
Algae Income Simulation Model: AISIM

**Research Report 12-4
December 2012**



Agricultural and Food Policy Center



Department of Agricultural Economics
Texas AgriLife Research
Texas AgriLife Extension Service
Texas A&M University

College Station, Texas 77843-2124
Telephone: (979) 845-5913
Fax: (979) 845-3140
<http://www.afpc.tamu.edu>

Algae Income Simulation Model: AISIM

AFPC Research Report 12-4

James W. Richardson
Myriah Johnson



**Agricultural and Food Policy Center
The Texas A&M University System**

Agricultural and Food Policy Center
Department of Agricultural Economics
Texas AgriLife Research
Texas AgriLife Extension Service
Texas A&M University

December 2012

College Station, Texas 77843-2124
Telephone: (979) 845-5913
Fax: (979) 845-3140
Web Site: <http://www.afpc.tamu.edu/>

Algae Income Simulation Model: AISIM

AISIM is a Monte Carlo firm level simulation model designed to simulate the annual activities of an algae farm. The model is designed to facilitate researchers analysis of the economic returns and costs of production for an algae farm under alternative management systems. The model can be thought of as a compilation of many techno-economic models of different phases of an algae farm. This appendix presents an overview of AISIM.

Programming

AISIM is programmed in Microsoft® Excel and depends upon the Simetar© add-in. The Excel workbook model is divided up into multiple worksheets that include: Input, Model, SimData, Prices, and others.

All input for an algae farm is entered in the INPUT worksheet and most all calculations are in the MODEL worksheet. Simetar is used to simulate the model by drawing annual stochastic prices, production, and costs from known probability distributions. Simetar functions are used to estimate parameters for the probability distributions to simulate the stochastic variables. The parameters for price probability distributions are estimated from historical data provided as input by the researcher. The best fitting distribution for simulating production and input prices is the multivariate empirical distribution (MVEMP).

The AISIM model is simulated recursively for 10 years. This means that the ending cash position of the business in year 1 is the beginning cash flow position for year 2, and so on. The 10 year planning horizon is repeated 500 times (iterations) using different stochastic prices and production values for each year. By simulating the 10 year planning horizon for 500 iterations, the model is able to simulate most all combinations of the stochastic variables, (i.e., the best and worst cases and those in between) based on their respective probabilities of being observed.

Input Requirements

The analyst must enter all of the data to describe the scenario to simulate for a farm. This includes data for the type of cultivation, harvesting, extraction, and co-products. A base scenario can be defined and copied multiple times with slight variations in the many management control variables. Simetar can then simulate all of the scenarios at once using the same risk for all of the stochastic variables. In this manner one can be guaranteed that the scenarios can be compared directly and that the only differences between scenarios is the input data changes. In the subsequent sections we describe the types of data required as input for AISIM.

- **Price Projections**

The annual average prices for soybean meal, electricity, nitrogen, phosphorous, CO₂, diesel, and other inputs are entered, as well as any discount or premium an algae co-product would have relative to these products. Annual values must be consistent with each other so it is recommended that the price projections all come from a common source, such as the FAPRI or USDA Baseline. In addition, annual rates of inflation for input costs and interest rates are

required. These data are used to calculate annual increases in input costs not associated with the stochastic prices.

- **Historical Price Data**

Historical prices for the stochastic price variables must be entered as input to AISIM. The model requires 10 years of historical prices for the stochastic price variables. AISIM uses the historical price data to calculate the parameters to simulate prices as a multivariate empirical distribution. Entering new historical data causes AISIM to automatically update the MVEMP distribution's parameters.

- **Options for Farming System**

AISIM is programmed to accommodate many different production systems. The production system options are listed at the beginning of the model. They are listed as yes/no or number options and trigger many different formulas within the model. The first option is the source of weather history data. Currently three weather history areas are available: Pecos and Corpus Christi, TX and Southeast New Mexico. These are used to calculate annual evaporation and precipitation. The second option is the source of the biomass production data; Pacific Northwest National Laboratory (PNNL) pre-loaded production data or specify your own growth rates. If the user chooses PNNL they are still be able to select a location. If the user chooses to use their own data they must specify whether those growth rates will be in annual (grams/liter), annual (grams/meter²), monthly (grams/liter), and at what pond concentration they will harvest each time (grams/liter). There is also the option to have a batch or continuous growth and harvesting process. Additionally, there is the option to choose whether water will or will not be recycled. Water cost can be calculated using the energy cost for pumping the water from a well for example, or the user can specify a \$/gal pumping cost.

