**Socio-Ecological Indicators for Historical Agro-Ecosystems**

Sustainable Farm Systems Project

Draft July 2015

**The Three Dimensions of SFS Data**

The datasets created for the Sustainable Farm Systems Project have three basic dimesions: **indicator**, **location**, and **date**. The project’s goal is to understand the agro-ecological and social-metabolic characteristics of particular farm systems and how they changed through time over the past 250+ years. Research focusses on a number of case studies in North America, Latin America, and Europe, and we would like to be able to accommodate case studies from other parts of the world in the future. This diversity requires the system to be able to handle many different types of farming and to handle both modern and historical agricultural structures.

For each case study location we assemble a set of descriptive and analytical measures of farm activity, structured as thematic groups of **indicators**. For example, core sets of analytical indicators address land use structures, livestock systems, energy dynamics, soil nutrient dynamics, and landscape ecology. Other sets of indicators provide important socio-economic and environmental context, such as population, social structure, economy, transportation, and climate.

A second dimension of datasets reflects **location**. Research teams are developing indicators for particular case study locations. The scale of these locations varies from small villages to larger aggregations of farms within local jurisdictions (such as counties or rural municipalities), to regional- and provincial-scale aggregations. We also would like to be able to employ these measures for national-scale datasets as well as at the level of a single farm.

Since this is a historical study, each location has data for multiple **dates**. At the low end, we may have data for only two or three time points, while other case studies may have as many as dozens or even a hundred dates. Normally the temporal unit of analysis is one year—we do not calculate indicators for specific days, weeks, or months, but for a full annual farm cycle. Sometimes the year represented is approximate, for example, the farm structure as it existed in the 1820s; most of the time we can specify an exact year.

Thus datasets consist of three dimensions: multiple indicators for multiple locations in multiple years. A single data point will represent all three dimensions, for example the area planted in wheat (indicator) in Rooks County, Kansas (location) in 1910 (date).

**Data Hierarchy**

**1st Order Indicators**

* raw transcriptions from primary sources (such as cadastres and maps, account books, census reports, air photos) and scientific observations (soil survey, weather data)
* estimates for missing data (down-scaling from national or provincial scales, estimates from nearby or similar locations, interpolations from other time periods, etc.)
* examples: acres and bushels of crops; number of livestock; number of machines; soil type; amount of rainfall
* structure, completeness, and reliability varies considerably across case studies and through time
* can reveal structural differences between case studies or major gaps in information

**1st Order Indicators, Standardized**

* simple mathematical conversions into standard units
* examples: convert volume to weight (bushels to pounds) and all units to metric (pounds to kilograms; acres to hectares); convert livestock numbers to 500kg Livestock Units; convert daily rainfall amounts to annual total; use soil texture to estimate bulk density
* requires user-specified conversion factors that may vary across locations and time
* standardized units allow direct comparisons between case studies

**2nd Order Indicators**

* mid-level aggregated or simplified indicators using Material and Energy Flow Accounting (MEFA) measures for analytical purposes
* examples: productivity measures (crop yield, area productivity, labour productivity); energy profile components (LP, LBP, TP, FP, BR, ASI, EI, TIC, L, etc.); soil nutrient inputs (free fixation, symbiotic fixation, manure, seeds, fertilizer, etc.) and outputs (leaching, harvest, denitrification, ammonia volatilization, etc.); biomass appropriation measures (NPPact, NPPh, etc.); freshweight and dry matter biomass measures; feed balance measures
* these synthesized indicators allow a multitude of analytical and comparative studies, and may be the end point for some research programs and indicators. They also support graphic representation (flow charts, graphs, maps) with powerful communication benefits.

