Supporting Information (SI)

Postcombustion Capture or Direct Air Capture in Decarbonizing US Natural Gas Power?

Habib Azarabadi, Klaus S. Lackner*

Corresponding author

Klaus Lackner- *School of Sustainable Engineering & the Built Environment, Arizona State University, Tempe, Arizona 85287, United States; Email: klaus.lackner@asu.edu*

Supplementary Information Contents:

The number of pages 19

The number of tables 6

The number of figures 7

The number of equations 5

S1. Power Plants Information Datamining

Python's Pandas package was used for datamining and analysis in this study. The main sources of data for NGCC units are the unit and generator information sheets ("UNT16" and "GEN16") from the eGRID 2016 metric data spreadsheet.

First, on the generator page, NGCC units with a MWh generation record for the year 2016 were identified. NGCC units are flagged with generator prime mover types of "CA" for combined cycle steam turbine, "CT" for combined cycle combustion turbine (or gas turbine) and "CS" for combined cycle single shaft. Generator information, including the plant's name and state, DOE/EIA facility code (ORISPL), generator ID, nameplate capacity in MW, 2016 electricity generation in MWh and age were collected. Generation and capacity data for CA and CTs which belong to the same NGCC unit were identified and aggregated and a combined dataset for generation and nameplate capacity of all NGCC units including 670 records was created.

Then the emission records for these NGCC units including the plant's name and state, ORISPL, unit ID, 2016 heat input in GJ and 2016 CO_2 emissions in ton were collected to merge the generator and emission data records. Similar to the generation records, emission records belonging to each NGCC unit were combined and the aggregate heat input and CO_2 emissions dataset of all NGCC units with 655 records was created.

The generation and emission datasets were processed to match as many records as possible and merged by using ORISPL code and other information such as the comparison between the heat input and generation values. A final database including generation and emission records for the NGCC fleet was created and the efficiency (η) and capacity factor (CF) values for each NGCC units were calculated:

$$CF = \frac{2016\ Generation\ (MWh)}{Nameplate\ Capacity\ (MW) \times 8766\ (h)} \tag{S1}$$

$$\eta = \frac{2016\ Generation\ (MWh)}{2016\ Heat\ Input\ (GJ) \times conversion\ factor\ (\frac{MWh}{GJ})} \tag{S2}$$

Missing heat input data and obvious inconsistencies resulted in 513 valid records out of 670. Invalid records have either very high or very low net efficiencies or in a few cases, capacity factor values above 1 or below 0. For this analysis and results in the paper, we only used the valid NGCC records and when results are shown in terms of total percentage, a conversion factor was used to convert the results from the valid units to all the units.

Table S1 summarizes some statistical information about the generated database. Median values are calculated by the number of units. Mean values for capacity factor and net efficiency are also calculated by the number of units, while the mean for Levelized Cost of Electricity (LCOE) values is an arithmetic mean weighted by the MWh electricity generation of each unit. Minimum and maximum net efficiency values are our assumptions; datapoints outside of the 35%-52% range are assumed to be invalid. LCOE values are only calculated for units considered for retrofit (larger than 25 MW and younger than 25 years).

Table S1 Statistical information for 513 valid records in the generated NGCC database

Parameter	Min	Median	Mean	Max
Nameplate Capacity (MW)	7.7	458	467	1850
Age (years)	0	14	15.3	51
Net Efficiency (HHV%)	35%	45.5%	44.5%	52%
Capacity Factor	0.002	0.47	0.46	0.96
Number of Gas Turbines in a Unit	1	2	1.9	12
Number of Steam Turbines in a Unit	1	1	1.1	4
LCOE pre-retrofit (2017$/MWh)	39.79	57.33	55.06	6,800
LCOE post-retrofit (2017$/MWh)	54.58	89.45	84.56	17,000

S2. Cumulative Generation Versus Capacity Factor

In this analysis, natural gas-fired generating units are divided into two groups: NGCC and non-NGCC. NGCC units are relatively larger and are mainly utilized for intermediate-load and baseload electricity demand. Non-NGCC units, on the other hand, are mainly small gas or steam turbines widely used for peak load electricity demand. There are about 6 times more non-NGCC units than NGCC units, while NGCC units produce 5.5 times more electricity.

Figure S1 illustrates the cumulative distribution of NGCC and non-NGCC units based on their capacity factor. As shown in Figure S1a, most NGCC units operate in the intermediate-load range. Only 20% of NGCC units are categorized as peaker (capacity factor < 0.2) and they produce an insignificant percentage of NGCC electricity. On the contrary, roughly 80% of non-NGCC units are peakers and they generate about 40% of total non-NGCC electricity (Figure S1b).

