[0005] To provide a few examples: hydrophobic interaction chromatography can be used to separate proteins based on differences in hydrophobicity, affinity chromatography can be used to separate molecules based on differences in affinity for a target ligand attached to a chromatography resin, and ion exchange chromatography can be used to separate molecules based on differences in molecular charge. As a more specific example, cation-exchange chromatography (CEX) is an ion exchange chromatography used when the molecule of interest is positively charged. Other common types of chromatography include size-exclusion chromatography (SEC), in which molecules in solution are separated by size and/or molecular weight, and Protein A chromatography.

[0006] In general, to ensure a robust commercial manufacturing process, it is important to characterize biological processes by, among other things, identifying how variability in the attributes of raw materials contributes to process performance and product quality. For various reasons, however, manufacturers/suppliers typically do not evaluate how raw materials manufactured at the edge of a certificate of analysis (or "CoA"), due to variability between manufacturing lots, can influence process consistency and product quality. Instead, the effects of raw material variability are evaluated, if at all, at target ranges using various risk-based analyses. Risk-based approaches of this sort may be insufficient for certain raw materials, such as the resin used as the stationary phase in column chromatography purification processes (e.g., the purification processes discussed above with reference to step 14 of FIG. 2). In particular, lot-to-lot variability of these raw materials can result in costly failure events, such as unacceptable purification performance indicators (e.g., excessive HCP levels) that require the discarding, or additional processing, of drug batches.

SUMMARY

[0006a] According to a first aspect of the present invention, there is provided a method of selecting raw materials for use in a column chromatography purification process, the method comprising:

for each of one or more candidate resins, receiving, by one or more processors of a computing system, a respective set of resin attribute values, the respective set of resin attribute values including at least one analytical measurement of the candidate resin;

for each of the one or more candidate resins, predicting, by the one or more processors applying the respective set of resin attribute values as inputs to a multivariate statistical model, a respective value of a performance indicator for the column chromatography purification process, wherein the multivariate statistical model includes one or more of a projection on latent structures (PLS) model, a regression tree model, an elastic net model, a lasso model, a ridge model, and/or a support vector machine (SVM) model;

selecting a resin of the one or more candidate resins based at least in part on the one or more predicted respective values of the performance indicator; and

performing the column chromatography purification process using the selected resin as a stationary phase.

[0006b] According to a second aspect of the present invention, there is provided a system comprising:

a computing system that includes one or more processors and one or more memories, the one or more memories storing instructions that, when executed by the one or more processors, cause the computing system to, for each of one or more candidate resins

receive a respective set of resin attribute values, the respective set of resin attribute values including at least one analytical measurement of the candidate resin,

predict, by applying the respective set of resin attribute values as inputs to a multivariate statistical model, a respective value of a performance indicator for a column chromatography purification process, wherein the multivariate statistical model includes one or more of a projection on latent structures (PLS) model, a regression tree model, an elastic net model, a lasso model, a ridge model, and/or a support vector machine (SVM) model, and

display the respective value of the performance indicator, or a result based on the respective value of the performance indicator, to a user to facilitate selection of a resin of the one or more candidate resins based at least in part on the one or more predicted respective values of the performance indicator; and

a column chromatography system configured to perform the column chromatography purification process using the selected resin as a stationary phase.

[0006c] According to a third aspect of the present invention, there is provided a non-transitory computer-readable medium storing instructions that, when executed by one or more processors or a computing system, cause the computing system to, for each of one or more candidate resins:

receive a respective set of resin attribute values, the respective set of resin attribute values including at least one analytical measurement of the candidate resin;

predict, by applying the respective set of resin attribute values as inputs to a multivariate statistical model, a respective value of a performance indicator for a column chromatography purification process, wherein the multivariate statistical model includes one or more of a projection on latent structures (PLS) model, a regression tree model, an elastic net model, a lasso model, a ridge model, and/or a support vector machine (SVM) model; and

display the respective value of the performance indicator, or a result based on the respective value of the performance indicator, to a user to facilitate selection of a resin of the one or more candidate resins, for use as a stationary phase in the column chromatography purification process, based at least in part on the one or more predicted respective values of the performance indicator.