**2nd Order Indicators, Standardized**

* converted to biophysical units and normalized for area (GJ/ha, kg N/ha)
* requires user-specified conversion factors that may vary across locations and time

**3rd Order Indicators**

* headline sustainability measures
* examples: crop yield; N, P, K balance; EROIs; HANPP; GDP; crop diversity index; Shannon index; effective mesh size index

**Input Data**

Because of the complexity of original historical primary sources and the considerable differences between farm systems in distant places, the shape and structure of 1st order indicators is highly variable across locations and time points. This complexity makes standard input structures very difficult to specify. For this reason, we propose to begin software development at the level of 2nd order indicators, which will make up the main input datasets. Individual research teams will have to wrestle with their own primary sources and create 1st order indicators using spreadsheet or database software. Once they have calculated key 2nd order indicators they can then employ the new software for comparison, analysis, graphic representation, and automated generation of 3rd order indicators.

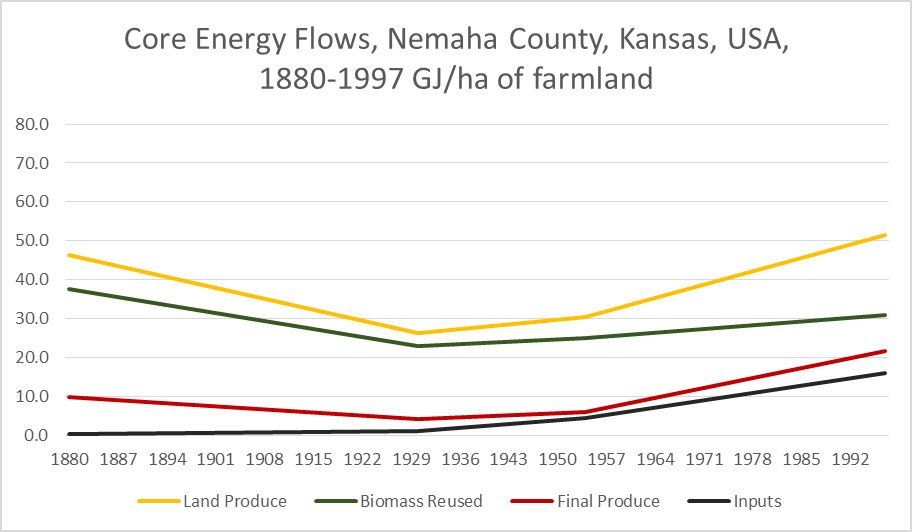
As a later phase of the software development project, we may identify particular subsets of 2nd order indicators that we all use frequently and whose calculation can be easily automated. For some themes there may be a high level of uniformity in primary sources and reporting across most case study locations that allows automatic generation of 2nd order indicators from 1st order inputs. For example, we might try to build a “freshweight biomass module” that takes individual crop production inputs, applies standard coefficients, and generates biomass totals. Such programming is for a later stage of the project, but could make the software much more practically useful for our teams and for outside researchers.

The attached spreadsheet contains a list of the 2nd order indicators that teams will supply, and the related 3rd order indicators to be calculated. It is imporant to note that many of the input datasets will not be complete. We may be able, for example, to input only the energy indicators for some case studies, without also inputing soil nutrient indicators. In this sense the software needs to be modular and flexible, allowing input of all or of only some indicators.