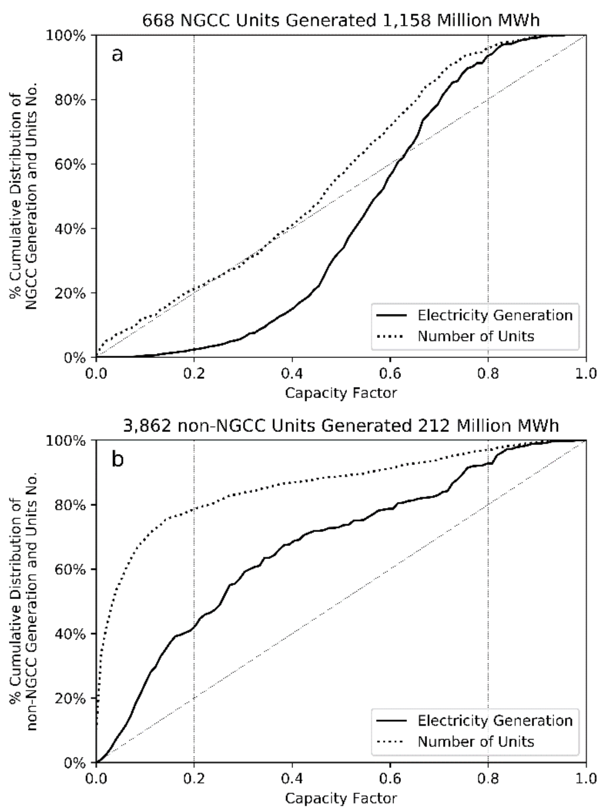

Figure S1 2016 cumulative distribution of (a) NGCC and (b) non-NGCC natural gas units versus capacity factor. Clearly, most of non-NGCCs provide electricity only during peak demand hours while most NGCC units are categorized as intermediate-load units.

S3. Use of Machine Learning for Validating All Power Plant Data Records

Missing heat input values result in invalid net efficiency values for some of the NGCC units. To address the issue and investigate the potential error from the invalid datapoints, we used a simple machine learning algorithm to estimate the net efficiency of these units. Then used the estimated efficiency values and units' 2016 generation record to calculate their heat input by equation S2. Then an emission factor (ton CO_2/GJ heat input) was used to calculate CO_2 emissions for these units.

Several machine learning algorithms were used, but we found the K-nearest neighbor regression method the most accurate.[1] K-nearest neighbor method uses net efficiency value(s) of the K nearby (similar) NGCC units to estimate the net efficiency of an invalid data record. K is an optimum value which results in the highest regression accuracy. We used nameplate capacity, capacity factor, and age as the determining attributes for net unit efficiency.

Typically, the valid datapoints are split into two categories, one for training and the other one for testing the trained algorithm. We used a range of K values from 0 to 100 and $K = 16$ was found to be the optimum value resulting in the most accurate regression for the testing set (16 nearest points used in the regression algorithm). Figure S2 shows r-squared scores of the training and testing sets for different K values.

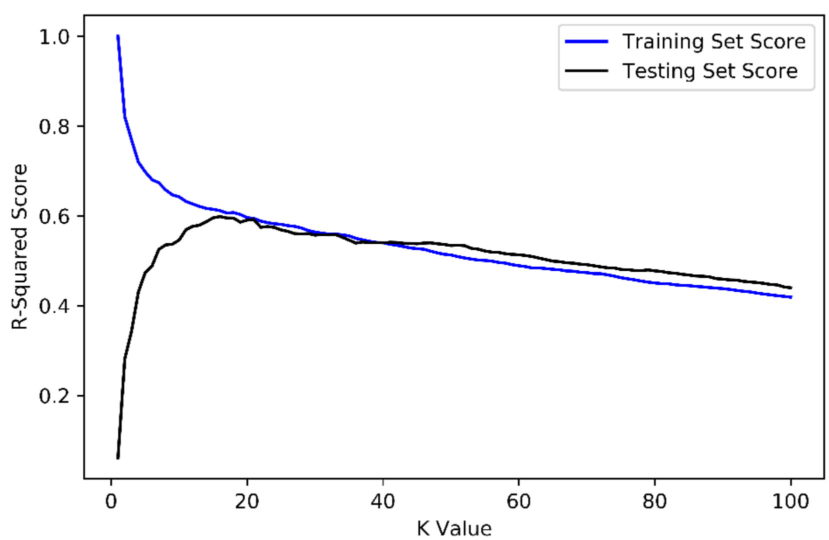

Figure S2 R-squared scores for different K values used for training the k-n algorithm and for testing the trained algorithm. The optimum K value is determined when the highest R-squared score is achieved for the testing dataset.

