PROGRESS IN PHOTOVOLTAICS: RESEARCH AND APPLICATIONS
Prog. Photovolt: Res. Appl. 2008; **16**:441–453
Published online 26 November 2007 in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/pip.806

Broader Perspectives

Accuracy of Progress Ratios Determined From Experience Curves: The Case of Crystalline Silicon Photovoltaic Module Technology Development

W. G. J. H. M. van Sark[1*,†], E. A. Alsema[1], H. M. Junginger[1], H. H. C. de Moor[2] and G. J. Schaeffer[2]
[1]*Department of Science, Technology and Society, Copernicus Institute for Sustainable Development and Innovation, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, the Netherlands*
[2]*Energy Research Centre of the Netherlands, ECN, P.O. Box 1, 1755 ZG Petten, the Netherlands*

Learning curves are extensively used in policy and scenario studies. Progress ratios (PRs) are derived from historical data and are used for forecasting cost development of many technologies, including photovoltaics (PV). Forecasts are highly sensitive to uncertainties in the PR. A PR usually is determined together with the coefficient of determination R^2, which should approach unity for a good fit of the available data. Although the R^2 is instructive, we recommend using the error in the PR determined from the fit because it is a direct measure of the range in PR values that is recommended to be used in sensitivity analyses within scenario studies. We present a simple equation to calculate the error in PR from the fit parameters. In the case of crystalline PV module technology development we find a PR $= 0{\cdot}794 \pm 0{\cdot}003$ by fitting price data of the period 1976–2006. A moving average approach with a 10-year time window shows that PR varies from $0{\cdot}818 \pm 0{\cdot}017$ up to a starting year of 1987, and is reduced considerably to a minimum value of $0{\cdot}704 \pm 0{\cdot}014$ for the starting year 1991. For the most recent starting year 1997, the average PR is considerably higher at $0{\cdot}884 \pm 0{\cdot}022$, highlighting the recent silicon feedstock supply problem. When available, error in individual data points can be used to perform weighted fits in order to decrease fitting errors. To illustrate this approach, an analysis of Dutch PV system price development over the period 1992–2002 shows that PR is $0{\cdot}876 \pm 0{\cdot}010$, where the error is decreased with respect to unweighted fitting. The PR $= 0{\cdot}794$ has been used to analyze the cost targets stated in the Strategic Research Agenda as formulated by the European PV Technology Platform for the years 2013, 2020 and 2030. Assuming that such a PR is maintained, it is concluded that these targets may be attained at sustained annual growth rates of 21–42%, which seems feasible. Copyright © 2007 John Wiley & Sons, Ltd.

* Correspondence to: W. G. J. H. M. van Sark, Department of Science, Technology and Society, Copernicus Institute for Sustainable Development and Innovation, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands.
†E-mail: w.g.j.h.m.vansark@chem.uu.nl

KEY WORDS: experience curve; learning curve; progress ratio; crystalline silicon PV modules; error propagation; fitting; PV technology development

Received 1 March 2007; Revised 2 October 2007

INTRODUCTION

Learning or experience curves are widely used in policy and scenario studies in many fields[1–6] to account for and forecast technology development. These curves illustrate that technical and economic performance of a technology increases substantially as producers and consumers gain experience with this technology. Typically, production costs are reduced sometimes by orders of magnitude. This phenomenon was first described by Wright:[7] he reported that unit labour costs in airframe manufacturing declined significantly with accumulated experience of the workers, and also that this cost reduction was a constant percentage with every doubling of cumulative output. Plotted on a double logarithmic scale this empirical relationship constitutes a straight line, nowadays referred to as 'learning curve', as only the effects of learning-by-doing are measured.[8] Of particular interest was to investigate 'the possible future of airplane cost'[7] by extrapolation of this line. Arrow introduced the notion to general economics, namely that this learning-induced cost reduction was the product of experience.[9] The Boston Consultancy Group further extended the learning curve concept in two ways.[1] Firstly, the concept was applied to the total cost of a product, thereby including other learning mechanisms (such as RD&D and economies of scale), and other cost factors (e.g. cost of capital, marketing, overhead). Secondly, the concept was applied not only within a single company, but also to entire industries. In order to distinguish the curves based on this broader approach from simple learning curves they were labelled 'experience curves'. Nevertheless, the term learning curve is sometimes also used as synonym for experience curve.

Long-range forecasts are used in planning possible future solutions to a range of socio-economic problems, among which the climate problem is the most pressing.[6,10–12] Developers of scenarios that include energy supply and demand and the use of renewable energy technologies, prefer to use endogenous learning in which only one parameter determines the rate of learning. Therefore experience curves are now used in most leading climate-economy models, as recently reviewed by Köhler *et al.* [12] Another use is reported by Sandén for the assessment of PV subsidies.[13] In many studies, experience curves have been constructed on the basis of historical data that sometimes span 2–3 decades. From these curves one deduces the so-called 'progress ratio' (PR), which is the relative amount of cost reduction per doubling of cumulative output. The 'learning rate' is then defined as one minus PR. Dutton and Thomas[14] have analysed over 100 experience curves for manufacturing firms and found PRs between 0·6 and 1·0, with a mean of 0·8. McDonald and Schrattenholzer[15] have collected data for energy technologies (26 data sets) and found a distribution of PRs also between 0·6 and 1·0, but with a slightly higher mean of 0·84.

Given the empirical nature of the experience curve data and related inherent uncertainties, the PR may vary somewhat when key parameters are changed such as the assumptions about initial capacity installed, the associated start-off costs, the method of aggregating annual data, correcting for inflation and varying exchange rates and changing the learning system boundaries.[16,17] Small changes in the PR value can lead to strongly varying results in (long-term) scenarios and energy models, which rely on experience curves to model endogenous learning.[18,19] In many climate-economy models the lack of uncertainty analysis is considered to be an important deficiency and incorporation of uncertainties into such models is viewed as a major challenge.[12]

As an example, why consideration of uncertainty is important we can consider PV implementation scenarios. In such scenarios, a relatively minor variation of the PR for PV has an enormous influence on the total 'learning investment', which is defined as the cumulative excess cost for PV generation above the break-even level where PV installations become competitive with conventional electricity-generating plants.[20,21] Van der Zwaan and Rabl report that if the break-even system cost is taken as 1 $/Wp, the learning investment in PV systems is calculated to be 211 billion US$ for a PR of 0·80, which is reduced to one-third for a PR of 0·75.[20] This example clearly shows the sensitivity of the PR, and indicates the need for adding an uncertainty indication to the PR value, for example $0{\cdot}80 \pm 0{\cdot}05$. Scenarios using PRs should always include sensitivity studies to show the effect of different PRs, and the given error should indicate the range of possible values. With the analyses presented in this paper we aim to contribute to a better understanding of the uncertainties that are introduced by the use of experience curve concepts into all kinds of scenario models.

