# CO2 output - an exploratory research report

Carbon emission is a multifaceted topic involving many causes and consequences. In this report, several of these facets will be touched upon and investigated, namely which predictors of co2 output there are and how they relate, which countries are showing efforts in reducing co2 output, and which renewable energy sources hold promise for the future in regards to costs and pricing.

## Predictors of CO2 output

***Introduction***

While it is clear that the main contributor to a large co2 output is the usage of fossil fuel as an energy source, co2 output and (fossil fuel) energy usage relate to a multitude of topics which are all, to an extent, intertwined. To shed some light on this “entangled mess”, it will be investigated how co2 output, energy usage and possibly relevant topics (GDP, urbanisation and education) relate.

***Data***

Various datasets were combined to obtain data on CO2 output, energy usage, GDP, urbanisation and education. These datasets contained information about many countries and many years. The most recent year for which at least some observations were available for all variables was 2017, and as such, this year was selected. Hence, all the following analyses were performed on data from 2017.

*Energy usage*

To confirm that fossil fuel energy usage is a major predictor of CO2 output and to determine which fossil fuel energy source is the largest predictor, a multiple regression analysis was performed with coal, oil and gas use per capita as predictors and CO2 output per capita as dependent variable. There were 75 countries that had the necessary data available to be included in the analysis. All variables were standardised in order to better compare the magnitude of the resulting coefficients.

*Socio-economic topics*

In addition to fossil fuel energy usage, three socio-economic topics that might be related to CO2 emissions were chosen to be investigated, namely GDP per capita, urbanisation and education. Urbanisation was measured in the percentage of the population that lives in an urban area. Education was measured in years of schooling. To gain initial insight into the interrelatedness of these variables and CO2 emissions, the correlations were calculated. To determine whether the chosen socio-economic topics are predictors of CO2 output, linear regressions were performed with each of these topics as an independent predictor, as well as a combined multiple regression analysis with all three variables as independent predictors. There were 171 countries that had the necessary data available to be included in the analysis. All variables were standardised in order to better compare the magnitude of the resulting coefficients.

## *Results*

*Fossil fuel energy usage*

The table below shows the statistics resulting from the multiple linear regression for fossil fuel energy usage and $CO_2$ emissions. All three energy sources were found to be significant predictors of $CO_2$ emission, with gas being the most influential predictor with a rather large standardized coefficient, indicative of a strong relationship with $CO_2$ emissions. For oil and coal, the coefficients ranged from small to moderate, indicative of at least some established relationship with $CO_2$ emission.

**Multiple linear regression results: Fossil fuel energy sources and CO2 emission**

| Predictor | Standardized coefficient | Standard error | T-value | P-value |
|---|---:|---:|---:|---:|
| Coal | 0.386 | 0.041 | 9.504 | <0.001 |
| Oil | 0.185 | 0.044 | 4.199 | <0.001 |
| Gas | 0.827 | 0.045 | 18.495 | <0.001 |

*Socio-economic topics*

The correlation matrix below shows that $CO_2$ emission per capita is correlated with the three socio-economic topics, with all correlations exceeding $r = 0.5$ Notably, the relationship between $CO_2$ emission per capita and GDP per capita is most prominent, with a correlation of $r = 0.75$. Since the data was standardised, the individual linear regression coefficients for the socio-economic variables are the same as their correlation with $CO_2$ emission. Through the regression it could be confirmed that these relationships were highly significant. It can also be seen in the correlation matrix that the three socio-economic variables are highly related to each other, with these correlations also exceeding $r = 0.5$. This is important to note, as this interconnectedness can cloud the true nature of the relationship with $CO_2$ emission. E.g. if a true relationship between GDP and $CO_2$ emission exists, and GDP is also related to schooling, schooling will show as having a relationship with $CO_2$ emission when not accounting for GDP, whether there is a true relationship between schooling and $CO_2$ emission or not. As such, a multiple regression analysis was performed with all socio-economic variables and $CO_2$ emission. Indeed, the relationships between schooling, urban percentage and $CO_2$ emission were no longer significant in this combined analysis and are completely driven by GDP per capita. These results do confirm that GDP per capita is a major predictor for $CO_2$ emission, as that relationship remained highly significant with a large standardized coefficient.

**Correlation matrix socio-economic topics and CO2 emission**

| | CO2 emission per capita | GDP per capita | Schooling | Urban percentage |
|---|---:|---:|---:|---:|
| **CO2 emission per capita** | 1 | 0.752 | 0.51 | 0.519 |
| **GDP per capita** | 0.752 | 1 | 0.657 | 0.648 |
| **Schooling** | 0.51 | 0.657 | 1 | 0.591 |
| **Urban percentage** | 0.519 | 0.648 | 0.591 | 1 |

**Linear regression results: Individual socio-economic variables and CO2 emission**

| Predictor | Standardized coefficient | Standard error | T-value | P-value |
|---|---|---|---|---|
| GDP per capita | 0.752 | 0.051 | 14.821 | <0.001 |
| Schooling | 0.51 | 0.066 | 7.707 | <0.001 |
| Urban percentage | 0.519 | 0.066 | 7.887 | <0.001 |

**Multiple linear regression results: Socio-economic variables and CO2 emission**

| Predictor | Standardized coefficient | Standard error | T-value | P-value |
|---|---|---|---|---|
| GDP per capita | 0.71 | 0.075 | 9.509 | <0.001 |
| Schooling | 0.014 | 0.071 | 0.195 | 0.845 |
| Urban percentage | 0.05 | 0.07 | 0.723 | 0.471 |

## Decreasing CO2 output

***Introduction***

A general consensus has been reached, at least in the scientific community, that the output of co2 must be significantly decreased in order to reduce further damage to the environment and slow down climate change and its devastating consequences, which are becoming more apparent by the year. By looking at countries' co2 output per capita over a timespan of 20 years (2001-2021), it will be investigated which countries have been consistently decreasing their co2 output and are making the biggest strides in further reducing carbon emissions.