We used $K = 16$ to estimate the net efficiency as well as heat input and CO_2 emissions of the invalid datapoints and generated an NGCC database with 668 valid records. Table S2 summarizes the same statistical information as in Table S1 (Table S1 is for the 513 originally valid records). Values in parenthesis show the percentage change compared to the values in Table S1.

Table S2 Statistical information for 668 valid records after using the K-nearest regression model

Parameter	Min	Median	Mean	Max
Nameplate Capacity (MW)	4.7 (-40%)	335 (-27%)	412 (-12%)	1850 (0%)
Age (years)	0 (0%)	14 (0%)	17.3 (13%)	60 (18%)
Net Efficiency (HHV%)	35% (0%)	46.2% (1.5%)	44.7% (0.5%)	52% (0%)
Capacity Factor	0.0 (0%)	0.46 (-2%)	0.43 (-6.5%)	0.96 (0%)
LCOE pre-retrofit (2017$/MWh)	39.79 (0%)	57.96 (1.1%)	55.01 (-0.09%)	567,000
LCOE post-retrofit (2017$/MWh)	54.58 (0%)	90.24 (0.9%))	84.69 (0.15%)	-

As shown, the statistical description for the partially regressed data of all NGCC units and that of the originally valid datapoints are sufficiently similar. Therefore, we only used valid datapoints in our analysis. Figure S3 illustrates the difference in the results when all datapoints are used after validation by the machine learning method. The black curve is the same as in Figure 1b. The curve includes 513 valid datapoints and with the help of the conversion factor worked out in section S1, extrapolates the results from the valid units to all units. The blue curve shows the same analysis including the valid and regressed datapoints. As shown, the initial point and the shape of the cost curve does not significantly differ. The main difference is the lower ratio of retrofittable emissions to nonretrofittable and residual NGCC emissions when all the datapoints are used. This difference, however, is not significant and makes the results slightly biased in favor of postcombustion capture. This is because a higher ratio of total natural gas emissions is retrofittable when valid data are used for the analysis.

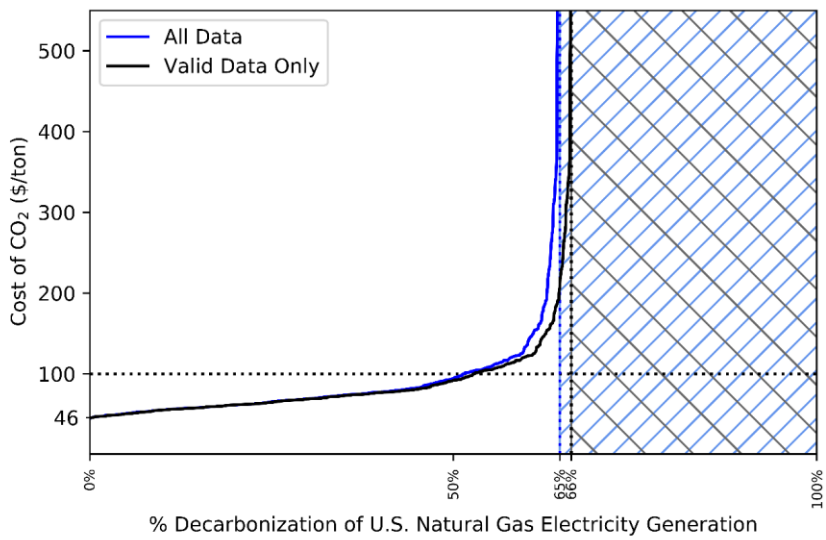

Figure S3 Comparison between cost curves plotted with the originally valid datapoints (black curve) and plotted with all data including the regressed datapoints with the machine learning algorithm (blue curve).

S4. Cost Analysis Model for Postcombustion Capture Retrofit Based on IECM

IECM allows the user to design a power plant with various types of fuel, power block design, cooling systems, and environmental control systems and provides outstanding flexibility for changing financing and cost parameters as well as design parameters for each section of the power plant.[2] The model has a systematic approach for calculating the cost, performance and mass balance around different fossil fuel power plants and emission control systems.