PRs for photovoltaic (PV) technology have been used to assess the prospects and diffusion of PV.[4,20–23] Harmon[24] and Parente *et al.*[25] recently updated PV experience curves on the basis of crystalline silicon module price data from Maycock up to the year 2000.[26] Harmon reported a PR of 0·798 with $R^2 = 0{\cdot}9927$,[24] while Parente[25] showed that a statistically significant break occurs in 1991: in the period 1981–1990, a PR was determined of 0·798 ($R^2 = 0{\cdot}977$), while in the period 1991–2000 a 3% lower PR was determined of 0·774 ($R^2 = 0{\cdot}978$). Fitting the complete curve (period 1981–2000) yielded PR = 0·772 with $R^2 = 0{\cdot}988$. From the given values one is tempted to associate an error of one digit, that is, 0·001. However, this is too accurate, as we will show below. More recently, Nemet[27] has compared crystalline silicon PV module experience curves on the basis of two datasets and found PR values of 0·74 and 0·83, for datasets from Maycock[28] and Strategies Unlimited[29], respectively. Although Nemet[27] does not attempt to explain the PR differences from these two datasets, they appear to be primarily caused by different data for the beginning of the experience curve, that is, below 30 MW cumulative capacity. A moving average analysis by Schaeffer *et al.*[21] with a time window of 10 years has shown that the PR may vary between 0·84 and 0·7 for 10-year periods starting from 1976 to 1992, with low PR values in the most recent time-windows. Associated R^2-values are between 0·84 and 0·98.

More recently, Swanson[30] has added crystalline silicon module cost data up to 2005, and reported a PR of 0·81. The recently occurring silicon feedstock supply shortage problem, which results from the sustained high growth of the PV industry[31–33] has led to module price increases in the past years. This is expected to lead to an increase of the PR, as was indeed clearly perceptible in Swanson's data.[30] This effect will become even more apparent from moving time analyses, as we will demonstrate below. It is generally expected that this price increase is only temporarily and the PV industry will continue riding down the experience curve within a few years. In addition, other types of modules based on thin films are expected to increase their present market share of about 10%.[33] These thin films modules have a lower cost per Wp figure as well as lower efficiency, and will definitely contribute to meeting mid- and long-term system cost goals. Although some attempts have been made to construct an experience curve separate from the crystalline silicon experience curve,[34] in this paper we focus on crystalline silicon, as the data set is more complete and gives better insight in experience curve development. Therefore, the PV experience curve presented in this paper implicitly refers to the crystalline silicon module experience curve.

We can conclude from this review that substantial differences are found in PR values, depending on the data source and the considered time period. In none of the reviewed analyses the uncertainty in the PR was determined, although the number of digits suggest errors as small as 0·01 or even 0·001.

The determination of PR involves fitting of historical data that span one of more decades, and resulting values for PR are given in two or three digits. In many cases, the correlation coefficient R^2 associated with reported values for PR is not given,[35] therefore data cannot be checked for reliability. In this paper, we show that one can easily deduce the error in PR from the definition of PR: it is expressed in the two constants that result from the fit. Subsequently, we will present an updated PV experience curve including data up to the year 2006, and we will determine the error in PR. We continue with a comparison of our data with reported data,[21,24,25] and deduce the error in PR from that data. Performing a moving average analysis with a time window of 10 years of the updated data will show the effect of the recent silicon feedstock shortage problem on the PR: it approaches 0·9 for the time windows that include the most recent years.

In most experience, curve studies data for 1 year are plotted as one data point, while apparently this data point is the result of averaging data collected from several sources for that particular year. Upon averaging data the standard deviation in the mean can be calculated, which can be shown in the graph by introducing error bars. Then, knowing these error margins, a weighted fit can be performed, which may yield a different value of PR in comparison to an unweighted fit. We will perform weighted fits using various errors to study the effect on the error in PR. As a further example, we will compare weighted and unweighted fitting for data of PV *system* price development in the Netherlands.

Finally, using projected module cost data for the years 2013, 2020 and 2030 from the recently published

DOI: 10.1002/pip

Strategic Research Agenda from the European Photovoltaic Technology Platform,[36] we will determine annual PV module market growth rates that will be needed to reach these module cost targets, based on extrapolation of the experience curve.

THEORETICAL CONSIDERATIONS

Experience curve

An empirical relation has been reported to exist between the cost and the cumulative production in a wide range of products.[1–3,5,7–9,14,15] A power function is generally used to describe this relation, although other functions have been proposed as well.[37] Usually double logarithmic graphs clearly demonstrate a linear relationship, where the slope is a measure of learning or experience; hence the term learning or experience curves. Such a curve can be described as:

$$c_x = ax^m \quad \text{or in logarithmic form} \qquad \log c_x = \log a + m \log x \tag{1}$$

in which c_x is the costs required to produce the xth unit of production, x the cumulative production up to and including the xth unit of production, a the costs required to produce the first unit and m the measure of the rate of costs reduction as cumulative production increases. The constant parameter m also is denoted learning or experience parameter, and is used to calculate the PR for cumulative doubling of production:

$$\mathrm{PR} = \frac{c_{x_2}}{c_{x_1}} = \frac{ax_2^m}{ax_1^m} = 2^m \quad (\text{for } x_2 = 2x_1) \tag{2}$$

The learning rate LR is then defined as

$$\mathrm{LR} = 1 - \mathrm{PR} \tag{3}$$

Both PR and learning rate are expressed in ratios or percentages. Values for PRs typically range from 1·0 (100%) to 0·6 (60%), with a mean around 0·8 (80%).[14,15] Note that in practice cost data are not readily available, and price is used as a *proxy* for cost.

The error σ_{PR} in the PR can be calculated from error propagation theory as given by, for example, Bevington:[38]

$$\sigma_{\mathrm{PR}} = \left(\frac{\mathrm{d}(2^m)}{\mathrm{d}m}\right)_m \sigma_m = \ln 2 \cdot 2^m \cdot \sigma_m = \ln 2 \cdot \mathrm{PR} \cdot \sigma_m \tag{4}$$

in which σ_m is the error in parameter m, as resulting from the fitting procedure.