The next set of options deals with the harvesting and extraction systems. The user can specify yes or no for each harvesting and extraction process depending on whether it is used. The user can also use more than one harvesting and extraction system, though they must remember to later specify what percent of the annual production is processed by the harvesting or extraction machine. The user can also specify whether the extraction will take place off-site or not. In the operating cost (OPEX) section the user specifies a \$/ton of biomass extraction cost. The last set of options for the model deals with the final products. The farm can produce multiple products, but the user must be cognizant that some of the co-products cannot be in production together or sold at the same time. For example, if the farm produces diesel it cannot sell oil because the oil has been used to produce diesel. Also, in order for fertilizer to be produced LEA must first be produced.

- **Algae Production Information**

The model simulates 10 years of algae production for the 500 iterations using monthly biomass probability distributions or the monthly biomass yield distributions from the BAT model at PNNL. If the analyst opts to simulate biomass production based on monthly probability distributions, the model requires the parameters for 12 monthly GRKS distributions, one for each

month. The GRKS distribution requires the parameters: the mode, the minimum (2.5% quantile) and the maximum (97.5% quantile). AISIM uses the monthly GRKS distributions to simulate monthly biomass production values based on biomass equations described in a subsequent section.

A second option is to use the 30 years of simulated biomass production from the BAT model. The researcher must indicate the region where the farm is located and then AISIM simulates biomass from the appropriate BAT file. The 30 years of monthly BAT biomass production are used to calculate parameters for a MVEMP distribution. The MVEMP is simulated for the necessary 12 months over 500 iterations and the random variables are used in the model.

- **Lipid Production**

A GRKS distribution parameterized by the user is used to simulate average annual percent lipid value. The average percent lipid is multiplied by the total biomass to calculate total lipid production.

- **Debt Financing Information**

The user must specify their financial and debt financing information. This information is used in ProForma Financial statements. Data needed are: dividends to the investor as a fraction of net cash income, dividends as a fraction of initial equity, discount rate, cash reserve for the first marketing year, fraction of the year for an operating loan, capital expenses (CAPEX) inflation rate, CAPEX loan length, the year the CAPEX loan is started, fraction of CAPEX financed, fraction of machinery replacements financed, number of years for machinery replacement loans, and the interest rate for machinery replacement loans.

- **Cultivation**

The cultivation input section of the model outlines the basis of the farm and is a key piece in the production and sizing of equipment for the model. First, the user must specify the desired acre feet of water for the algae production facility. Next, the number of blocks of ponds and number of harvested ponds per block must be specified. This option is provided for users to specify harvesting as a batch process. If the farm has a continuous harvesting process users specify their total number of ponds in the number of blocks of ponds entry, and the number of harvested ponds per block is specified as one. For a batch process the user specifies the number of blocks of ponds and the number of harvested ponds in each block. The number of blowdown ponds must also be specified. To accommodate an ARID raceway type situation (University of Arizona), the acres per pond is split into two categories: acres per pond and acres per trench. In the ARID raceway system the water flows through the pond and trench during the day, but is all put in the trench during the nighttime. With this system there are different exposed surface areas and, thus, different amounts evaporation during the daytime versus the nighttime. The acres per pond, along with the specified depth of water in the pond in inches, is used to calculate the total acre feet of ponds for the facility and thus is used as a check against what the user specified.

A pond concentration (g/L) at harvest must also be entered. This value is used to calculate the quantities of water and algae that move through the system. In conjunction with the accommodations for the ARID Raceway system the user must also specify the percent of daily evaporation that occurs during the daytime and at night. Lastly, a number of days of operation per year must be entered.

- **Land Area**

The user must also specify the number of acres needed just for ponds and then for the whole facility. Land not for ponds is used for space between the ponds, harvesting and extraction equipment, office or storage buildings, and anything else that might be needed.

