

Claims

1. A method for adjusting the glucose concentration to a target value during the mammalian cell cultivation, comprising the following steps
 - 5 a) determine the current values at least for the process variables 'Time', 'CHT.PV', 'ACOT.PV', 'FED2T.PV', 'GEW.PV', 'CO2T.PV', 'ACO.PV', 'AO.PV', 'N2.PV', 'LGE.PV', 'CO2.PV', 'FED3T.PV', 'OUR' and 'PH.PV' in the cultivation,
 - b) determine the current glucose concentration in the cultivation medium using the measured values from a) by means of a data-driven model for
10 the mammalian cell cultivation, which was generated using a feature matrix comprising the process variables 'Time', 'CHT.PV', 'ACOT.PV', 'FED2T.PV', 'GEW.PV', 'CO2T.PV', 'ACO.PV', 'AO.PV', 'N2.PV', 'LGE.PV', 'CO2.PV', 'FED3T.PV', 'OUR' and 'PH.PV',
 - 15 and
 - c) adding glucose until the target value is reached if the current glucose concentration from b) is lower than the target value, and thus adjusting the glucose concentration to a target value.
- 20 2. The method according to claim 1, characterized in that the process variable(s) is/are selected from the group comprising the process variables viable cell density, viable cell volume, glucose concentration in the cultivation medium, and lactate concentration in the cultivation medium.
- 25 3. The method according to one of claims 1 through 2, characterized in that the method is carried out without sampling and exclusively using on-line measured values from this cultivation.
4. The method according to one of claims 1 through 3, characterized in that the data-driven model is generated by means of machine learning.
5. The method according to one of claims 1 through 4, characterized in that the data-driven model is generated with the random forest method.

6. The method according to one of claims 1 through 5, characterized in that the data-driven model is generated with a training dataset, which comprises at least 10 cultivation runs.
7. The method according to any one of claims 1 through 6, characterized in that
 - 5 a) the datasets available for modeling are randomly divided into training and test datasets in a ratio between 70:30 and 80:20,
 - b) the model is formed,
 - c) the mean value and the standard deviation for the determination of the process variable for the datasets from the training dataset is determined and the mean value and the standard deviation for the determination of the process for the datasets is determined from the test dataset,
 - 10 d) steps a) to c) are repeated until comparable mean values and standard deviations with regard to the division between test and training dataset are achieved, wherein the division obtained under a) is different with each new run.
 - 15
8. The method according to one of claims 1 through 7, characterized in that the datasets, which are used to generate the data-driven model, each contain the same number of data points.
9. The method according to one of claims 1 through 8, characterized in that the data points in the datasets, which are used to generate the data-driven model, are each for the same times of the cultivation.
- 20 10. The method according to one of claims 1 through 9, characterized in that missing data points in the datasets are supplemented by interpolation.
11. The method according to claim 10, characterized in that missing data points for the glucose concentration and/or viable cell volume are obtained by a third degree polynomial fit, that missing data points for the lactate concentration are obtained by univariate spline fit, and/or that missing data points for the viable cell density can be obtained through a Peleg fit.
- 25 12. The method according to one of claims 1 through 11, characterized in that the datasets contain a data point at least every 144 minutes.
- 30

- 55 -

13. The method according to any one of claims 1 through 12, characterized in that the mammalian cell is a CHO-K1 cell.
14. The method according to any one of claims 1 through 13, characterized in that the mammalian cell expresses and secretes an antibody.
- 5 15. The method according to any one of claims 1 through 14, characterized in that the data-driven model is generated with a training dataset that contains complex IgG cultivation runs and standard IgG cultivation runs.
16. The method according to any one of claims 1 through 15, characterized in that the cultivation volume is 300 mL or less.