Fitting

The experience curve as shown in Equation (1) can be generalized as:

$$y = f(x; a, m) \tag{5}$$

in which y is the dependent variable, x the independent variable, $f()$ the function describing the dependency between y and x, and a and m the parameters used in the function $f()$. The chi-square minimization is widely used as a standard way of defining the best fit: it minimizes the sum of the squares of the vertical deviations between the experimental curve and a (non)-linear theoretical curve of choice. The reduced chi-square χ_r^2 is defined as:[38]

$$\chi_r^2(a, m) = \frac{1}{n-p}\sum_i w_i[y_i - f(x_i; a, m)]^2 \tag{6}$$

in which $n-p$ is the degrees of freedom, n is the number of data points (x_i, y_i), p the number of parameters (in this case 2, i.e. a and m) and w_i the weight associated with the ith data point. Minimization of χ_r^2 is often done by using the Levenberg-Marquardt method,[39,40] which is implemented in many (non)-commercial data analysis software tools, see for example Reference [41]. Standard errors in parameters can be calculated as well using co-variance matrices,[38] in which the goodness of the fit is reflected, that is, small errors correspond in general to a good fit. The error σ_m is used in determining the error in PR (σ_{PR}) as defined in the Equation (4).

Two types of weighting methods will be applied: (1) no weighting, with $w_i = 1$ and 2) instrumental weighting, with $w_i = 1/\sigma_i^2$, where σ_i are the errors in each data point when available, otherwise $\sigma_i = 1$.

Another way of determining a best fit involves the use of the coefficient of determination R^2 (also known as goodness-of-fit parameter), which is defined as the ratio of the regression sum of squares to the total sum of squares:[38]

$$R^2 = \frac{\sum_i \left[f(x_i; a, m) - \frac{1}{n}\sum_i y_i\right]^2}{\sum_i \left[y_i - \frac{1}{n}\sum_i y_i\right]^2} \tag{7}$$

The coefficient of determination R^2 varies between 0 and 1 and denotes the strength of association between y

DOI: 10.1002/pip

and $f(x;a,m)$. Fitted data with R^2 values larger than 0·8 are considered strongly correlated, whereas fitted data with $R^2 < 0{\cdot}25$ are weakly correlated.[38]

RESULTS AND DISCUSSION

Updated global PV experience curve

The recent silicon feedstock supply shortage has had, and still has, a large influence on the feedstock price, and consequently on the module price. The theory underlying experience curves does not account for an *increase* in raw material costs. Instead, it is implicitly assumed that raw material prices (e.g. silicon, steel, copper, plastics etc.) remain constant, and that production cost can be lowered incrementally by either becoming more and more efficient in the use of the material (thinner wafers), or substituting one material by another. Nemet has reported that indeed plant size and module efficiency together account for 73% of the module price reductions in the period 1980–2001; only 12% reduction is due to silicon feedstock price reductions.[27] A module cost analysis in the year 2000 showed that the feedstock constitutes about 30% to the module cost.[42] Considering the present high feedstock price, an updated bottom-up study on cost distribution is warranted. It is desirable to update the PV experience on a regular basis to assess the progress in the PV industry, and to show how especially the crystalline silicon PV industry has coped with the feedstock problem. High material costs have already prompted a faster development in reducing wafer thickness leading to silicon usage of 10 g/Wp, compared to about 13 g/Wp a few years ago.[36] When feedstock supply capacity is extended to cater for the growing needs of the crystalline silicon PV industry, the module price is expected to be decreasing faster than projected due to the wafer thickness developments.

For the construction of an updated crystalline silicon module experience curve including data up to 2006, we used as a starting point the data that were used in the PHOTEX project,[16] which were based on the Strategies Unlimited dataset (1976–2001),[29] which consists of average selling prices of PV power modules as a function of cumulative shipments. We then added data reported by Swanson for the years 2002–2005.[30] For the year 2006 we have used the market volume data reported by Photon International.[32] Following the price trends of the average module price as revealed on the Solarbuzz website,[43] we assumed that the module price of 2006 is equal to that of 2005, as reported by Swanson.[30] Thus, we arrive at the updated crystalline silicon PV module experience curve, as shown in Figure 1. Note that the data have been converted to 2006 US$, using appropriate deflators.[44] Fitting the complete dataset yields PR $= 0{\cdot}794 \pm 0{\cdot}003$, where we have determined the error using Equation (4).

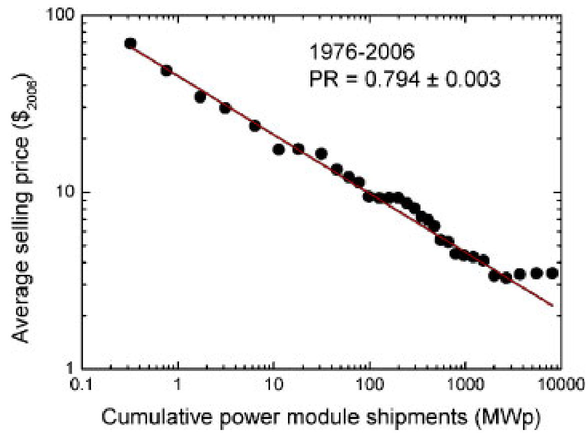

Figure 1. Updated crystalline silicon PV module experience curve showing average module price in 2006 US$ as a function of cumulative power module shipments. Data from Strategies Unlimited[29] are combined with data from Swanson[30] and Hirschman *et al*.[32]

Dependence of progress ratio error on R^2

In this section, the effect of fit quality as expressed by the correlation coefficient R^2 as a function of data spread around the best-fitted line is studied. The data used are the average selling prices of PV power modules in 2006 US$ as a function of cumulative shipments as shown in Figure 1. Fit results for the original data points are used as the starting data and are depicted as the solid line in Figure 2.

Data spread is introduced by changing the original data y_i according to:

$$y_i^r = y_i\left(1 + \left(\text{rand}() - \frac{1}{2}\right)\frac{r}{100}\right) \tag{8}$$

where rand() is a function that generates a random number between 0 and 1, r is the randomness parameter. In Figure 2 seven datasets are shown that were generated with values of the randomness parameter r ranging from 1 to 200. Generally, the larger r, the wider the data spread around the original data. Results of fitting these seven curves are given in Table I, where χ_r^2, R^2, PR, σ_{PR} and $|\Delta_{PR}|$ are listed.

DOI: 10.1002/pip

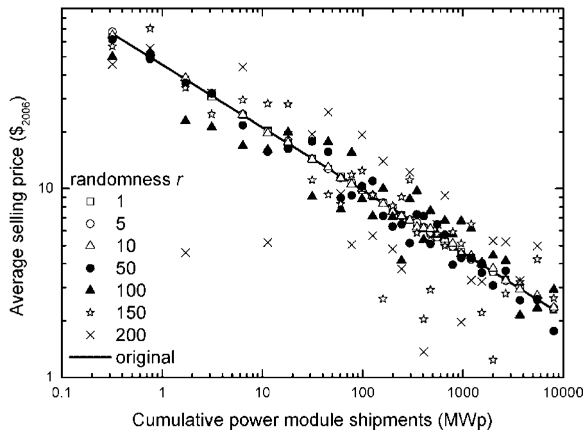

Figure 2. Effect of randomness parameter r on the average selling price in 2006 US\$ as a function of cumulative power module shipments (MWp). Experience curve data refer to the data in Figure 1

Here, $|\Delta_{PR}|$ is the difference of the PR determined from data with various randomness parameters with the PR determined from the original data. For most of the data in Table I $|\Delta_{PR}|$ is larger than the error σ_{PR} determined from the fits. Note that σ_{PR} increases for increasing randomness of the data, while R^2 decreases, as could be expected.

