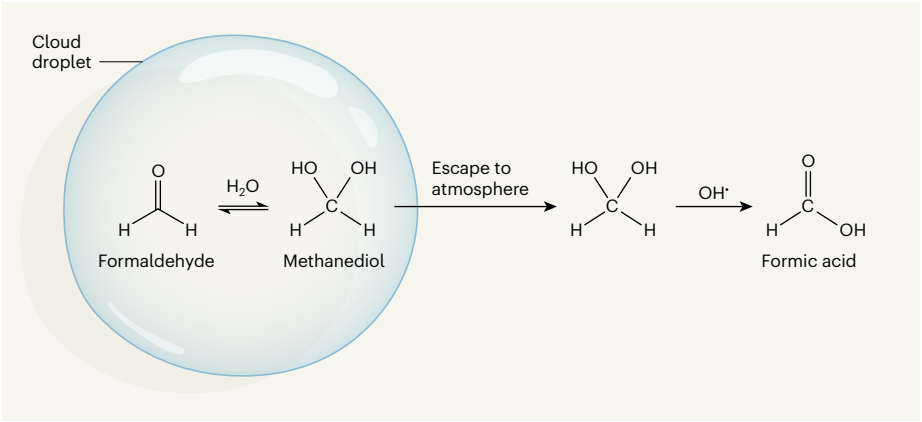

Figure 1 | A pathway for the formation of formic acid. Franco *et al.*[4] report evidence of a previously unknown mechanism for the production of atmospheric formic acid. The authors propose that formaldehyde reacts with water in cloud droplets to form methanediol. This compound escapes to the atmosphere, where it reacts with oxidants such as hydroxyl radicals ($OH^{\bullet}$) to form formic acid. The mechanism potentially explains why atmospheric levels of formic acid are higher than would be expected on the basis of previously known sources.

chemical pathways has continued[10].

Formic acid constitutes a substantial fraction of the organic carbon in the atmosphere[11]. Organic compounds in the gas phase have a key role in atmospheric oxidation reactions. They therefore influence the lifetime of greenhouse gases and the formation of ozone and aerosols — microscopic particles suspended in the air — in the troposphere, the lowest approximately 10 kilometres of the atmosphere. Organic compounds in the aerosol phase are air pollutants and have a cooling effect on climate. The simple fact that formic acid formation is so poorly understood calls into question our understanding of atmospheric organic carbon as a whole. Formic acid also contributes to acidity in the atmosphere and to the deposition of acids in ecosystems, particularly in remote regions.

Franco *et al.* now show that hydrolysis of formaldehyde in cloud droplets can lead to the formation of formic acid (Fig. 1). Inside cloud droplets, formaldehyde (HCHO) converts to its hydrated form, methanediol ($HOCH_2OH$). The authors demonstrate that methanediol can outgas from the cloud droplets and react with hydroxyl radicals ($OH^{\bullet}$) in the atmosphere to form formic acid (HCOOH) in the gas phase.

Why did it take so long to uncover this pathway? It had previously been assumed that two other processes short-circuit this chemistry: dehydration of methanediol, which re-forms formaldehyde; and rapid oxidation of methanediol inside cloud droplets. Simply put, methanediol was thought to be so short-lived that it never leaves the cloud droplet. However, by carefully considering the timescale for dehydration, and combining this analysis with results from experiments in an atmospheric-simulation chamber, Franco and colleagues show that methanediol can volatilize into the gas phase from cloud droplets and then oxidize to formic acid.

Franco *et al.* go on to incorporate methanediol oxidation into a global climate–chemistry model (a computational climate model that incorporates atmospheric-chemistry pathways) and evaluate the predicted levels of formic acid by comparing them with measurements obtained from a space-based instrument. The authors adopted an innovative approach to do this in more detail than was previously possible: they identified columns of formic acid in the atmosphere by analysing satellite data using a computer algorithm known as an artificial neural network. Such approaches are increasingly being used to quantify, at relatively low computational cost, atmospheric chemical species that have weak spectroscopic signatures. Crucially, the authors show that levels of atmospheric formic acid simulated by models that include methanediol oxidation agree with observations much more closely than with levels predicted by models that do not include this chemistry.