[0007] Embodiments described herein relate to systems and methods that facilitate the selection of a resin for the stationary phase of a column chromatography purification process when manufacturing a therapeutic protein, such as a monoclonal antibody ("mAb"), or a bispecific or other multi-specific antibody, for example. In these embodiments, a multivariate statistical model enables the selection of resins (e.g., the selection of specific resin lots) that will not degrade (or overly degrade) performance of the purification process, by accounting for variability (e.g., lot-to-lot variation) in resin manufacturing. The multivariate statistical model predicts a performance indicator, such as a level of HCP and/or one or more other impurities, for a column chromatography purification process (e.g., a CEX, SEC, Protein A, or other suitable chromatography process that uses a resin as the stationary phase), based on various resin attributes and possibly one or more other types of inputs (e.g., harvest filtrate loading material factors and/or chromatography process parameters). The resin attributes may be provided by the manufacturer within a Certificate of Analysis (CoA), for example.

[0008] Resin "selection" may refer to selecting one or more resins out of multiple candidate resins, or confirming whether a single candidate resin is acceptable for use (i.e., "screening" the candidate resin prior to commercial-scale use). For example, resin lots received from a supplier may be screened using the multivariate statistical model and CoA data provided by the manufacturer, to determine which lots are acceptable and which lots should be rejected/replaced (or will necessitate further purification steps to meet requirements, etc.). As another example, specific resin lots may be selected (e.g., ordered) in the first instance based on which lots provide the most clearance relative to acceptability thresholds (e.g., by choosing the resin lots for which the multivariate statistical model predicts the lowest HCP levels). As still other examples, resins from different manufacturers, and/or different types or formulations of resins, may be selected by applying different, corresponding sets of resin attribute values to the multivariate statistical model.

[0009] Using techniques such as these, the amount of drug substance that must be rejected/discarded, and/or the amount of time and other resources needed to ensure acceptable purification performance, may be substantially reduced. For example, the

What is claimed is:

1. A method of selecting raw materials for use in a column chromatography purification process, the method comprising:

for each of one or more candidate resins, receiving, by one or more processors of a computing system, a respective set of resin attribute values, the respective set of resin attribute values including at least one analytical measurement of the candidate resin;

for each of the one or more candidate resins, predicting, by the one or more processors applying the respective set of resin attribute values as inputs to a multivariate statistical model, a respective value of a performance indicator for the column chromatography purification process, wherein the multivariate statistical model includes one or more of a projection on latent structures (PLS) model, a regression tree model, an elastic net model, a lasso model, a ridge model, and/or a support vector machine (SVM) model;

selecting a resin of the one or more candidate resins based at least in part on the one or more predicted respective values of the performance indicator; and

performing the column chromatography purification process using the selected resin as a stationary phase.

2. The method of claim 1, wherein predicting the respective value of the performance indicator further includes applying one or more harvest filtrate parameter values as inputs to the multivariate statistical model.

3. The method of claim 1, wherein predicting the respective value of the performance indicator further includes applying one or more purification process parameter values as inputs to the multivariate statistical model.

4. The method of claim 1, wherein predicting the respective value of the performance indicator includes applying (i) the respective set of resin attribute values, (ii) one or more harvest filtrate parameter values, and (iii) one or more purification process parameter values as inputs to the multivariate statistical model.

5. The method of any one of claims 1 through 4, wherein predicting the respective value of the performance indicator includes predicting a level of host cell protein resulting from the column chromatography purification process.

6. The method of any one of claims 1 through 5, wherein selecting the resin includes comparing the one or more respective values of the performance indicator to a predetermined acceptability threshold.

7. The method of any one of claims 1 through 6, wherein:

predicting the respective value of the performance indicator includes predicting a respective range of values of the performance indicator; and

selecting the resin includes comparing each of the one or more respective ranges of values to a predetermined acceptability threshold.

8. The method of any one of claims 1 through 7, wherein predicting the respective value of the performance indicator includes applying the respective set of resin attribute values as inputs to the PLS regression model.

9. The method of any one of claims 1 through 8, wherein, for each of the one or more candidate resins, receiving the respective set of resin attribute values includes receiving resin attribute values provided by a manufacturer or supplier of the candidate resin.

10. The method of any one of claims 1 through 9, wherein:

the one or more candidate resins include a plurality of candidate resins that correspond to different manufacturing lots of a single resin type.

