# Applied Scientist (Quantitative Genetics)

**Location:** San Carlos, CA
**Company:** Heritable Agriculture
**U.S. work eligibility:** Required

## About Heritable:

At Heritable Agriculture, we're reinventing the way the world grows food by making plants programmable. Our team blends AI, multi-omics, and remote sensing to create crops that thrive where others struggle. If you're ready to build the future of farming, where data meets dirt and innovation takes root, come grow with us.

## About the role:

We're looking for a proactive, curious, and fast-learning scientist to help build production-grade models and data pipelines at the intersection of plant breeding, environmental data characterization, and machine learning. You'll help design, implement, and scale robust, well-tested analytical systems that power phenotype prediction and decision support. This role blends quantitative genetics, biostatistics, and applied ML with strong software engineering practices to deliver reliable, maintainable tools used in production.

## What you'll do:

- Build and maintain end-to-end pipelines for phenotype (genomic) prediction and GxE modeling across multi-environment trials.
- Develop advanced modeling approaches for phenotypic prediction and breeding decisions.
- Integrate diverse environmental covariates and remote-sensing data into prediction models.
- Implement high-quality data wrangling and feature engineering across phenotypic, genotypic, and environmental datasets
- Write production-grade Python and R code with unit/integration tests, CI/CD, reproducible workflows, and clear documentation.
- Collaborate with breeders, agronomists, product managers, and data engineers to translate questions into robust analytical designs and decision tools.
- Proactively anticipate edge cases, failure modes, and data quality issues; add safeguards, validations, and monitoring.
- Optimize performance and cost using HPC and/or cloud services; containerize and automate recurring analyses.

## Required Qualifications:

- PhD or equivalent experience in Quantitative Genetics, Statistical Genetics, Biostatistics, Applied Machine Learning, Computer Science, or a related field.
- Strong grounding in quantitative genetics, plant breeding, and genomic prediction methodology (experimental design, enviromics, selection indices, multi-trait/multi-environment modeling).
- Proven experience analyzing large, heterogeneous datasets and building ML/statistical models end-to-end.
- Proficiency in model tuning and selection, including hyperparameter optimization
- Expert data wrangling and statistical computing in both Python and R.
- Deep understanding of cross-validation design, leakage prevention, stratification, and statistical rigor in benchmarking.
- Strong software engineering habits: testing, version control, code review, documentation, reproducibility.
- Clear communication of complex methods and results to technical and non-technical partners.
- Collaborative, mission-driven partner who thrives on shared impact.
- Curious self-starter who proposes new ideas, builds solid prototypes independently, and partners to productionize.
- Ability to relocate to the San Francisco Bay Area and work in the office at least 4 days per week.

## Preferred Skills:

- Proven experience with geospatial data analysis and environmental data tooling (xarray, GeoPandas, Rasterio, GDAL, STAC APIs), specifically in agricultural context; experience with satellite data feature extraction.
- Knowledge of GIS software (ArcGIS, QGIS)
- Experience with scalable data/compute: HPC (e.g., SLURM), distributed computing (Dask/Spark/Ray), and cloud platforms (GCP/AWS/Azure).
- Expertise with workflow orchestration and MLOps (Airflow/Prefect, MLflow/W&B, DVC), containers (Docker), and CI/CD.
- Experience working with APIs: design and build inference APIs, define and validate JSON payloads and versioned schemas
- Additional programming languages: Julia, C++, SQL; performance profiling and optimization.

## To Apply:

Email your application materials to applied-scientist@heritable.ag