While IECM does not offer a cost analysis for a retrofitted NGCC unit, it provides information about the capital and operating cost of an amine system designed for that unit. The IECM software is also useful in quantifying the retrofit-induced changes in the power block and cooling system of an NGCC unit. We extracted this information and built a cost model around a known NGCC unit that calculates the cost of retrofit by postcombustion capture. The cost model was used to calculate the cost of retrofit for each existing US NGCC unit with valid data in our database.

IECM 11.2 only offers two models of NGCC gas turbine with fixed MW outputs, General Electric 7FB and 7FA. We chose the more efficient 7FB model which also has a higher capacity. The model only allows discrete values for the total capacity of an NGCC unit since the capacity of the steam turbine is fixed and only 1 to 5 gas turbines can be added to a unit. Therefore, cost information for only five different nameplate capacity values is available in the IECM (i.e., 295 MW, 590 MW, 885 MW, 1180 MW, and 1475 MW). The US NGCC units in our database, however, have a spectrum of nameplate capacities. We extracted the cost information for the five capacity values and used this information to interpolate the cost information of US NGCC units. Our cost model is mainly based on the total nameplate capacity and not the exact number of combustion and steam turbines.

In using the model, whenever a financial or technical/operational variable was not known, we used the default value in the IECM. We used the “Typical New Plant” option which includes an NGCC unit with a wet cooling tower as a default. Due to lack of unit-specific data, the cost of land was excluded from the analysis, but the introduced error is very small (the cost of land makes difference on the order of a few cents in $/MWh value of LCOE). We changed the natural gas composition to match the average US natural gas higher heating value (HHV) extracted from the eGRID database (22,442 Btu/lb Natural Gas).[3] Table S3 shows the natural gas composition used in our cost model.

Table S3 Natural gas composition

Natural Gas Component	Volumetric Percentage
Methane	87%
Ethane	9%
Propane	1.5%
Carbon Dioxide	1%
Nitrogen	1.5%
Total	100%

The units' retirement age and economic book life (amortization duration) were assumed to be 30 years. The book life for postcombustion units is assumed to be the remaining life of the NGCC unit and it cannot be lower than 5 years. In other words, only NGCC units younger than 25 years old are considered for retrofit. The relationship between the age and amortization level of an NGCC unit was assumed as shown in Figure S4.[2,4]

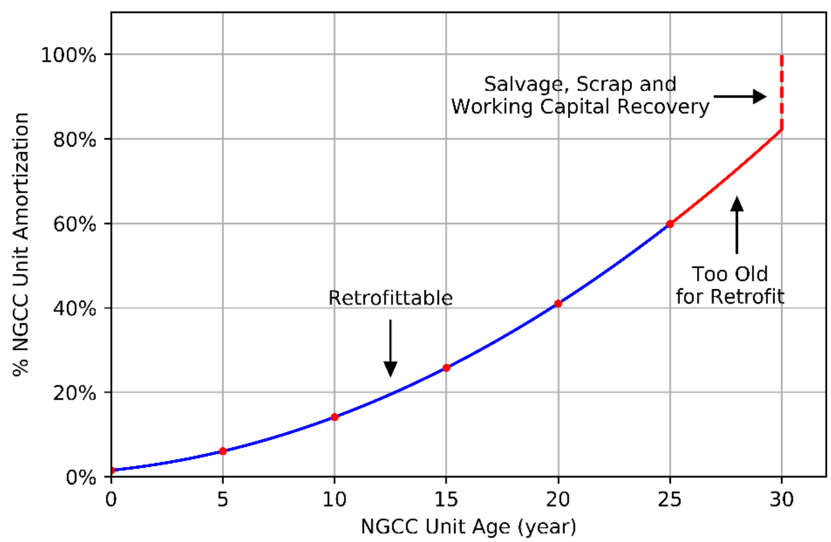

Figure S4 Relationship between age and amortization level of an NGCC unit in the cost model.

The Fixed Charge Factor (FCF), which is the fraction of the capital cost that must be recovered every year, is a function of a unit's remaining lifetime as well as the discount rate and the rate of return on different bonds and stocks and taxes. We used IECM default values to calculate the FCF for each unit.[2,5]

Capital and O&M costs for different unit sizes are extracted from IECM. Based on a zero net present value, the LCOE can be calculated for each unit:[5]

$$LCOE = \frac{NAMEPCAP \times CAP \times FCF + F_{O\&M}}{GENNTAN} + V_{O\&M} \tag{S3}$$

Where NAMEPCAP is unit's capacity (MW), CAP is the capital cost ($/MW), FCF is in (fraction/year), $F_{O\&M}$ is fixed O&M cost ($/MW/year), GENNTAN is the amount of electricity generated in one year (MWh) and $V_{O\&M}$ is variable O&M cost ($/MWh). Fuel cost is embedded in the variable O&M.