The dataset presented in Figure 2 was extended by generating more sets with randomness parameters in the range of 1–200. The PR and associated error as a function of correlation coefficient R^2 for fits of these datasets are shown in Figure 3. It can be inferred that a larger randomness parameter generally leads to worse fits. In addition, Figure 3 illustrates that associated errors in PR can be larger than the difference between original and generated data.

The errors in PR determined from the fit reflect the quality of the fit. This can be clearly seen in Figure 4 in which the dependence of PR error σ_{PR} on correlation coefficient R^2 is shown. PRs with associated $R^2 < 0{\cdot}9$ have appreciable errors larger than 0·01. Such apparently small errors may lead to significant differences in scenario outcomes. From our data we infer that fits with $R^2 \sim 0{\cdot}95$ lead to an error of close to 0·01 in the value of PR. Fits with $R^2 \sim 0{\cdot}98$ lead to an error of 0·005 in PR. Hence, values of PR reported in three digits such as by Parente *et al.*[25] of PR = 0·772 with $R^2 = 0{\cdot}988$ for the period 1981–2000 are less accurate than suggested. Based on their R^2 value an error of at least 0·003 seems reasonable.

Analysis of global PV experience curve

The global crystalline silicon PV module experience curve is further analysed by a moving average analysis. This is performed by moving a time window of 10-year duration over the data and by determining PRs from the best fits. The result is shown in Figure 5. Three periods can be discerned of apparent constant PR. In the first period, with starting years from 1976 to 1983 the average PR is $0{\cdot}810 \pm 0{\cdot}014$. In the second period, (starting years 1984–1987) we find an average PR of $0{\cdot}834 \pm 0{\cdot}005$. Although the difference in PR values is significant, it is small. When averaging over the two periods we arrive at a PR of $0{\cdot}818 \pm 0{\cdot}017$. For the period with starting years from 1989 to 1992, the average PR is considerably lower at $0{\cdot}717 \pm 0{\cdot}004$. Although the PRs of the first two periods do not differ much, albeit significant, the PR in the period 1989–1992 clearly is much lower. This break already appears in the starting year 1988, to be full from 1989 onwards. From the starting year 1992 onwards, the PR is steadily increasing to reach $0{\cdot}75 \pm 0{\cdot}02$. For later starting years (1995–1997) the increase is more rapidly, due to the price increase in the years 2004–2006 resulting from the silicon feedstock problem. Note, that also R^2 decreases for the later starting years. For a moving average analysis a duration of 10 years may be too short, as market

Table I. Results of fitting the data of Figure 2 as a function of randomness parameter r. Experience curve data refer to the data in Figure 1

Parameter	Starting data	Randomness parameter r							Experience curve data
		1	5	10	50	100	150	200	
χ^2_r	0	0·00094	0·0686	0·30429	2·592	15·10668	31·14189	77·89322	1·66594
R^2	1	1·00000	0·99970	0·99839	0·98738	0·89775	0·88557	0·63282	0·99246
PR	0·79440	0·79467	0·7919	0·7948	0·797	0·814	0·800	0·83	0·794
σ_{PR}	0	0·00008	0·0007	0·0014	0·004	0·011	0·014	0·02	0·003
$\|\Delta_{PR}\|$	0	0·00030	−0·0025	0·0004	0·003	0·020	0·006	0·03	—

DOI: 10.1002/pip

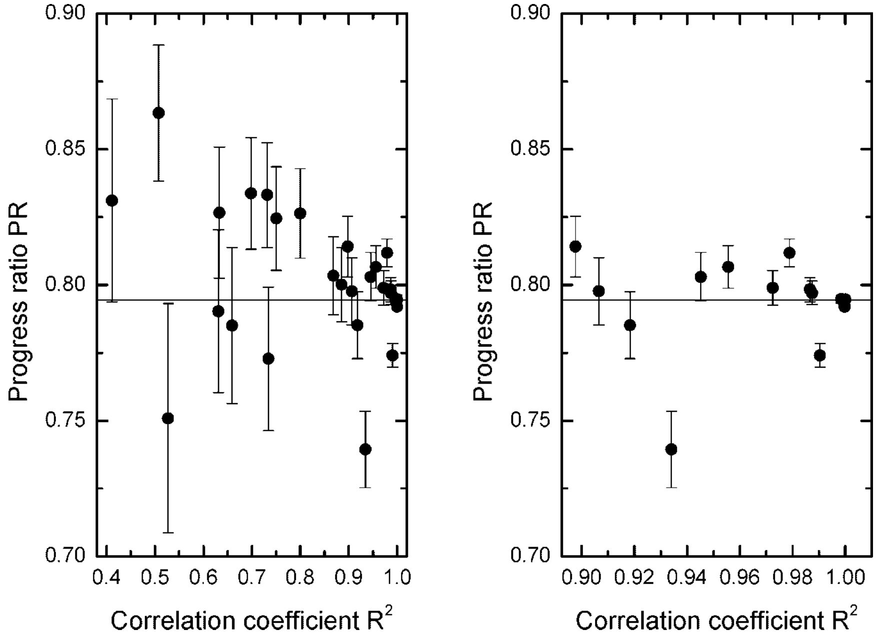

Figure 3. Calculated progress ratio and associated error as a function of correlation coefficient R^2 for fits of data generated with various randomness parameters in the range of 1–200. The right-hand side figure is an expansion of the left-hand side figure for $0{\cdot}9 < R^2 < 1$

effects such as a varying price-cost difference may be too influential. We have also used a 15 and 20 year time period and found smaller variations in both PR and R^2 compared to the 10-year period.

