

DISS. ETH NO. 18391

**ENVIRONMENTAL ASSESSMENT OF CHEMICAL PRODUCTION UNDER  
DATA-SCARCE CONDITIONS**

A dissertation submitted to

ETH ZURICH

for the degree of  
Doctor of Sciences

presented by

Gregor Peter Wernet

Dipl. Chem., Eberhard-Karls-Universität Tübingen

born on June 2<sup>nd</sup>, 1978

Citizen of Germany

accepted on the recommendation of

Prof. Dr. Konrad Hungerbühler

Prof. Dr. Alexander Wokaun

Prof. Dr. Stefanie Hellweg

Dr. Rolf Frischknecht

2009

## Summary

The production of chemicals is a source of environmental hazards. In addition to emissions directly from the plant, efforts to increase the sustainability of production today have led to an increased focus on the impacts of raw material use, energy use and waste treatment. Increased efficiency – often with economic benefits as a driving force – and stricter regulations have done much to improve the environmental profile of chemical production in developed countries. Nevertheless, with Global Warming becoming a major issue, new ways of assessing the impacts of chemical production are necessary.

Using a technique called Life Cycle Assessment (LCA), the full production of a chemical can be assessed, including all raw materials and emissions. This “cradle to factory gate” approach includes all processes and materials required, from solvents used in the plant up to the oil in the ground extracted to create the solvents. The result is an assessment of all environmental impacts caused by the production of the chemical, be it direct or indirect. First, an inventory of all mass and energy flows over the whole production is gathered, listing all raw materials used and emissions caused by the summation of all the processes required. These inventories are then assessed using one or several impact assessment methods. These methods can focus on single environmental issues, such as the Global Warming Potential (GWP) or resource use in the Cumulative Energy Demand (CED). Other methods assess multiple impact categories and aggregate their results into a single impact score using normalization and weighting schemes. For the results to be meaningful, data on the whole production cycle are required. This includes emissions, energy and ancillary uses, and all other mass and energy flows. Detailed production measurements are ideal, but often models have to be used. Databases of inventory data can be used to cover background and upstream processes. However, public inventory data are available for only a few hundred of the over 100 000 chemicals in production today.

The goal of this work is to create an algorithm for estimating these data and the resulting impacts in conditions where data are scarce, e.g. in situations where measurements may not be feasible or too costly, or where data are confidential and therefore not available. To assist this task, an analysis of many chemical productions is carried out to acquire information about the general composition of chemical production impacts.

To achieve this, basic chemical production impacts are compared to the impacts of advanced chemicals from a life cycle perspective in a first analysis. The environmental impacts of organic and inorganic chemicals are characterized according to various impact categories and the main contributors to the overall impacts are determined. The study shows that energy use is a major factor for the sustainability of chemical production in industrialized countries, mostly due to resource use and contributions to Global Warming and air pollution. Process emissions from the chemical plant are often less important, most likely due to modern emission controls and regulations in these countries. The analysis reveals that the relative contribution of energy-related impacts to the total impacts increases with the complexity of the chemicals. To confirm this observation, an analysis of the production of more complex chemicals is needed. Therefore, the production inventory of an active pharmaceutical ingredient (API), as a representative of highly complex chemistry, is compiled and the composition of its environmental profile analyzed. The total impacts per kilogram are more than 1 and sometimes 2 orders of magnitude higher than for a basic chemical. This is one of the first detailed and publicly available LCA studies of a pharmaceutical. Similar to the analysis of the basic chemicals, the contributions to the overall impacts were determined. The results concur with the previous finding that energy use is a major contributor to the environmental impacts of chemical production.

Energy use is therefore a critical part of chemical production, but energy use data are especially scarce. Tools to approximate energy use and environmental impacts are therefore needed. This problem is approached in this work with the creation of Molecular Structure-Based Models (MSMs). These models do not require data on the production process. Instead, they operate using only the molecular structure as input to deliver estimations of key production parameters. These models are not designed to replace detailed process analyses and process models but rather to complement them in cases of severe lack of data.