- **Harvesting**

AISIM has three alternative harvesting systems: centrifuge, polymer flocculation, and electrocoagulation. The model allows for each of the systems to be used as the sole harvesting system or to be used in combination. The user must specify what percent of the annual algae production will be harvested by each harvesting system. CAPEX and OPEX associated with each process, such as electricity, chemicals, and maintenance must be entered by the user. In addition, the user must also specify the throughput capacity (L/hr), harvest time (hours/day), effective recovery rate of harvested algae (%), and percent solids of the algae output (%). These parameters are used to determine the number of harvesting units needed for the facility. Effective recovery rate also allows the biomass actually harvested to be tracked so the extraction machinery can be sized appropriately.

- **Harvesting Method – Centrifuge**

The centrifuge is a piece of equipment for harvesting that has been used in many different industries. Because of its well established nature it is often used as a base scenario to compare to alternative harvesting methods. The centrifuge uses centrifugal force to cause the algae and water to separate from one another. Water from the centrifuge can be recycled. Information needed for programming, as well as data for the centrifuge harvesting method, was obtained from Dr. Ron Lacey in the Agricultural Engineering Department at Texas A&M University. Data given by Dr. Lacey was for an “at-scale” algae production facility with 1,000 acre feet of water. All data needed for the centrifuge process is common to the alternative harvesting methods and was listed above.

- **Harvesting Method – Polymer Flocculation**

The polymer flocculation method of harvesting is also a process that has been used extensively in other industries. In polymer flocculation a chemical is added to the algae media which causes the algae particles to “floc” or clump together. Once the algae has concentrated it can be separated from the water for further processing. Water from the flocculation process can be recycled. All data needed for the polymer flocculation process is common to the alternative harvesting methods and was listed above. Data for flocculation are based on trials at Pecos, Texas, and was provided by Mr. Lou Brown.

➤ **Harvesting Method – Electrocoagulation**

Electrocoagulation harvesting utilizes aluminum ions, which are released during electrolysis between two aluminum plates submerged in the media. The positively charged aluminum ions attract the negatively charged algae and create a floc. The floc then settles to the bottom of the media and is separated by decanting the clear media from the top, which can be recycled to the cultivation process. The algae are concentrated from approximately 1 g/L in cultivation to around 8% solids (80 g/L) in the sediment. The capacity of the electrocoagulation unit is determined by the size of the plates and the volume of the reactor chamber. Information for modeling of this process and data for the model was provided by Dr. Lacy and by Kaselco, via Dr. Lacey. All data needed for the electrocoagulation process is common to the alternative harvesting methods and was listed above.

• **Extraction**

In AISIM, there are three different extraction options to choose from: Solution Recover Services (SRS), Hydro-Thermal Liquefaction – Catalytic Hydro-thermal Gasification (HTL-CHG), and Pyrolysis. The parameters for the extraction system are more specific than harvesting to each process. However, parameters that remain constant between processes are: throughput capacity (tons/day) or (liters/hour), and extraction (hours/day) to determine the number of units needed, along with capital cost (\$/unit), life of the machine (years), and annual maintenance cost.

➤ **Extraction Method – SRS**

In the SRS process, harvested algae biomass in water is pretreated with conditioning chemicals at precise pH, temperature and residence times followed by non-polar solvent addition and proprietary extracting techniques. This conditioning step hydrates and chemically solubilizes the algae cell to enable non polar solvent to remove lipids. The extracted algal mass is then phase separated and lipids are recovered from the non-polar solvent by distillation. Value added products are derived from processed biomass while solvent, water and conditioning chemicals are recycled. The oil is further processed for fuel production.

In addition to the extraction parameters previously listed above the following are also required for SRS: effective extraction (%), electricity usage (kwh/kg of biomass processed) or (kwh/ton biomass processed), chemical cost (\$/ton of chemical), units of chemical/unit of biomass processed, and the efficiency of the chemicals. Similar to the harvesting processes, the percent effective extraction is used to determine the final amount of output leaving the SRS system that will be available to sell. Electricity is tracked from this process into the overall electricity consumption as a part of the operating costs. AISIM can accept electricity consumption in either kwh per kg or kwh per ton of biomass processed to offer the user some flexibility. The cost per ton entered for chemicals is inflated in the model using a stochastic inflation rate over the 10 years.