Retrofit-induced changes:

The IECM model and literature suggest an energy penalty equivalent to roughly 7-percentage point loss in a unit's efficiency after the retrofit.[2,6] This decreases the maximum available generation capacity (MW), while the amount of CO_2 produced per MWh of electricity increases. We used a 10-percentage point net efficiency loss since retrofitting an existing unit is typically harder than building a new unit with postcombustion capture.

Since the amount of CO_2 produced per MWh of electricity increases, after postcombustion capture, the amount of CO_2 released to the atmosphere is more than 10% of the CO_2 emission per MWh of electricity for the reference unit before the retrofit. This net CO_2 removal efficiency is typically around 88% since the postcombustion unit captures 90% of the already increased CO_2 production, not the initial CO_2 production. As mentioned in the main body of the article, the horizontal axes in Figures 1b and 4 do not take the additional CO_2 into account and the percentage values are relative to the initial CO_2 emissions of the reference units.

After the retrofit, the unit's new capital and operating costs were used in equation S3 to calculate the LCOE of the retrofitted unit. The additional capital cost after retrofit is not only due to the amine scrubber equipment but also due to the additional cooling capacity that is required. Fixed and variable O&M costs for a retrofitted unit were also estimated by comparison between a unit with and without postcombustion capture for each cost component. The capital cost of an amine scrubbing unit was multiplied by 1.15 to account for retrofit difficulties.[7] Capital and operating costs due to transportation and storage are not included in this analysis.

When LCOE and the rate of CO_2 emission for the reference and retrofitted units are calculated, the cost of avoided CO_2 (COC) can be determined by equation 1.

S5. Postcombustion Cost Model Sanity Check

To assess the reliability of the retrofit cost analysis model, we used two similar NGCC retrofit cost analysis studies.[8,9] We recalculated LCOE values before and after retrofit using the unit characteristics in these studies. As summarized in Table S4, the recalculated LCOE values and the difference between LCOE before and after retrofit are close to the published values. Based on equation 1, the difference between $LCOE_{retrofit}$ and $LCOE_{ref}$ is the key parameter in determining the cost of avoided CO_2. Therefore, the recalculated COC values are also close to the values in these studies. This validates the reliability of our retrofit cost model.

Table S4 Comparison between the retrofit cost analysis model in this study and models in the literature

Nameplate Capacity (MW)	Net Eff	Capacity Factor	Fixed Charge Factor	NG Price ($/MMBtu)	T&S Cost	$LCOE_{ref}$ ($/MWh)		$LCOE_{retrofit}$ ($/MWh)		$LCOE_{diff}$ ($/MWh)		COC ($/ton)		Ref
						Lit	Recalc	Lit	Recalc	Lit	Recalc	Lit	Recalc	
383	56.2%	0.85	0.13	6.35	10	53.9	52.1	76.6	76.3	22.7	24.2	78.8	83.7	8
171.3	52.7%	0.85	0.13	6.35	10	57.9	58.1	87.9	89.2	30.0	31.1	88.4	101.2	
77.9	48.1%	0.85	0.13	6.35	10	63.9	67.3	102.7	108.5	38.8	41.2	105	123	
970	45.6%	0.59	-	13.77	-	104	123	151	159	47.0	36	128	101.4	9
780	43.6%	0.68	-	13.77	-	106	126	152	162	46.0	36	119	97	

We used our default Fixed Charge Factor values for the last two recalculations. The relatively larger difference between COC values is due to different assumptions for the rate of emissions (ROEs) in different studies.

S6. Sensitivity to Natural Gas Price

We assumed a constant price of $4/MMBtu for natural gas. The estimated LCOE values are sensitive to the natural gas price, however, the cost of CO_2 capture, especially when analyzed for all NGCC units is not significantly sensitive to the natural gas price. Figure S5 illustrates the same results shown in Figure 1a with two natural gas prices, $2 and $8 per MMBtu.