11. The method of any one of claims 1 through 10, wherein:

the column chromatography purification process is a commercial-scale chromatography purification process; and

the method further comprises training the multivariate statistical model using historical small-scale and commercial-scale chromatography purification process data.

12. A system comprising:

a computing system that includes one or more processors and one or more memories, the one or more memories storing instructions that, when executed by the one or more processors, cause the computing system to, for each of one or more candidate resins

receive a respective set of resin attribute values, the respective set of resin attribute values including at least one analytical measurement of the candidate resin,

predict, by applying the respective set of resin attribute values as inputs to a multivariate statistical model, a respective value of a performance indicator for a column chromatography purification process, wherein the multivariate statistical model includes one or more of a projection on latent structures (PLS) model, a regression tree model, an elastic net model, a lasso model, a ridge model, and/or a support vector machine (SVM) model, and

display the respective value of the performance indicator, or a result based on the respective value of the performance indicator, to a user to facilitate selection of a resin of the one or more candidate resins based at least in part on the one or more predicted respective values of the performance indicator; and

a column chromatography system configured to perform the column chromatography purification process using the selected resin as a stationary phase.

13. The system of claim 12, wherein predicting the respective value of the performance indicator further includes applying one or more harvest filtrate parameter values as inputs to the multivariate statistical model.

14. The system of claim 12, wherein predicting the respective value of the performance indicator further includes applying one or more purification process parameter values as inputs to the multivariate statistical model.

15. The system of claim 12, wherein predicting the respective value of the performance indicator includes applying (i) the respective set of resin attribute values, (ii) one or more harvest filtrate parameter values, and (iii) one or more purification process parameter values as inputs to the multivariate statistical model.

16. The system of any one of claims 12 through 15, wherein the respective value of the performance indicator is a level of host cell protein resulting from the column chromatography purification process.

17. The system of any one of claims 12 through 16, wherein predicting the respective value of the performance indicator includes predicting a respective range of values of the performance indicator.

18. The system of any one of claims 12 through 17, wherein the multivariate statistical model includes the PLS regression model.

19. The system of any one of claims 12 through 18, wherein, for each of the one or more candidate resins, receiving the respective set of resin attribute values includes receiving resin attribute values provided by a manufacturer or supplier of the candidate resin.

20. The system of any one of claims 12 through 19, wherein the one or more candidate resins include a plurality of candidate resins that correspond to different manufacturing lots of a single resin type.

21. A non-transitory computer-readable medium storing instructions that, when executed by one or more processors or a computing system, cause the computing system to, for each of one or more candidate resins:

receive a respective set of resin attribute values, the respective set of resin attribute values including at least one analytical measurement of the candidate resin;

predict, by applying the respective set of resin attribute values as inputs to a multivariate statistical model, a respective value of a performance indicator for a column chromatography purification process, wherein the multivariate statistical model includes one or more of a projection on latent structures (PLS) model, a regression tree model, an elastic net model, a lasso model, a ridge model, and/or a support vector machine (SVM) model; and

display the respective value of the performance indicator, or a result based on the respective value of the performance indicator, to a user to facilitate selection of a resin of the one or more candidate resins, for use as a stationary phase in the column chromatography purification process, based at least in part on the one or more predicted respective values of the performance indicator.

22. The non-transitory computer-readable of claim 21, wherein predicting the respective value of the performance indicator further includes applying one or more harvest filtrate parameter values as inputs to the multivariate statistical model.

23. The non-transitory computer-readable of claim 21, wherein predicting the respective value of the performance indicator further includes applying one or more purification process parameter values as inputs to the multivariate statistical model.

24. The non-transitory computer-readable of claim 21, wherein predicting the respective value of the performance indicator includes applying (i) the respective set of resin attribute values, (ii) one or more harvest filtrate parameter values, and (iii) one or more purification process parameter values as inputs to the multivariate statistical model.

25. The non-transitory computer-readable of any one of claims 21 through 24, wherein the respective value of the performance indicator is a level of host cell protein resulting from the column chromatography purification process.

26. The non-transitory computer-readable of any one of claims 21 through 25, wherein the multivariate statistical model includes the PLS regression model.