The analysis presented by Parente *et al.*[25] also showed, using the Chow structural break test, that a statistically significant break occurs in the learning curve, however, they reported the break to occur in 1991. Our analysis (Figure 5) showed that the break already started in 1988. Further, Parente *et al.*[25] reported for the period 1981–1990 a PR of 0·798 ($R^2 = 0{\cdot}977$), while in the period 1991–2000 a 3% lower PR is determined of 0·774 ($R^2 = 0{\cdot}978$). As they reported errors in the fitting parameter m, we can use these to determine the error in PR with Equation (4). We thus arrive at the values shown in Table II, in which

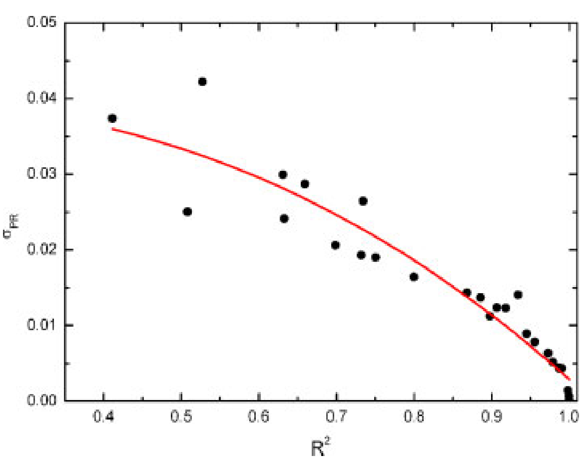

Figure 4. Dependence of progress ratio error σ_{PR} on correlation coefficient R^2.

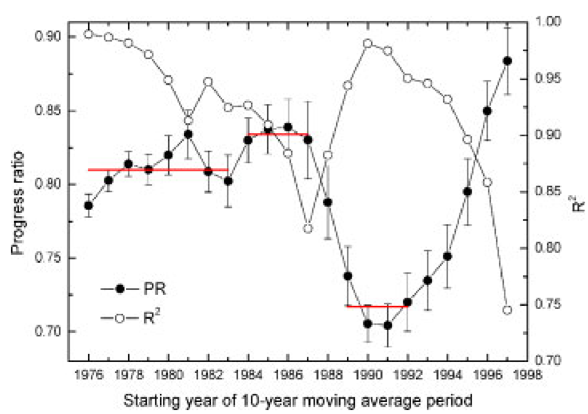

Figure 5. Consecutive set of PRs and associated R^2 values determined from fits using 10-year time windows over the experience curve data presented in Figure 1.

DOI: 10.1002/pip

Table II. Comparison of progress ratios determined from fitting various time periods. The errors in the results from Parente *et al.*[25] are calculated from their data. The present study uses original data for the periods indicated as reported by Strategies Unlimited[16,29]

	1981–1990	1991–2000	1981–2000
Parente *et al.*			
PR	$0{\cdot}798 \pm 0{\cdot}010$	$0{\cdot}774 \pm 0{\cdot}011$	$0{\cdot}772 \pm 0{\cdot}010$
R^2	0·977	0·978	0·988
Present study			
PR	$0{\cdot}834 \pm 0{\cdot}016$	$0{\cdot}704 \pm 0{\cdot}015$	$0{\cdot}816 \pm 0{\cdot}009$
R^2	0·913	0·975	0·954

results of our own analysis of the original[29] data of Figure 1 are given for the indicated periods. Clearly, there are considerable differences between the analysis results. One possible explanation of this is the sources of data, which are not the same; however, this cannot be ascertained.

An open issue is whether or not PRs remain constant over time. Some authors assume that experience curves will flatten out with increasing market penetration.[45] This has also been predicted for crystalline silicon modules,[34] however, present data for PV technology contradict this, when regarding data up to the starting year 1992. This improvement of PR in the 1990s coincides with a period of relatively more R&D support in the early 1990s and a relatively lower market growth rate, see for more details References 16,21,23. The increase in PR in the most recent years is entirely due to price increase due to the silicon feedstock shortage, and not to a large market penetration, as electricity generated by PV constitutes a share which is far below 1%.

Effect of errors in data on weighted fits

We introduce various errors in the original data to assess the importance of weighting the fits and the influence of errors on the fitting results. The used errors for each data point are shown in Figure 6. Constant relative errors are used for all data of 1, 5, 10, 15, 20 and 25%. In addition, linearly varying errors are used such that older data are considered more unreliable. Also, a combination of constant relative and linearly decreasing errors is used, denoted 'slow linear', to add another error series. A Poisson-type error is calculated as $\sigma_{y_i} = 1/\sqrt{y_i}$. A random error is generated between 0 and 50%. Finally, a constant error of 1 $\$_{2006}$/Wp is used.

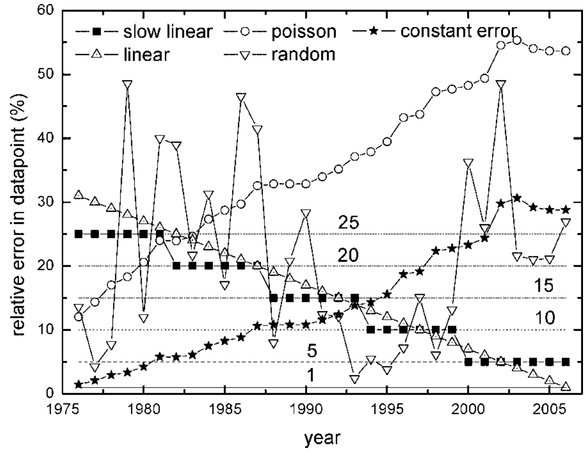

Figure 6. Relative errors used in comparing weighted and unweighted fitting of the experience curve data of Figure 1

The results of fitting the data where the errors are used as weights according to $w_i = 1/\sigma_i^2$ are shown in Table III. Using a constant error of 1 $\$_{2006}$/Wp is equivalent to using weights equal to unity, and therefore should yield identical results as obtained with non-weighted fitting. The results confirm this. An increasing value of the relative error, which is the same for every data point, yields increasing errors in PR; the value of PR and R^2 do not change. This is in correspondence with the definition of χ_r^2.

For the constant relative error cases, the error in the PR σ_{PR} is smaller or similar than the difference $|\Delta_{PR}|$ of the PR determined from the weighted fits with the PR determined from the original data, except for the result with a constant relative error of 25%.

Actual errors in the data points are not known. Assuming larger errors in earlier years and smaller errors in more recent years may be reasonable, if we assume that data quality is better for more recent years. The fitting results for the linearly decreasing errors show a quite low R^2 of 0·425 and a large $|\Delta_{PR}|$ of 0·0103. This result is entirely due to the inclusion of the data of the years 2003–2006, with relatively small error. If we leave out the latter period and only fit the data between 1976 and 2002 we obtain $R^2 = 0{\cdot}95978$ and $PR = 0{\cdot}789 \pm 0{\cdot}007$ with $|\Delta_{PR}| = -0{\cdot}005$. The Poisson-type and constant error lead to larger errors in later years. This may be warranted by the fact that the present variation in system size and price may be larger than in earlier years, which may be substantiated by the fact that fitting the data leads to a smaller error in PR than in the case of smaller errors in the data in later years.