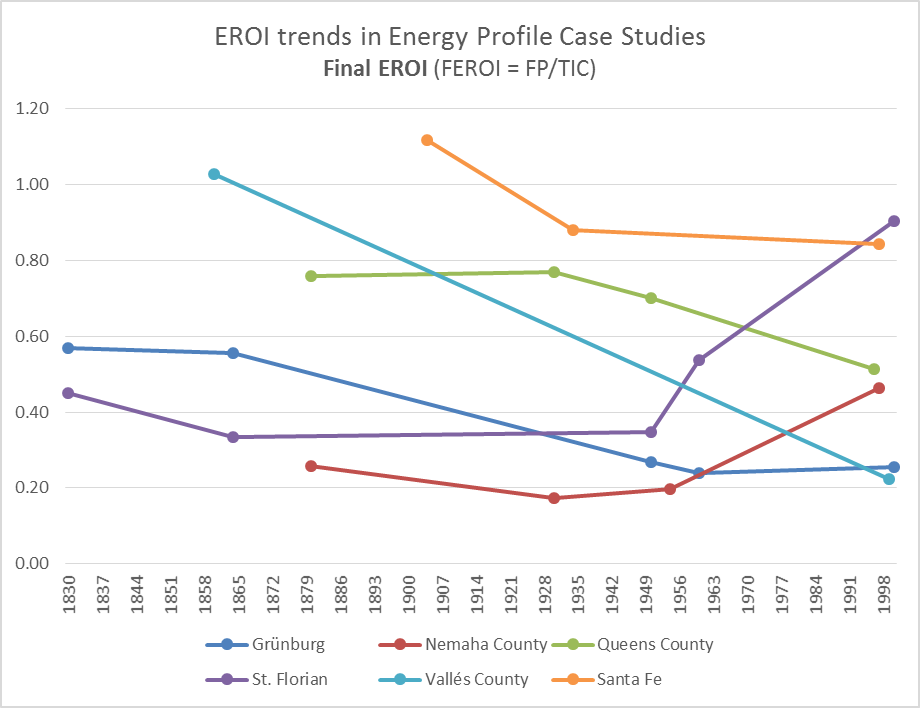
**Output Capabilities**

I. Basic graphing functions for each indicator

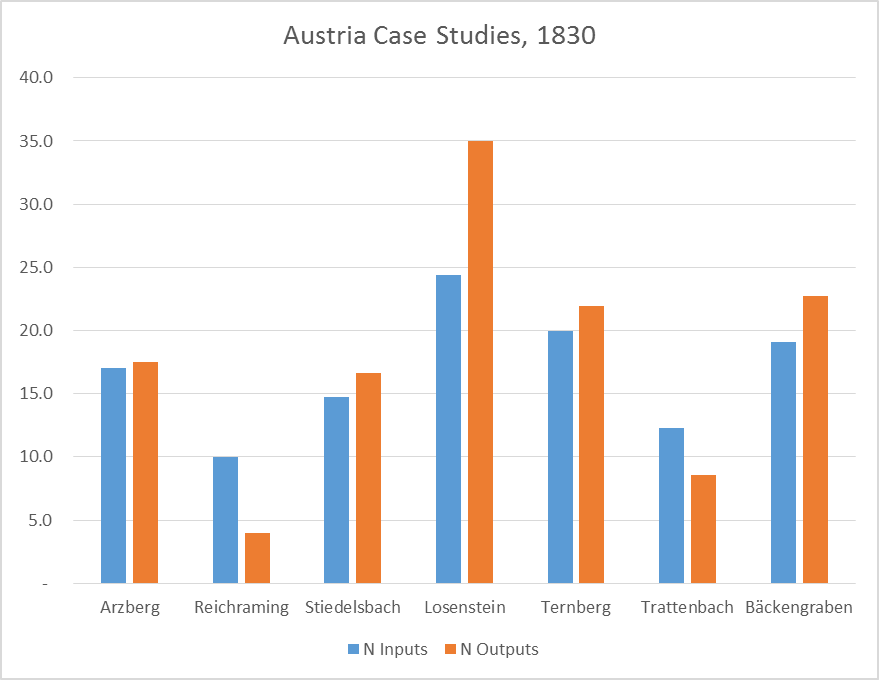
I. A. Multiple indicators, multiple time points, for 1 location:



I. B. Multiple locations, multiple time points, for 1 indicator:



I. C. Multiple locations, multiple indicators, for 1 time point:



II. Basic mathematical functions for multiple variables:

* a formula-building dialog that allows addition, subtraction, multiplication, and division for pairs or groups of data points that can then be graphed, as in section I.

III. Basic statistical measures for user-specified groups of data points:

* range, mean, median
* standard deviation
* regression relationships, multi-factor analysis, cluster analysis
* etc.

IV. Sankey-style flow charts for major analytical groups of variables:

* cyclical energy model
* linear soil nutrients model.
* default design values (colour, line width, text font, etc.) but also the option for user-specified design values