➤ **Extraction Method - HTL-CHG**

HTL-CHG is a process in which algae biomass is transformed into oil, methane and electricity. Other output streams from this process include phosphate, CO₂, water, and other nutrients that can be recycled back to the ponds. The harvested algae are brought in and are made into a slurry for the process. The algae slurry is heated and pressurized. The output from this step is an oil formation consisting of oil and effluent water, along with a solid precipitate, phosphate. The phosphate can be recovered and sent for remake, after which it can be recycled to the ponds. Once the oil and water are separated the oil can be upgraded just like crude oil. The effluent water is then sent on to CHG for further processing. Once the effluent water reaches the CHG process it is again made into a slurry that is heated and pressurized to produce a liquid and precipitate, phosphorus, that can be processed and recycled. The liquid is then combined with a catalyst in the gasifier to produce CH₄ and CO₂. This CH₄ can be sold or turned into electricity, heat or CNG. The CO₂ can be recycled to the ponds and the remaining gas/water mixture can be further separated to obtain the CO₂ from the water and nutrients mixture, all of which can then be recycled to the ponds.

The two different parts of this process require information. The information that has been used for this process is from Genifuel. The HTL system requires that the user enter an algae slurry preparation cost (\$/ton of algae). This cost is inflated over the 10 year horizon. Additionally, the user must enter a value for the HTL process, such as: electricity and natural gas as a percent of the system's crude oil output, the elemental composition of the algae, percent nitrogen and phosphorus and the phosphorus remake cost as a percent of new phosphorus cost. The last parameter needed from the user for the HTL part is the percent of the total oils extracted based upon the algae input.

The CHG process requires some of the same information as HTL, such as electricity and natural gas usage as percent of the system's methane output. A percent water loss value is needed to determine the amount of water that can be recycled to the ponds. The thermal efficiency of the electricity generator is used to determine the electricity production from methane per year. Lastly, a catalyst is used in the gasification step. As specified by the user, the catalyst has a life and must be replaced at the end of its life. Intermediately, the catalyst can be remanufactured to extend its life. The user must specify the cost of the initial catalyst load along with the cost of remanufacturing it and the probability of needing to remanufacture the catalyst in a given year.

➤ **Extraction Method - Pyrolysis**

Pyrolysis is an extraction process that chemically decomposes algae by heat in the absence of oxygen. In the pyrolysis process the harvested algae is dried to less than 10% moisture using heat from the combustion of syngas. Syngas is a by-product from the pyrolysis process and is recycled for the purpose of drying the algae. The pyrolyzed algae will produce syngas, char (which can be sold as a soil amendment), and bio-oil. The bio-oil can be separated in a decanter to split the aqueous and organic fractions. These organic fractions are then further processed and upgraded to produce fuel.

The information and parameters needed for modeling pyrolysis were provided by Dr. Sergio Capareda of the Agricultural Engineering Department at Texas A&M University. Parameters for the pyrolysis processes include extraction machine operation (hours/day), bio-oil yield (gallons/dry ton of algae processed) and char yield (pounds/dry ton of algae processed). Additionally, each of the following are needed: organics as a percent of the bio-oil, aqueous solution as a percent of the bio-oil, percentage of the aqueous fraction that is water, and the percentage of organics in the aqueous solution. An upgrade cost from the organic bio-oil to crude must also be provided (\$/gal).

There is also a startup cost for each pyrolysis unit. This is due to the fact that there isn't any syngas available to dry the first batch of algae. Thus, the user must specify what the startup cost (\$/unit) is, as well as the number of startups per year they plan to have. Lastly, general information such as capital cost (\$/unit), life of machine (years) and annual maintenance cost (\$) must be given.