DOI: 10.1002/pip

Table III. Results of weighted fitting of the experience curve data of Figure 1 with the errors used as shown in Figure 6

	Original data	Constant error (1 $/Wp)	Constant relative error (%) 1	5	10	15	20	25	Linear decreasing error	Linear decreasing error, slow	Poisson	Random
χ_r^2	1·66594	1·66594	114·262	4·57048	1·14262	0·50783	0·28566	0·18282	5·51392	2·33622	1·94139	0·84941
R^2	0·99246	0·99246	0·96278	0·96278	0·96278	0·96278	0·96278	0·96278	0·42531	0·79193	0·99652	0·97038
PR	0·794	0·794	0·8029	0·803	0·803	0·803	0·803	0·803	0·898	0·817	0·764	0·803
σ_{PR}	0·003	0·003	0·0004	0·002	0·004	0·006	0·008	0·010	0·007	0·007	0·002	0·003
$\|\Delta_{PR}\|$	—	0	0·0085	0·0085	0·0085	0·0085	0·0085	0·0085	0·103	0·022	−0·030	0·0085

From the weighted fits presented here we tentatively infer an absolute error of between 0·005 and 0·010 to be a reasonable estimate for the error in PR, which is around 1%. Even if the errors in the data points are of the order of 10–20%, the resulting error in PR is an order of magnitude lower. We therefore can conclude that a relatively large variance in the individual data points (i.e. yearly averages) can be accepted without compromising the accuracy of the overall PR value. Further, from the table we can conclude that using weighted fits when errors are available leads to increased values of σ_{PR} with respect to unweighted fitting for errors in data points larger than about 10%.

PV system price development in the Netherlands

As an illustration of the weighted fitting approach described above we present an experience curve analysis of systems installed in the Netherlands in the period 1991–2002. Data are taken from a database, which was constructed in the framework of the European PHOTEX project.[16,23] The database contains price data for modules, system components and complete grid-connected systems. Some 3500 records are available, being a representative sample for PV systems installed in France, Germany, Italy and the Netherlands. The original price data, in various local currencies, were converted to Euro of the year 2000, Euro$_{2000}$. Most of the data are end-user prices of typical rooftop systems. Additional data was compiled from Dutch sources.[46] It should be noted that determining a PR from *system* prices may not be correct as learning can be different for the various parts that constitute the system. Module price development takes place on a global level, whereas inverter and mounting development is much more local.[16] Nevertheless, the analysis here serves purely as an example of applying weighted fitting.

Figure 7 shows the development of Dutch PV system price as a function of cumulative installed capacity, price information of 173 individual systems is depicted. For every year of installation the data were averaged to one data point. The standard deviation of the mean is used as error bars in these data points, see Figure 7. Three ways of fitting are performed: (1) using individual data points; (2) unweighted fitting of averaged data points; (3) weighted fitting of averaged data points where the errors are used for weighting according to $w_i = 1/\sigma_i^2$. Results are shown in Table IV. The PR determined from fitting all individual data points is somewhat larger than the one determined from averaged data, as a result of the fact that the majority of data points is of recent date, which emphasizes the recent data. Weighted fitting yields a PR of $0{\cdot}876 \pm 0{\cdot}010$. Weighting in this case lowers the associated error of PR from 0·016 to 0·010. Note that the global module PR is considerably lower, suggesting that in this case study either module prices

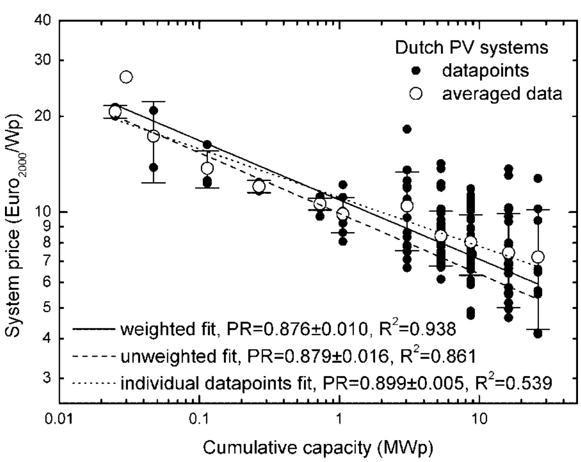

Figure 7. System price in Euro$_{2000}$/Wp of PV systems installed in the Netherlands between 1992 and 2002 as a function of cumulative installed capacity (MWp)

DOI: 10.1002/pip

Table IV. Results of fitting the Dutch system price data as depicted in Figure 7

Parameter	Individual data points	Averaged data with error bars	
		Unweighted	Weighted
χ_r^2	4·58689	5·48399	0·7734
R^2	0·53853	0·86144	0·93798
PR	0·899	0·879	0·876
σ_{PR}	0·005	0·016	0·010
$\lvert\Delta_{PR}\rvert$	0·023	0·002	0

did not follow world market prices or that balance-of-system (BOS) components such as inverter and mounting structure learned at a slower rate than modules.

PV technology development outlook

Using the experience curve as shown in Figure 1, one can extrapolate the price development beyond the 1000 GW range, assuming that the PR will remain constant, within say 5%. Figure 8 shows an extrapolation up to 10 000 GWp cumulative installed capacity, which would constitute about 1% of the global energy demand in 2050.[47] It is then interesting to compare the extrapolated data with presently used PV technology R&D roadmaps. As an example, the Strategic Research Agenda as formulated by the European PV Technology Platform[36] specifies cost targets for systems, modules and BOS to be realistically reachable in the short (2013), medium (2020) and long (2030) term. The cost targets for modules are used for *all* flat-plate PV module technologies considered. They are 0·8–1·0 €/Wp, 0·60–0·75 €/Wp and 0·3–0·4 €/Wp, for 2013, 2020 and 2030, respectively, and are shown in Figure 8 as horizontal lines. For the conversion from € to US$ we used the 2006 annual average interbank exchange rate of 1 US$ = 0·797043 €.[48] The cost ranges reflect the ranges in efficiency. Modules with lower efficiency need to be cheaper than modules with higher efficiency to yield comparable overall system cost.

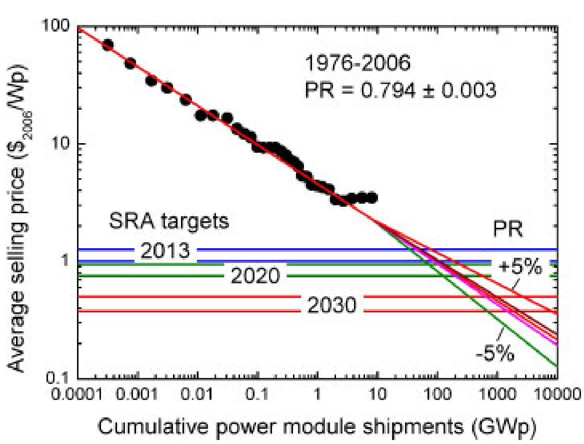

Figure 8. Extrapolated experience curve up to a cumulative installed capacity of 10 000 GWp, extending Figure 1. Also indicated are the cost targets for the years 2013, 2020 and 2030 as specified in the Strategic Research Agenda (SRA) from the European Photovoltaic Technology Platform,[36] and the effects of a 1% and 5% up- and downward variation of the progress ratio.