First, the functionality of the models is established. As MSMs are a new concept in the context of environmental impact estimation, their capabilities are tested on a dataset of 103 inventory datasets of basic organic chemicals. Several approaches are compared and neural networks are chosen as the algorithm for the models. This experiment confirms the feasibility of MSMs. The small dataset used limits the predictive capabilities of the models. To increase the effectiveness of the models, a co-operation with several partners from the chemical industry is established. The industry partners provide production data that are used to improve the models' capabilities as well as their range of chemicals covered. The improved models are

based on a dataset of 392 inventories and are able to estimate the CED, the GWP and Eco-indicator 99 scores, an aggregated impact assessment method. As the dataset used for the creation of these models is still somewhat limited, there are some classes of chemicals for which they are not yet applicable. However, due to their nature, the models can be continually improved as more and more inventory data become available. The MSMs can be used to perform screening LCAs and can be useful in cases where data are lacking. They are a valuable tool for industrial ecologists and process designers working on new processes or for assessing the sustainability of supply chains.

The models and results of this work provide the means to estimate aggregated production parameters and to assess chemical production even when data availability is low. They can therefore close gaps in the area of chemical production data which have long been a problem. The models can be used for environmental impact assessment, CO<sub>2</sub>-accounting and process design.

# Zusammenfassung

Bei der Produktion von Chemikalien können eine Vielzahl von umweltschädlichen Emissionen verursacht werden. Neben Emissionsanalysen im Chemiewerk selbst werden heutzutage zur Förderung der Nachhaltigkeit auch die Auswirkungen von Abfallbehandlung, Ressourcen- und Energieverbrauch untersucht. Effizienzsteigerungen in der Produktion – oft aus finanziellen Beweggründen – und strengere Regulierungen haben dazu beigetragen, dass sich das Umweltprofil der chemischen Produktion in Industriestaaten verbessert hat. Doch neue Herausforderungen wie der Treibhauseffekt fordern eine erweiterte Betrachtungsweise der Umweltfolgen, welche durch die chemische Industrie entstehen.

Mittels „Life Cycle Assessment“ (LCA oder Ökobilanzen) kann die Produktion von Chemikalien inklusive aller verbrauchten Rohmaterialien und aller Emissionen bewertet werden. Dieser „cradle to factory gate“-Ansatz beinhaltet alle benötigten Prozesse und Materialien; vom Lösemittel bis hin zum Öl, welches zur Herstellung des Lösemittels gefördert wurde. Das Ergebnis ist eine Bewertung aller direkt und indirekt durch die Produktion verursachten Umweltschäden. Hierzu wird zuerst ein Inventar aller Massen- und Energieflüsse in der Produktion zusammengestellt, für das der gesamte Ressourcenverbrauch und sämtliche Emissionen über alle Produktionsschritte hinweg betrachtet werden. Diese Inventare werden dann mit einer oder mehreren Bewertungsmethoden beurteilt. Manche dieser Methoden betrachten ausschliesslich einzelne Problembereiche, wie zum Beispiel die Emission von Treibhausgasen im „Global Warming Potential“ (GWP) oder den Ressourcenverbrauch im „Cumulative Energy Demand“ (CED). Andere Methoden wiederum haben mehrere Schadenskategorien und aggregieren diese durch Normalisierung und Gewichtung zu einem Einzelwert. All diese Methoden benötigen Informationen über die gesamte Produktion, inklusive Emissionen, Verbrauch von Energie und Hilfsstoffen und andere Massen- und Energieflüssen, um aussagekräftige Ergebnisse liefern zu können. Genaue Messergebnisse sind die beste Quelle für solche Daten, oft müssen allerdings Modelle zu Hilfe gezogen werden. Datenbanken mit Inventaren von Hintergrundprozessen können hilfreich sein; jedoch sind nur für wenige hundert der über 100 000 heute produzierten Chemikalien Inventardaten frei verfügbar.

Ziel dieser Arbeit ist es, eine Methode zu entwickeln, mit der Inventardaten und Umweltauswirkungen auch bei Datenmangel abgeschätzt werden können. Diese sollen in Situationen, in denen Messungen nicht machbar sind oder in denen kein