"This oxidation chemistry potentially explains some previously reported field observations."

Methanediol oxidation potentially explains some of the previously reported field observations. Formaldehyde is predominantly formed by photochemistry (light-induced reactions in the atmosphere), and a key precursor of formaldehyde on a global scale is isoprene — a compound emitted by trees and shrubs. Methanediol oxidation involves formaldehyde, and so it makes sense that concentrations of formic acid over forested regions are large. Moreover, formaldehyde formation in air is enhanced when atmospheric levels of the pollutant nitrogen oxide are high[12]. This could explain why formic acid formation is observed in urban air at nitrogen oxide concentrations that would prohibit formation of this acid by other gas-phase reactions (such as reactions between alkene compounds and ozone[13]).

Nevertheless, many questions remain. Formic acid can form in cloud-free atmospheres[8] and in the canopies of forests[14]. Could methanediol oxidation occur not only in cloud droplets, but also in aqueous aerosol particles or on wet surfaces? And are atmospheric concentrations of methanediol as high as the levels implied by Franco and colleagues' study? This will need to be evaluated by measurements.

Several other organic acids behave similarly to formic acid in the atmosphere[8]. Could their formation be explained by analogous pathways involving other aldehydes? The devil is in the detail — the rates of the various processes involved will determine whether these putative pathways contribute much to the atmospheric levels of other acids. And what are the sinks of formic acid? Franco and colleagues suggest that formic acid is removed from the atmosphere by scavenging (absorption by precipitation and cloud droplets) and photochemical reactions more quickly than was reported in previous studies, which implies that formic acid sinks should be further investigated. The authors' findings will motivate many follow-up studies, both in the laboratory and in the field, and methanediol oxidation adds to the array of pathways that atmospheric chemists can use to explain observations of formic acid.

Joost de Gouw is at the Cooperative Institute for Research in Environmental Sciences and in the Department of Chemistry, University of Colorado, Boulder, Colorado 80309, USA. **Delphine Farmer** is in the Department of Chemistry, Colorado State University, Fort Collins, Colorado 80523, USA.
e-mail: joost.degouw@colorado.edu

1. Paulot, F. *et al. Atmos. Chem. Phys.* **11**, 1989–2013 (2011).
2. Yuan, B. *et al. Atmos. Chem. Phys.* **15**, 1975–1993 (2015).
3. Millet, D. B. *et al. Atmos. Chem. Phys.* **15**, 6283–6304 (2015).
4. Franco, B. *et al. Nature* **593**, 233–237 (2021).
5. Graedel, T. E. & Eisner, T. *Tellus B* **40B**, 335–339 (1988).
6. Akagi, S. K. *et al. Atmos. Chem. Phys.* **11**, 4039–4072 (2011).
7. Crisp, T. A. *et al. Atmos. Environ.* **98**, 426–433 (2014).
8. Veres, P. R. *et al. Geophys. Res. Lett.* **38**, L17807 (2011).
9. Link, M. F., Nguyen, T. B., Bates, K., Müller, J.-F. & Farmer, D. K. *ACS Earth Space Chem.* **4**, 730–740 (2020).
10. Shaw, M. F. *et al. Nature Commun.* **9**, 2584 (2018).
11. Heald, C. L. *et al. Atmos. Chem. Phys.* **8**, 2007–2025 (2008).
12. Wolfe, G. M. *et al. Atmos. Chem. Phys.* **16**, 2597–2610 (2016).
13. Jenkin, M. E., Saunders, S. M. & Pilling, M. J. *Atmos. Environ.* **31**, 81–104 (1997).
14. Fulgham, S. R. *et al. Geophys. Res. Lett.* **47**, e2020GL088745 (2020).