

[Skip to Main Content](#)

Wiley Online Library

Wiley
Online
Library

Search

- [This Journal](#)
- [Anywhere](#)

- Search term
[Advanced Search](#) [Citation Search](#)

- Search term
[Advanced Search](#) [Citation Search](#)

[Login / Register](#)

STATISTICAL ANALYSIS AND DATA MINING

Original Article

A weighted random forests approach to improve predictive performance

[Stacey J. Winham](#)

Corresponding Author

E-mail address: winham.stacey@mayo.edu

Department of Health Sciences Research, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

Department of Health Sciences Research, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

[Search for more papers by this author](#)

[Robert R. Freimuth](#)

Department of Health Sciences Research, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

[Search for more papers by this author](#)

[Joanna M. Biernacka](#)

Corresponding Author

E-mail address: biernacka.joanna@mayo.edu

Department of Health Sciences Research, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

Department of Psychiatry and Psychology, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

Department of Health Sciences Research, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

[Search for more papers by this author](#)

[Stacey J. Winham](#)

Corresponding Author

E-mail address: winham.stacey@mayo.edu

Department of Health Sciences Research, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

Department of Health Sciences Research, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

[Search for more papers by this author](#)
[Robert R. Freimuth](#)

Department of Health Sciences Research, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

[Search for more papers by this author](#)
[Joanna M. Biernacka](#)

Corresponding Author

E-mail address:biernacka.joanna@mayo.edu

Department of Health Sciences Research, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

Department of Psychiatry and Psychology, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

Department of Health Sciences Research, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, USA

[Search for more papers by this author](#)

First published: 08 July 2013

<https://doi.org/10.1002/sam.11196>

Cited by: 9

[About](#)

[Access](#)



[PDF](#)



[PDF](#)

[Tools](#)

- [Request permission](#)
- [Export citation](#)
- [Add to favorites](#)
- [Track citation](#)

[Share](#)

Give access

[Share full text access](#)



Share full text access

Please review our [Terms and Conditions of Use](#) and check box below to share full-text version of article.

I have read and accept the Wiley Online Library Terms and Conditions of Use.

Shareable Link

Use the link below to share a full-text version of this article with your friends and colleagues. [Learn more.](#)

Copy URL

Share a link

- [Email to a friend](#)
- [Facebook](#)
- [Twitter](#)
- [Linkedin](#)
- [Google+](#)
- [Reddit](#)
- [CiteULike](#)

Abstract

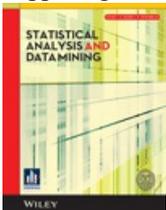
Identifying genetic variants associated with complex disease in high-dimensional data is a challenging problem, and complicated etiologies such as gene–gene interactions are often ignored in analyses. The data-mining method random forests (RF) can handle high dimensions; however, in high-dimensional data, RF is not an effective filter for identifying risk factors associated with the disease trait via complex genetic models such as gene–gene interactions without strong marginal components. In this article we propose an extension called weighted random forests (wRF), which incorporates tree-level weights to emphasize more accurate trees in prediction and calculation of variable importance. We demonstrate through simulation and application to data from a genetic study of addiction that wRF can outperform RF in high-dimensional data, although the improvements are modest and limited to situations with effect sizes that are larger than what is realistic in genetics of complex disease. Thus, the current implementation of wRF is unlikely to improve detection of relevant predictors in high-dimensional genetic data, but may be applicable in other situations where larger effect sizes are anticipated. © 2013 Wiley Periodicals, Inc. *Statistical Analysis and Data Mining*, 2013

[Citing Literature](#)