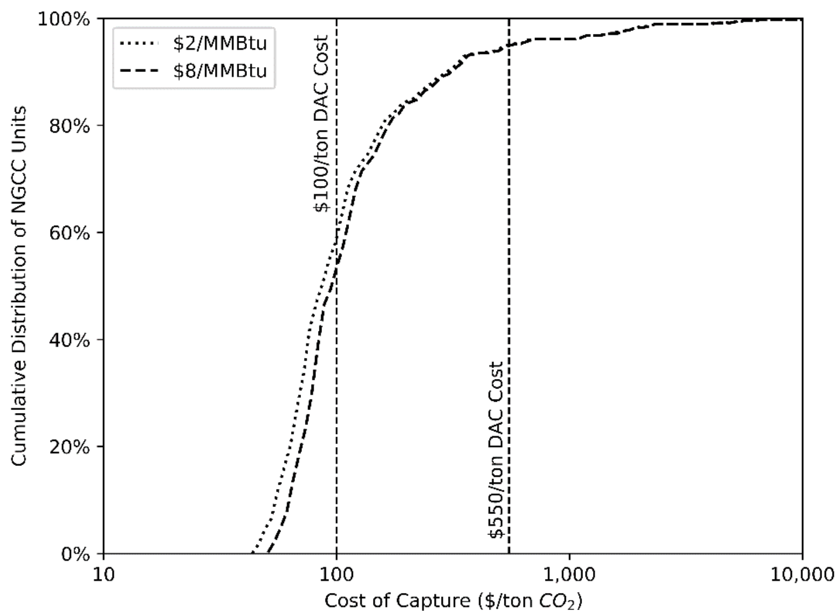

Figure S5 Cumulative distribution of the cost of CO_2 capture for the NGCC units considered for retrofit. The sensitivity of the results relative to the cost of natural gas is investigated with two natural gas prices.

As shown, the percentage of units with COC below $100/ton and below $550/ton is not strongly dependent on the price of natural gas.

S7. Incorporating Learning into the LCOE

Reduction in the COC through Learning-by-doing is reflected by a reduction in the cost of electricity for a retrofitted unit (smaller $LCOE_{retrofit}$) and lowering the retrofit energy penalty (smaller $ROE_{retrofit}$). This can be shown by considering the equation used for COC calculations:

$$COC\ (\$/ton\ CO_2)\ = \frac{LCOE_{retrofit} -\ LCOE_{ref}\ (\$/MWh)}{ROE_{ref} -\ ROE_{retrofit}\ (ton\ CO_2/MWh)} \tag{1}$$

Learning affects the capital and O&M costs of the postcombustion capture and as a result, each unit will have a lower post-retrofit LCOE compared to the post-retrofit LCOE when learning is not considered. All cost components in equation S3 are affected by learning when postcombustion capture is implemented in scale. Capital, fixed O&M and variable O&M costs of amine scrubbing units have different learning rates which are extracted from the work by van den Broek et al.[12] and summarized in Table 2 of the article. Due to the similarities between capital and fixed O&M costs, we assumed $F_{O\&M}$ has the same learning rate range as capital cost. We also assumed learning starts after retrofitting the first (cheapest) 3 GW of NGCC capacity and ends after retrofitting 100 GW of cumulative NGCC capacity. This means the first doubling in the learning equation (equation 2) happens when 6 GW of the NGCC capacity will have been retrofitted.

$$Y =\ a\,\varepsilon^{\log_2 X} =\ a\,X^{\log_2 \varepsilon} \tag{2}$$

When learning rates are known, the cost of each component after learning can be determined by equation 2 at any cumulative implementation level (parameter X). The problem in the case of postcombustion capture is that initial cost components (Cap, $F_{O\&M}$, and $V_{O\&M}$) at the beginning of learning (parameter a in equation 2) are not unique values and are different for each NGCC unit. To address this issue, we had to calculate the mean and standard deviation of each cost component over all the NGCC units considered for postcombustion capture. Then we used two constant values for each cost component, 1.5 standard deviations above and below the mean for that component. The table below, which is a part of Table 2 of this article summarizes these values. It also shows the final value of each component when the learning endpoint (100 GW) is achieved.

Table S5 Approximations for postcombustion cost components used in the learning calculations

Cost Component	Initial Value (3 GW Cumulative Capacity)	Final Value (100 GW Cumulative Capacity)
Cap ($\$/kW$)	\$550-\$1090	\$214-\$792
$F_{O\&M}$ ($\$/\mathrm{kW/year}$)	\$5.2-\$52.7	\$2.0-\$39
$V_{O\&M}$ ($\$/MWh$)	\$2.4-\$4.8	\$0.45-\$2.93
Energy Penalty	10%-point	7.1%-9.1%

In other words, we used the six cost values as proxies to simplify our model and provided an approximate cost range for each unit's postcombustion capture retrofit. To test the accuracy of the chosen proxy values, we used them to estimate a lower and an upper limit for the cost of postcombustion capture for each unit without learning. Namely, we tested them with the accurate postcombustion results calculated with the cost model and shown in Figure 1b. Figure S6 illustrates this comparison. The black curve shows the cost of postcombustion capture for the NGCC units (no learning effect included) versus the level of decarbonization as shown in Figure 1b. The shaded area around the curve shows the lower and upper limit estimates with the cost components in Table S5. The cost components are used in equation S3 to estimate the post-retrofit LCOE and then used in equation 1 to estimate a lower and an upper approximation for the cost of CO_2.