- **CAPEX**

Values for CAPEX must be specified by the user, except for in the case of the harvesting and extraction equipment, gators, vehicles, pumps, and land which are populated based on information entered elsewhere in the model. CAPEX categories are: dirt moving construction, raceway construction, photo bio-reactor (PBR) tubes or bags, sump construction, liner, perimeter fence, dividers between ponds, paddlewheels, CO₂ delivery system, nutrient storage and distribution, piping system, algae inoculum stations, water wells, storage buildings, anaerobic digester, power generation, electrical lines, office building, backhoes, motor graders, ponds sweepers, lab building and equipment, field expenses, diesel plant capital cost, contingency costs, and other capital costs. CAPEX costs are summed and split into the appropriate machinery replacement categories based upon useful life. These summations are used to compute the values of the machinery replacement loans and cash flow requirements for maintaining current equipment.

- **OPEX**

There are several different operating cost categories. In each, the user may either enter a lump sum annual operating cost for that category or be more specific and enter information that will be used to calculate usage and costs.

- **CO₂**

Three pieces of information are needed for the CO₂ calculations. The first is the pounds of CO₂ required per pound of biomass produced and the second is the efficiency of the CO₂ (%). These are used in determining the total amount of CO₂ that is needed for the year. The third piece is the contract cost for CO₂ (\$/ton).

➤ **Media**

There are three options besides the annual media cost. The first is to enter a ratio of growth media to biomass and lipid media to biomass. The cost of the lipid and growth media (\$/liter) must be entered. Additionally, it must be specified what percentages of nitrogen, phosphorus, and potassium are in the solution. This is so the media cost can be appropriately inflated according to stochastic prices for these ingredients.

The second option is to enter a ratio of nitrogen, phosphorus, potassium and other nutrients to biomass and the cost of these inputs on a \$/ton basis. The final option requires the startup growth media and lipid media liters/batch. For this option, the number of growth media, lipid media, recycled growth media, and recycled lipid media batches per year must also be given along with the cost (\$/liter) of each. The percents of nitrogen, phosphorus, and potassium in each media mix must be specified to appropriately simulate the costs.

➤ **Labor**

In the labor section the user must specify the number of employees by category. The labor categories are: CEO, legal/accounting/permitting, project manager, operations manager, administrative assistant, procurement, marketing, cultivation, harvesting, extraction, aquatic biology, fisheries biology, lab technicians, and maintenance. For each category a salary must be given to calculate a total labor cost. The assumed annual salaries for the categories are: CEO, \$150,000, legal/accounting/permitting, \$40,000, project manager, \$60,000, operations manager, \$100,000, administrative assistant, \$42,500, procurement, \$106,000, marketing, \$84,000, cultivation, \$46,000, harvesting, \$60,000, extraction, \$70,000, aquatic biology, \$64,000, fisheries biology, \$58,000, lab technician, \$80,000, and maintenance, \$43,000. The number of employees by category is also used to determine the number of vehicles needed. The vehicle costs are then linked in with the CAPEX section. The cost per vehicle is: SUV, \$45,000, pickup, \$30,000, and gators, \$9,000.

➤ **Electricity**

Electricity for cultivation can be entered as a kwh/ton of biomass produced. Harvesting and extraction electricity is not entered in this section because they are included in each processes' respective sections. There is also an other electricity section where non-water, cultivation, harvesting, and extraction electricity can be entered in kwh/ton of biomass.

➤ **Waste Water Disposal**

Parameters for calculating the cost of blowdown ponds are listed in this section. These parameters are: pond depth (ft), acres of blowdown ponds, price of dirt removal (\$/cubic foot), area of liner needed (ft²), liner cost (\$/ft²), liner installation cost (\$/ft²) and the number and cost of pumps.

➤ **Natural Gas**

The amount of natural gas in cubic feet/ton of biomass is required.

➤ **Property Tax Rate**

The property tax percentage rate must be entered. The tax is calculated as a percentage of the total property value.

➤ **Workman's Compensation/Unemployment Tax**

The workman's compensation/unemployment tax percentage must be entered. The tax is calculated as a percentage of the total labor cost.

➤ **Non-Harvesting and Extraction Maintenance Costs and Crash Cleanup Costs**

All annual maintenance costs are entered here. The crash cleanup cost represents the cost of disposing of the contaminated algae and restarting a pond(s) if it were to crash.