Zugang zu den Produktionsunterlagen des Herstellers möglich ist, hilfreich sein. Zur Unterstützung dieses Ziels wird eine Reihe von Chemikalienproduktionen, von denen Daten verfügbar sind, untersucht. So können die generelle Beschaffenheit und Zusammensetzung der Umweltauswirkungen von Chemikalienproduktionen bestimmt werden. In einer ersten Studie werden die Umweltauswirkungen der Produktion von Basischemikalien mit denen komplexerer Chemikalien verglichen. Organische und anorganische Chemikalien werden nach verschiedenen Kriterien bewertet und die grössten Beiträge zu den Umweltfolgen bestimmt. Die Studie zeigt, dass Energieverbrauch heute vor allem durch Ressourcenverbrauch sowie durch Beiträge zu Global Warming und Luftverschmutzung ein wichtiger Faktor in der Nachhaltigkeit von Chemikalienproduktion ist. Emissionen aus dem Werk selbst sind in fortschrittlichen Industrienationen dank moderner Emissionskontrollen und Regulierungen oft weniger relevant. Die Studie zeigt ausserdem, dass der relative Beitrag von Energieverbrauch zu den Gesamtumweltfolgen in solchen Staaten bei steigender Komplexität der Chemikalie zunimmt. Um diese Beobachtung zu bestätigen, ist in einer weiteren Studie das Produktionsinventar eines pharmazeutischen Wirkstoffes zusammengestellt und bewertet worden. Die Auswirkungen pro Kilogramm sind um 1 bis 2 Grössenordnungen höher als bei Basischemikalien. Dies ist eine der ersten detaillierten und frei erhältlichen LCA-Studien eines Pharmazeutikums. Die Ergebnisse bestätigen die Beobachtungen bei den Basischemikalien, dass der Energieverbrauch ein kritischer Parameter für die Nachhaltigkeit von Chemikalienproduktion ist.

Energieverbrauch ist demnach ein wichtiger Parameter, doch gerade Energiedaten sind besonders schwer zu bestimmen und seltener verfügbar als Massenflüsse. Dem Problem des Datenmangels bei der Erstellung von Chemikalieninventaren wird in dieser Arbeit mit der Erstellung Molekülstruktur-basierter Modelle (MSMs) begegnet. Diese Modelle benötigen keine Informationen über den Produktionsablauf. Stattdessen schätzen sie an Hand der Molekülstruktur wichtige Produktionsparameter ab. MSMs können Prozessanalysen und detailliertere Modelle nicht ersetzen, sondern ergänzen diese in Fällen, in denen keine Daten zu Verfügung stehen.

Zur Erstellung der MSMs wird zuerst eine Machbarkeitsstudie durchgeführt. Da MSMs ein neuartiges Konzept sind, ist ihre Leistungsfähigkeit an einem simplen Datensatz von 103 Inventaren getestet worden. Mehrere Ansätze werden verglichen und neuronale Netze als Basis für die Modelle gewählt. Die MSMs zeigen in dieser Studie grosses Potential bei der Vorhersage von Produktionsparametern, der kleine Datensatz limitiert jedoch die Leistung der Modelle. Zur Optimierung der Modelle

werden daher mehrere Partner aus der chemischen Industrie gewonnen. Diese haben Produktionsdaten zur Verfügung gestellt, mittels derer ein erweiterter Datensatz aus 392 Produktionsinventaren erstellt wird. Durch die vergrösserte Datenbasis können die Fähigkeiten der MSMs erhöht und der Anwendungsbereich der Modelle erweitert werden. Die Modelle können den CED, das GWP und Resultate für Eco-indicator 99, eine aggregierende Ökobilanz-Methode, abschätzen. Durch die immer noch begrenzte Zahl von Daten, die den Modellen zu Grunde liegen, sind diese noch nicht für alle Klassen von Chemikalien anwendbar. Sie können jedoch kontinuierlich erweitert werden, wann immer weitere Inventardaten verfügbar werden. MSMs können bei Datenmangel zur Durchführung von Ökobilanzen verwendet werden. Sie sind ein wertvolles Werkzeug für Umweltbeauftragte und Prozessentwickler, die neue Prozesse entwerfen oder die die Nachhaltigkeit ihrer Zulieferkette bewerten wollen.

Die Modelle und Resultate dieser Arbeit erlauben eine Abschätzung von aggregierten Produktionsparametern und eine Bewertung von Chemikalienproduktionen selbst dann, wenn keine direkten Produktionsdaten verfügbar sind. Das Problem des grossen Mangels an Chemikalieninventaren kann so gemindert werden und die Modelle erlauben eine Studie der vielen tausenden von Chemikalien, für die keine Inventardaten erhältlich sind. Die Modelle können bei Prozessdesign, CO<sub>2</sub>-Emissionsstudien und Ökobilanzen behilflich sein.