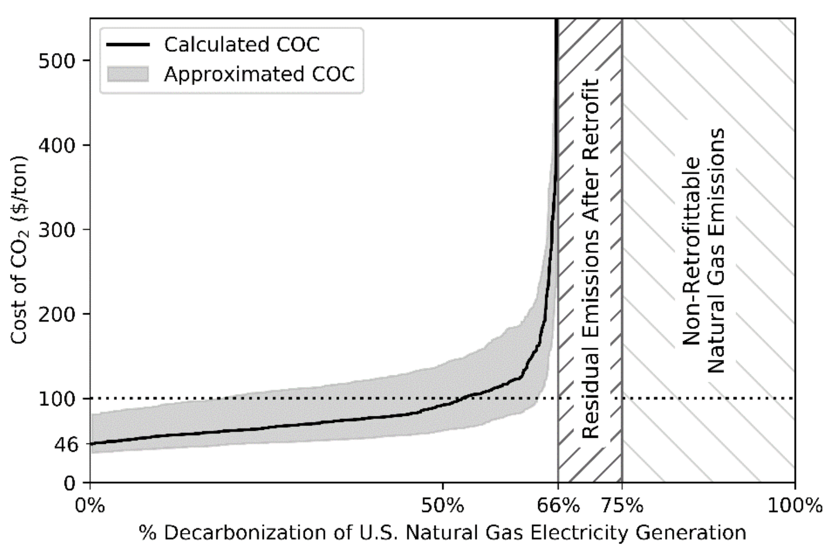

Figure S6 Comparison between the accurate calculation of COC and approximation with the proxy cost component values.

As shown in Figure S6, the upper and lower limit COCs are accurate approximations for the real cost of capture. Therefore, we expect the cost of postcombustion captures, with the impact of learning-by-doing, will fall between the upper and lower limit costs approximations.

The three lower cost values in addition to the three higher learning rates (quicker cost reduction) were used for a lower limit cost of CO_2 scenario and the three higher cost values with the lower learning rates (slower cost reduction) were used for an upper limit cost of CO_2 scenario. The learning results are shown in Figure 4.

S8. An Alternative Approach to Incorporate Learning in the Cost of Retrofit

As discussed in section S7, the unit-specific cost component values before learning (Cap, $F_{O\&M}$, and $V_{O\&M}$) make it difficult to project the impact of learning by equation 2. Here we propose an alternative approach to address this issue. By dividing equation 2 by parameter a (the initial cost before learning), we can reorganize this equation:

$$Y/a = \varepsilon^{\log_2 X} = X^{\log_2 \varepsilon} \tag{S4}$$

On the left side of equation S4, we have the ratio of a cost component after learning to its initial value with no learning. In the case of learning-by-doing, this ratio will be smaller than 1.0 and gets smaller when more experience is achieved (higher X). In this new approach, we recalculate the Y/a ratio after retrofitting each unit when more experience is accumulated. To calculate the cost of retrofit after learning for each unit, we simply multiply this ratio by the unique cost component of that unit:

$$Cost\ component\ after\ learning = (Y/a) \times Cost\ component\ before\ learning \tag{S5}$$

Equation S5 is used for each cost component of each unit to project the retrofit cost reduction after learning. Similar to the previous method, when cost components after learning are calculated, equations S3 and 1 are used to calculate the cost of CO_2. The Y/a ratio starts at 1.0 for the first 3 GW per-learning phase and its final values for different cost components are summarized in Table S6.

Table S6 Values of Y/a learning ratio for different cost components at learning endpoint

Cost Component	Learning Rate	Final Y/a Ratio (100 GW Cumulative Capacity)
Cap & $F_{O\&M}$	6%-17%	0.38-0.72
$V_{O\&M}$	10%-30%	0.18-0.60
Energy Penalty	2%-7%	0.70-0.91

Figure S7 illustrates the probable range of the postcombustion capture cost with learning. The difference between the two learning implementation methods can be noticed by comparing this figure with Figure 4a. Even though the average curves in the figures are almost identical, the authors prefer the method discussed in section S7 and Figure 4a for further analysis since they provide a larger range of uncertainty.