The amount of cumulative installed capacity required to attain the SRA cost targets can be calculated using Equations (1) and (2) to find:

$$x_t = x_{2006}\left(\frac{c_t}{c_{2006}}\right)^{\frac{1}{m}} \tag{9}$$

where x_t and x_{2006} are the cumulative capacities corresponding to the target year t and the year 2006, respectively, and c_t and c_{2006} are the cost targets in $/Wp corresponding to the target year and the year 2006, respectively. Note that two cost targets are set per target year. Assuming a constant annual market growth rate r_a one can find this rate knowing that the cost targets are set at a certain target year t, as follows:

$$r_a = \left(\frac{x_t}{x_{2006}}\right)^{\frac{1}{(t-2006)}} - 1 \tag{10}$$

We thus derive a range in installed capacity of 49–96, 117–228 and 774–1841 GWp, for 2013, 2020 and 2030, respectively. To reach these large amounts of a sustained *annual* growth rate of the PV module capacity would be needed of 29–42, 21–27 and 21–25%, for 2013, 2020 and 2030, respectively. These numbers are somewhat lower than the 30–40% annual increase of capacity as realized in the past 5–10 years,[31,32] and may therefore be considered as feasible. Note that the experience curve used is based on price data, as cost data are not available. Assuming costs to be lower than prices, the numbers presented here would have to be adjusted accordingly, that is, lowered somewhat, which brings the targets even closer in reach. A variation in PR of 1 or 5% up- or downwards yields values shown in Table V, again illustrating the large effect of varying PR. Clearly, lower values of PR yield lower cumulative capacities for a specific target,

Table V. The effect of a small (1 and 5%) up- and downward variation in progress ratio on the cumulative installed capacity and sustained annual growth rate for the target years 2013, 2020 and 2030 as specified in the Strategic Research Agenda from the European Photovoltaic Technology Platform.[36] The 2006 data used are 8·07 GWp installed capacity at 2·28 $/Wp

	Cost target (€/Wp)	Cumulative capacity (GWp)					Annual growth rate (%)				
PR variation		−5%	−1%	—	+1%	+5%	−5%	−1%	—	+1%	+5%
PR value		0·771	0·786	0·794	0·802	0·818	0·771	0·786	0·794	0·802	0·818
2013	1·0	13·87	31·20	49·02	79·85	240·7	8·05	21·3	29·4	38·7	62·4
	0·8	25·11	59·39	95·98	161·2	520·3	17·6	33·0	42·4	53·4	81·3
2020	0·75	29·82	71·55	116·6	197·4	650·3	9·78	16·9	21·0	25·7	36·8
	0·6	53·98	136·2	228·3	398·6	1406	14·5	22·4	26·9	32·1	44·6
2030	0·4	158·7	438·9	774·0	1428	5707	13·2	18·1	20·9	24·1	31·4
	0·3	341·0	1007	1841	3532	15421	16·9	22·2	25·4	28·8	37·0

with concomitant lower annual growth rates. The 0·8 €/Wp target in 2013 for 1 or 5% larger values of PR yields unrealistically large annual growth rates. Overall, the 2013 targets constitute a large challenge for the industry to reach.

CONCLUSION

The consequences of a small variation of the assumed PR for PV technology can be enormous in forecasting energy scenarios, such as the prediction of the societal cost of reaching break-even. Inclusion of an error in the PR provides scenario developers with the smallest necessary range over which sensitivity studies should be done.

The determination of PR from experience curves should go along with determination of its error. This can easily be done in standard spreadsheet software, and we have presented the equation that can be used for the calculation of the error σ_{PR}.

Fitting of data always yields a coefficient of determination R^2, which should approach unity for a good fit. We have studied the relation of the error in the PR with R^2. This error was found to be as large as 0·005 for R^2 values of 0·98, while such values are considered representative of an excellent fit.

In the case of crystalline silicon PV module technology development we have fitted the available price data of the period 1976–2006 and have determined $PR = 0{\cdot}794 \pm 0{\cdot}003$. A moving average approach with a 10-year time window showed that the PR is not constant. It varies from $0{\cdot}818 \pm 0{\cdot}017$ up to a starting year of 1987. For the period up to a starting year of 1992, the average PR is considerably lower at $0{\cdot}717 \pm 0{\cdot}004$. This has been explained by an increased R&D support in the early 1990s and a relatively lower market growth rate leading to weak prices.[16,21,23] For the most recent starting year 1997, the PR is considerably higher at $0{\cdot}884 \pm 0{\cdot}022$, which can be explained by the silicon feedstock supply problem and capacity bottlenecks that causes module prices to rise.

When available, error in individual data points can be used to perform weighted fits to lower the error in PR. An analysis of Dutch PV system price development over the period 1992–2002 shows that PR is $0{\cdot}876 \pm 0{\cdot}010$, where the error is improved with respect to unweighted fitting.

The found PR of $PR = 0{\cdot}794$ has been used to analyse the cost targets stated in the Strategic Research Agenda as formulated by the European PV Technology Platform.[36] Meeting these targets for the years 2013, 2020 and 2030 requires sustained annual growth rates of 21–42%, which is realistic. However, an increase of only 5% of the value of the PR reaching 0·818 implies unrealistically high growth rates. Further, none of our analyses included non-crystalline silicon PV technologies, which also have large potential to meet the stated European targets.

Acknowledgements

We gratefully acknowledge Dick Swanson for supplying module price data,[30] R. Gerwig for his thorough analysis of Dutch PV system price development and an anonymous reviewer for many useful comments.

REFERENCES

1. Boston Consulting Group. Perspectives on Experience. 1972; BCG.
2. Cunningham JA. Using the learning curve as a management tool. *IEEE Spectrum* 1980; **17**: 45–48.