Number of times cited: 9

- Shiyang Xuan, Guanjun Liu and Zhenchuan Li, Refined Weighted Random Forest and Its Application to Credit Card Fraud Detection, *Computational Data and Social Networks*, 10.1007/978-3-030-04648-4_29, (343-355), (2018).
[Crossref](#)
- Ranadip Pal, Predictive modeling based on random forests, *Predictive Modeling of Drug Sensitivity*, 10.1016/B978-0-12-805274-7.00007-5, (149-188), (2017).
[Crossref](#)
- Robin Genuer, Jean-Michel Poggi, Christine Tuleau-Malot and Nathalie Villa-Vialaneix, Random Forests for Big Data, *Big Data Research*, 10.1016/j.bdr.2017.07.003, **9**, (28-46), (2017).
[Crossref](#)
- Stacey J. Winham, Gregory D. Jenkins and Joanna M. Biernacka, Modeling X Chromosome Data Using Random Forests: Conquering Sex Bias, *Genetic Epidemiology*, **40**, 2, (123-132), (2015).
[Wiley Online Library](#)
- Gérard Biau and Erwan Scornet, A random forest guided tour, *TEST*, 10.1007/s11749-016-0481-7, **25**, 2, (197-227), (2016).
[Crossref](#)
- Kumar Dheenadayalan, G. Srinivasaraghavan and V. N. Muralidhara, Pruning a Random Forest by Learning a Learning Algorithm, *Machine Learning and Data Mining in Pattern Recognition*, 10.1007/978-3-319-41920-6_41, (516-529), (2016).
[Crossref](#)
- Christos K. Aridas, Sotiris B. Kotsiantis and Michael N. Vrahatis, Increasing Diversity in Random Forests Using Naive Bayes, *Artificial Intelligence Applications and Innovations*, 10.1007/978-3-319-44944-9_7, (75-86), (2016).
[Crossref](#)
- Raziur Rahman, Saad Haider, Souparno Ghosh and Ranadip Pal, Design of Probabilistic Random Forests with Applications to Anticancer Drug Sensitivity Prediction, *Cancer Informatics*, **14s5**, (CIN.S30794), (2015).
[Crossref](#)
- Robert M Cronin, Jacob P VanHouten, Edward D Siew, Svetlana K Eden, Stephan D Fihn, Christopher D Nielson, Josh F Peterson, Clifton R Baker, T Alp Ikizler, Theodore Speroff and Michael E Matheny, National Veterans Health Administration inpatient risk stratification models for hospital-acquired acute kidney injury, *Journal of the American Medical Informatics Association*, **22**, 5, (1054), (2015).
[Crossref](#)

Supporting Information



[Volume6, Issue6](#)

[Special Issue: JSM 2012](#)

December 2013

Pages 496-505

Journal Recommendation service

Find the right journal for your research

[Learn more](#)

WILEY

- [Access](#)
- [Related](#)
- [Information](#)
- - [Log in to get access](#)
 - [INSTITUTIONAL LOGIN >](#)

Personal login

Wiley Online Library

Log in to your account

Email or Customer ID

Password

[Forgot password?](#)

[NEW USER >](#)

[Purchase instant access](#)

The following formats are available to purchase through ReadCube:

\$6	Rent for 48 hours ⓘ	<input type="radio"/>
	Printing and saving restrictions apply	
\$15	Buy cloud access ⓘ	<input type="radio"/>
	Printing and saving restrictions apply	
\$38	Buy PDF	<input checked="" type="radio"/>

If you previously purchased this article, [Sign In to ReadCube](#).

CHECKOUT

Powered by  readcube

- - [Recommend to a librarian](#)

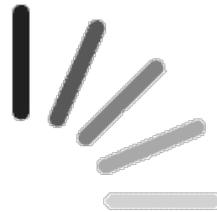
Wiley Editing Services

Ensure your manuscript is ready for submission

✓ English Language Editing

[Learn more](#)

WILEY



[Caption](#)

Additional links

About Wiley Online Library

- [Privacy Policy](#)

- [Terms of Use](#)
- [Cookies](#)
- [Accessibility](#)

Help & Support

- [Contact Us](#)

Opportunities

- [Subscription Agents](#)
- [Advertisers & Corporate Partners](#)

Connect with Wiley

- [The Wiley Network](#)
- [Wiley Press Room](#)

Copyright © 1999-2018 [John Wiley & Sons, Inc.](#) All rights reserved

WILEY

Wiley Online Library

Log in to Wiley Online Library

Email or Customer ID

Password

[Forgot password?](#)

[NEW USER > INSTITUTIONAL LOGIN >](#)

Wiley Online Library

Change Password

Old Password

New Password

Too Short Weak Medium Strong Very Strong Too Long

Congrats!

Your password has been changed

Create a new account

Email or Customer ID

[Returning user](#)

Wiley Online Library

Forgot your password?

Enter your email address below. If your address has been previously registered, you will receive an email with instructions on how to reset your password. If you don't receive an email, you should register as a new user

Email or Customer ID

RESET PASSWORD

Please check your email for your password reset instructions.

Request Username

Can't sign in? Forgot your username?

Enter your email address below and we will send you your username

Email or Customer ID

Submit

[Close](#)

If the address matches an existing account you will receive an email with instructions to retrieve your username