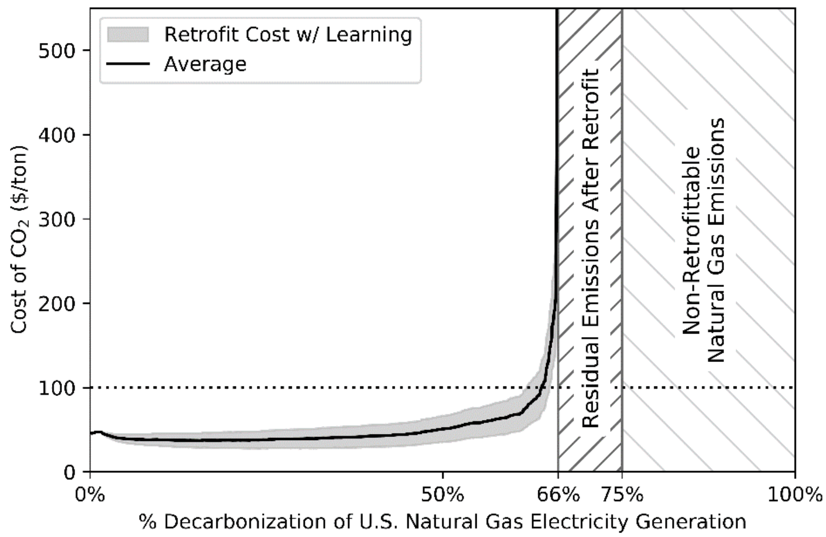

Figure S7 Cost of CO2 plotted against decarbonization level considering the impact of learning-by-doing for postcombustion capture retrofit. In this figure, we used an alternative method to project the cost reduction due to learning (compare with Figure 4a).

References

(1) 1.6. Nearest Neighbors — scikit-learn 0.22 documentation https://scikit-learn.org/stable/modules/neighbors.html#regression (accessed Dec 20, 2019).

(2) Carnegie Mellon University. *Integrated Environmental Control Model (IECM)*; 2018.

(3) *Emissions & Generation Resource Integrated Database (EGRID).*; 2016.

(4) Peters, M. Stone.; Timmerhaus, K. D.; West, R. E. *Plant Design and Economics for Chemical Engineers*; McGraw-Hill: Boston, 2006.

(5) Rubin, E. S. Understanding the Pitfalls of CCS Cost Estimates. *International Journal of Greenhouse Gas Control* **2012**, *10*, 181–190. https://doi.org/10.1016/j.ijggc.2012.06.004.

(6) Rubin, E. S.; Zhai, H. The Cost of Carbon Capture and Storage for Natural Gas Combined Cycle Power Plants. *Environ. Sci. Technol.* **2012**, *46* (6), 3076–3084. https://doi.org/10.1021/es204514f.

(7) Middleton, R. S.; Bielicki, J. M. A Scalable Infrastructure Model for Carbon Capture and Storage: SimCCS. *Energy Policy* **2009**, *37* (3), 1052–1060. https://doi.org/10.1016/j.enpol.2008.09.049.

(8) Carapellucci, R.; Giordano, L.; Vaccarelli, M. Studying Heat Integration Options for Steam-Gas Power Plants Retrofitted with CO2 Post-Combustion Capture. *Energy* **2015**, *85*, 594–608. https://doi.org/10.1016/j.energy.2015.03.071.

(9) Gerbelová, H.; Versteeg, P.; Ioakimidis, C. S.; Ferrão, P. The Effect of Retrofitting Portuguese Fossil Fuel Power Plants with CCS. *Applied Energy* **2013**, *101*, 280–287. https://doi.org/10.1016/j.apenergy.2012.04.014.

(10) Rubin, E. S.; Chen, C.; Rao, A. B. Cost and Performance of Fossil Fuel Power Plants with CO2 Capture and Storage. *Energy Policy* **2007**, *35* (9), 4444–4454. https://doi.org/10.1016/j.enpol.2007.03.009.

(11) *Updated Cost and Performance Estimates for Clean Coal Technologies Including CO2 Capture-2009*; EPRI, 2009.

(12) van den Broek, M.; Hoefnagels, R.; Rubin, E.; Turkenburg, W.; Faaij, A. Effects of Technological Learning on Future Cost and Performance of Power Plants with CO2 Capture. *Progress in Energy and Combustion Science* **2009**, *35* (6), 457–480. https://doi.org/10.1016/j.pecs.2009.05.002.