3. Liebermann MB. The learning curve and pricing in the chemical processing industries. *The RAND Journal of Economics* 1984; **15**(2): 213–228.
4. Neij L. Use of experience curves to analyse the prospects for diffusion and adoption of renewable energy technology. *Energy Policy* 1997; **25**(13): 1099–1107.
5. Wene C-O. *Experience Curves for Energy Technology Policy*. OECD/IEA: Paris, 2000.
6. Van der Zwaan B, Gerlagh R. Climate sensitivity uncertainty and the necessity to transform global energy supply. *Energy* 2006; **31**: 2235–2251.
7. Wright TP. Factors affecting the cost of airplanes. *Journal of the Aeronautical Sciences* 1936; **3**(4): 122–128.
8. Neij L. *Dynamics of energy systems*. Ph.D. Thesis, Lund University, Lund, Sweden, 1999.
9. Arrow KJ. The economic implications of learning by doing. *Review of Economic Studies* 1962; **29**: 155–173.
10. Grubler A, Nakicenovic N, Victor DG. Dynamics of energy technologies and global change. *Energy Policy* 1999; **27**(5): 247–280.
11. Van Vuuren DP, De Vries B, Eickhout B, Kram T. Responses to technology and taxes in a simulated world. *Energy Economics* 2004; **26**: 579–601.
12. Köhler J, Grubb M, Popp D, Edenhofer O. The transition to endogenous technical change in climate-economy models: a technical overview to the innovation modeling comparison project. *The Energy Journal* 2006; **18**: 17–55.
13. Sandén BA. The economic and institutional rationale of PV subsidies. *Solar Energy* 2005; **78**: 137–146.
14. Dutton JM, Thomas A. Treating progress functions as a managerial opportunity. *Academy of Management Review* 1984; **9**(2): 235–247.
15. McDonald A, Schrattenholzer L. Learning rates for energy technologies. *Energy Policy* 2001; **29**(4): 255–261.
16. Schaeffer GJ, Alsema EA, Seebregts A, Beurskens L, De Moor H, Van Sark WGJHM, Durstewitz M, Perrin M, Boulanger P, Laukamp H, Zuccaro C. Learning from the sun: analysis of the use of experience curves for energy policy purposes: the case of photovoltaic power. Report No. ECN-C-04-035, 2004; Energy research Centre of the Netherlands (ECN).
17. Junginger HM. *Learning in renewable energy technology development*. Ph.D. Thesis, Utrecht University, Utrecht, the Netherlands, 2005.
18. Neij L. Cost dynamics of wind power. *Energy* 1999; **24**(5): 375–389.
19. Van der Zwaan BCC, Seebregts A. Endogenous learning in climate-energy-economic models—an inventory of key uncertainties. *International Journal of Energy Technology and Policy* 2004; **2**(1/2): 130–141.
20. Van der Zwaan B, Rabl A. Prospects for PV: a learning curve analysis. *Solar Energy* 2003; **74**(1): 19–31.
21. Schaeffer GJ, De Moor HHC. In *Proceedings Nineteenth European Photovoltaic Solar Energy Conference*, Paris, France, 2004; 3415–3418.
22. Poponi D. Analysis of diffusion paths for photovoltaic technology based on experience curves. *Solar Energy* 2003; **74**(4): 331–340.
23. De Moor HHC, Schaeffer GJ, Seebregts AJ, Beurskens LWM, Durstewitz M, Alsema EA, Van Sark WGJHM, Laukamp H, Boulanger P, Zuccaro C. In *Proceedings 3rd World Conference on Photovoltaic Energy Conversion, WCPEC-3*, Osaka, Japan, 2003; 2624–2627.
24. Harmon C. *Experience curves for photovoltaic technology*. Report No. IR-00-014, 2000; IIASA.
25. Parente V, Goldemberg J, Zilles R. Comments on experience curves for PV modules. *Progress in Photovoltaics: Research and Applications* 2002; **10**: 571–574.
26. Maycock P. *PV News* (several issues over the years 1995–2002).
27. Nemet GF. Beyond the learning curve: factors influencing cost reductions in photovoltaics. *Energy Policy* 2006; **34**: 3218–3232.
28. Maycock PD. *The world photovoltaic market*. 2002; PV Energy Systems.
29. Strategies unlimited. Photovoltaic Five-Year Market Forecast 2002-2007. Report No. PM-52, 2003.
30. Swanson RM. A vision for crystalline silicon photovoltaics. *Progress in Photovoltaics: Research and Applications* 2006; **14**: 443–453.
31. Hirschman WP, Schmela M. Silicon shortage—so what! Market survey on cell and module production 2005. *Photon International* March 2006; 100–125.
32. Hirschman WP, Hering G, Schmela M. Gigawatts—the measure of things to come. Market survey on global solar cell and module production in 2006. *Photon International* March 2007; 136–166.
33. Van Sark WGJHM, Brandsen GW, Fleuster M, Hekkert MP. Analysis of the silicon market: will thin films profit? *Energy Policy* 2007; **35**(6): 3121.
34. Green MA. In *Proceedings 3rd World Conference on Photovoltaic Energy Conversion, WCPEC-3*, Osaka, Japan, 2003; O.1–O.5.
35. Snik A. *Experience curves for photovoltaic energy technology*. M.Sc. Thesis, Utrecht University, Dept. Science, Technology and Society, Utrecht, the Netherlands, 2002.
36. EU PV technology platform. *A strategic research agenda for photovoltaic solar energy technology*. 2007.
37. Yelle LE. The learning curve: historical review and comprehensive survey. *Decision Science* 1979; **10**: 302–328.
38. Bevington PR. *Data reduction and error analysis for the physical sciences*. McGraw-Hill: New York, NY, USA, 1969.
39. Levenberg K. A method for the solution of certain nonlinear problems in least squares. *Quarterly of Applied Mathematics* 1944; **2**: 164–168.

DOI: 10.1002/pip

40. Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society of Industrial and Applied Mathematics* 1963; **11**: 431–441.
41. Ledvij M. Curve fitting made easy. *Industrial Physicist* 2003; **9**: 24–27.
42. Frantzis L, Jones ECL, Wood M, Wormser P. In *Proceedings 16th European Photovoltaic Solar Energy Conference*, Glasgow, UK, 2000; 2100–2103.
43. Solarbuzz. Solar module price highlights website http://www.solarbuzz.com/ModulePrices.htm [Accessed in June 2007].
44. Johnson LD, Williamson SH. The annual real and nominal GDP for the United States, 1790—Present, Economic History Services http://www.eh.net/hmit/gdp/ [Accessed in June 2007].
45. McDonald A, Schrattenholzer L. Learning curves and technology assessment. *International Journal of Technology Management* 2002; **23**(7/8): 718–745.
46. Gerwig R. *PV experience curves for the Netherlands*. M.Sc. Thesis, Utrecht University, Dept. Science, Technology and Society, Utrecht, the Netherlands, 2005.
47. Graßl H, Kokott J, Kulessa M, Luther J, Nuscheler F, Sauerborn R, Schellnhuber H-J, Schubert R, Schulze E-D. *World in Transition, Towards Sustainable Energy Systems*. German Advisory Council on Global Change (WBGU) Earthscan: London, UK, 2004.
48. Oanda Corporation. Oanda.com, the currency database http://www.oanda.com [Accessed in June 2007].

DOI: 10.1002/pip