➤ **Harvesting and Extraction Costs**

If any harvesting and extraction costs were not captured earlier in their respective sections they can be entered in this section as an annual cost. Alternatively, the cost of outsourcing extraction (\$/ton of biomass) can be entered.

• **Pro Forma Financial Statements**

AISIM calculates three pro forma financial statements: income statement, cash flow, and balance sheet. These statements summarize the values described above.

➤ **Income Statement**

The income statement provides separate lines for each source of revenue, such as: diesel, electricity, PUFAs, LEA, whole algae, etc. Thus the analyst can see where receipts are being generated and can make changes to input data as needed. The receipts in the statement are calculated earlier in the model and are summarized in this location.

The second part of the income statement has separate lines for each cash expense. The cash expenses include: nutrients, labor, fixed costs, electricity, etc. Total cash expenses are used to calculate operating interest costs based on the projected interest rate in the relevant year and total cash expenses.

Other interest costs included in the expenses section are interest for the initial loan for CAPEX and interest for cash flow deficit loans. The latter occurs when the business has insufficient cash reserves and net cash income to pay required cash outflows described in the

cash flow statement. If a cash flow deficit exists in year t , then the interest for a short-term loan against the deficit is calculated and included as an interest cost in year $t+1$.

Net cash income equals total receipts minus total cash costs and total interest expenses. Net income equals net cash income minus depreciation. For this calculation depreciation is calculated using a straight-line method.

➤ **Cash flow Statement**

The cash flow statement is divided into two parts: cash inflows and cash outflows. Cash inflows includes cash on hand January 1, net cash income for the year, and interest earnings on beginning cash reserves. Cash outflows include: investor dividends, principal payments, repayment of cash flow deficits, income taxes, and down payments for machinery and equipment replacement. These items are cash outlays but are not tax deductible so they do not appear in the income statement. The last line of the cash flow statement calculates the ending cash balance on December 31 as total inflows minus total outflows.

➤ **Balance Sheet**

The balance sheet is divided into assets and liabilities. The first asset is positive ending cash reserves for December 31. If ending cash is negative this value is zero. The ending cash value from this line is what becomes beginning cash reserves next year in the cash flow statement. Other assets include the market value of land, capital improvements, and machinery.

The liability section of the balance sheet shows the current balance for the original loan and the cash flow deficits if ending cash reserves are negative. Net worth is the final value in the balance sheet and equals assets minus liabilities.

Key Output Variables (KOVs)

The KOVs are the variables from the model that are sent to Simetar to collect during simulation and then calculate summary statistics after the last iteration. The KOVs include variables such as: net present value, rate of return on equity, annual net cash income, annual ending cash reserves, probability of positive ending cash reserves each year, present value of ending net worth, probability of increase in real net worth, annual cost of biomass, lipids, and diesel. Any other variables in the model can be included in the KOV table.

Net present value (NPV) is calculated as:

$$NPV = - \text{Beg Net Worth} + \sum (\text{dividends and cash withdrawals}_t / \text{discount factor}_t) + \text{Ending Net Worth} / \text{discount factor}$$

Present value of ending net worth is calculated of $PVENW = \text{Ending Net Worth} / \text{discount factor}$

The discount factors above are calculated as: $1 / (1 + \text{discount rate})^t$

Income Taxes

The income taxes for the business are calculated assuring the business is taxed as a corporation. The taxable income equals net cash income minus depreciation calculated based on IRS code for the reasonable life of each piece of machinery. The income tax rates in the IRS code for corporations are used directly. At this time there are no state income taxes being calculated for the model.

Model Output

Each time the input data changes the model must be re-simulated using Simetar. The simulation takes 3—45 seconds and the stochastic results are presented in the SimData worksheet. The results in SimData include summary statistics for each KOV and the 500 actual simulated values for each KOV.

Probability charts (PDFs, CDFs, Fan Graphs, and StopLight charts) can be developed from the 500 simulated values. Tables of the summary statistics for selected KOVs can be developed.

If the summary tables are developed using “cell reference” formulas relating back to the values in SimData, the tables will be updated automatically each time Simetar re-simulates the model. The charts developed using data in SimData will automatically update each time the model is re-simulated by Simetar.