

Using text analytics, scientists map the spread of potato blight prior to the Irish potato famine

February 15 2024



Late blight lesion on a potato leaf. Credit: Jean Ristaino, NC State University

North Carolina State University researchers used text analytics on both historic and modern writing to reveal more information about the effects and spread of the plant pathogen—now known as *Phytophthora infestans*—that caused the 1840s Irish potato famine and that continues to vex breeders of potatoes and tomatoes.

The study examined keyword terms like "potato rot" and "potato disease" after digitizing historic farm reports, news accounts and U.S. Patent Office agricultural records from 1843 to 1845 to show how the pathogen first spread across the northeast United States before causing the devastating famine in Ireland in 1845. The study also used [text analysis](#) to track [social media feeds](#) for the modern-day spread of late blight.

Textual analysis holds promise as a useful tool to help researchers track and visualize both historic and current plant diseases, the researchers say.

"We went back to original descriptions of the potato disease outbreaks in the United States because they occurred between 1843 and 1845, before outbreaks occurred in Europe," says Jean Ristaino, William Neal Reynolds Distinguished Professor of Plant Pathology at North Carolina State University and corresponding author of a [paper](#) in *Scientific Reports* that describes the study.

"We searched those descriptions by keywords, and by doing that we were able to recreate the original outbreak maps using location coordinates mentioned in the documents.

"We were also trying to learn what people were thinking about the disease at the time and where it came from."



An infected potato tuber like one that might have sparked the first outbreaks in US. Credit: Jean Ristaino

The analysis documents late blight disease on potatoes in five states—New York, Delaware, Massachusetts, New Jersey and Pennsylvania—before it spread to the rest of the northeastern U.S. and into Canada between 1843 and 1845. The pathogen later wreaked havoc on Europe—especially Ireland.

The paper also examined tweets from 2012 to 2022 to learn more about modern spread of *P. infestans*. They mined tweets for both common and scientific names of the pathogen and were able to geolocate the sources.

"The social media mining was interesting because we found that most

people talking about this disease are scientists in developed countries promoting their own work on Twitter (now X)," Ristaino said. "It was also interesting to note that states where the disease appeared all those many years ago still have the disease now."

The study also used Google Ngram search terms to reveal a surprising finding. The researchers saw a spike in late blight disease reported in 1950s documents. Drilling down into the relevant academic literature cited in the documents, Ristaino saw evidence of a large late blight outbreak in tomatoes in the United States after World War II.

"That could have been the emergence of a new North American strain of the pathogen, known as U.S. 1, that became really widespread after that," Ristaino said.

Ristaino added that she and her team plan to continue this type of work and expand the analytic tools to other plant diseases and pests.

More information: Jean Ristaino, Reconstructing historic and modern potato late blight outbreaks using text analytics, *Scientific Reports* (2024). DOI: [10.1038/s41598-024-52870-2](https://doi.org/10.1038/s41598-024-52870-2).
www.nature.com/articles/s41598-024-52870-2

Provided by North Carolina State University

Citation: Using text analytics, scientists map the spread of potato blight prior to the Irish potato famine (2024, February 15) retrieved 21 May 2024 from <https://phys.org/news/2024-02-text-analytics-scientists-potato-blight.html>

| |
|---|
| This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only. |
|---|