***Data***

To determine the changes in co2 output per capita for each country, the relative percentage changes were calculated for the years between 2000 and 2021 for each country. With the available data, it was possible to calculate these for 218 out of 232 countries. To determine which country had the most consistent decrease in their co2 output, the median percentage change for the 20 years was calculated. The median was chosen over the mean, because the mean is more likely to be influenced by outlier years in a series with relatively little data points like this one. Additionally, a linear regression was performed on the relative percentage change over time to determine whether there was a negative trend, indicative of an increased effort to reduce co2 output relative to the prior year.

***Results***

*Consistent decrease*

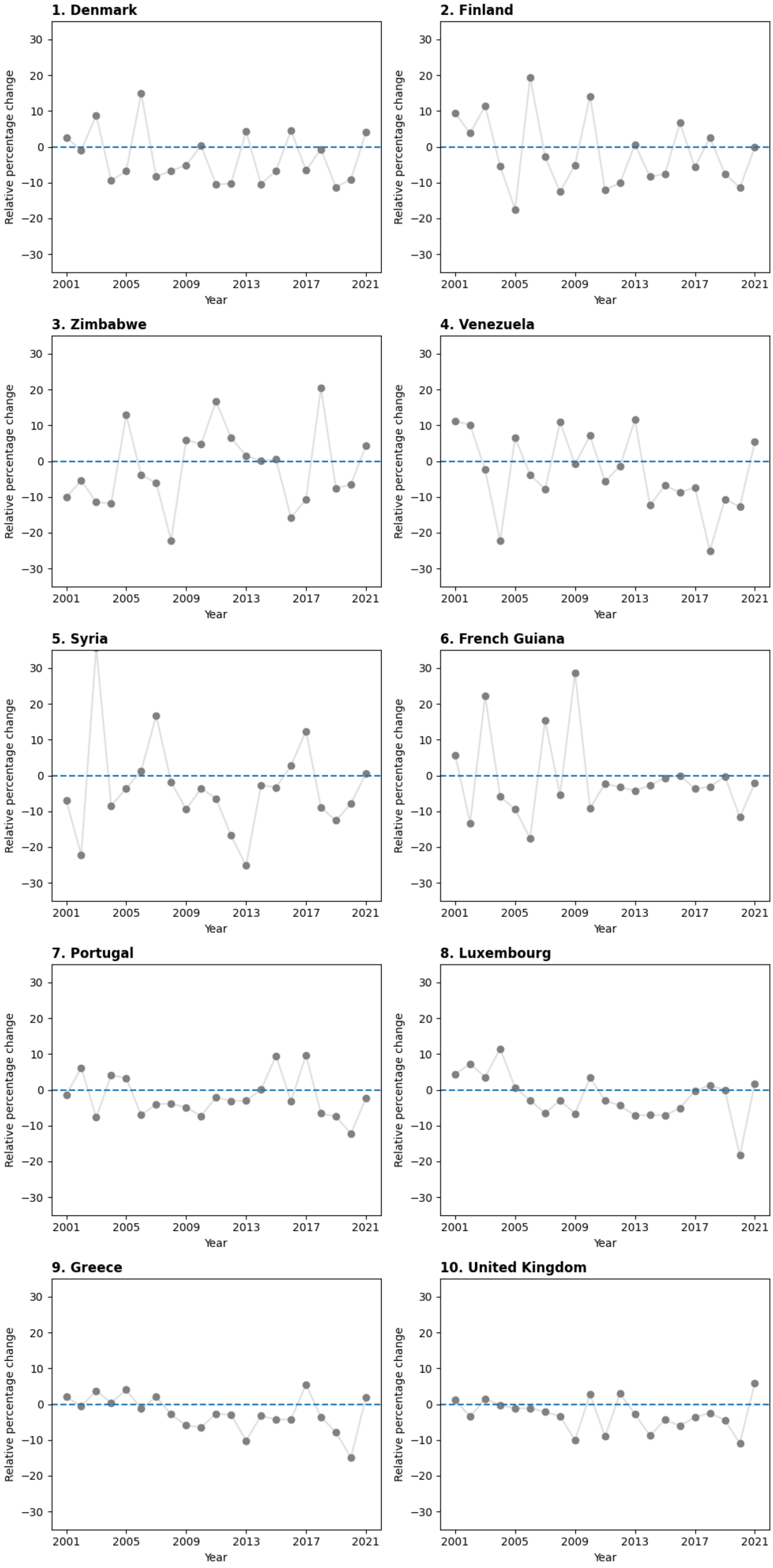
The top 10 countries with the largest median relative percentage change can be seen in the table and the figure below. These countries have been most consistently decreasing their CO2 output, relative to the year prior.

### Top 10 countries consistently decreasing CO2 output

| Country | Median change | Mean change | Years with decrease |
|---|---|---|---|
| Denmark | -6.58 | -3.02 | 14 |
| Finland | -5.09 | -1.82 | 13 |
| Zimbabwe | -3.89 | -1.81 | 11 |
| Venezuela | -3.79 | -3.11 | 14 |
| Syria | -3.66 | -3.37 | 15 |
| French Guiana | -3.2 | -1.1 | 16 |
| Portugal | -3.16 | -2.08 | 15 |
| Luxembourg | -2.95 | -1.8 | 13 |
| Greece | -2.87 | -2.45 | 14 |
| United Kingdom | -2.77 | -2.85 | 16 |

**Top 10 countries consistently decreasing CO2 output**

Relative percentage change (2001-2021) in CO2 output



*Increased efforts*

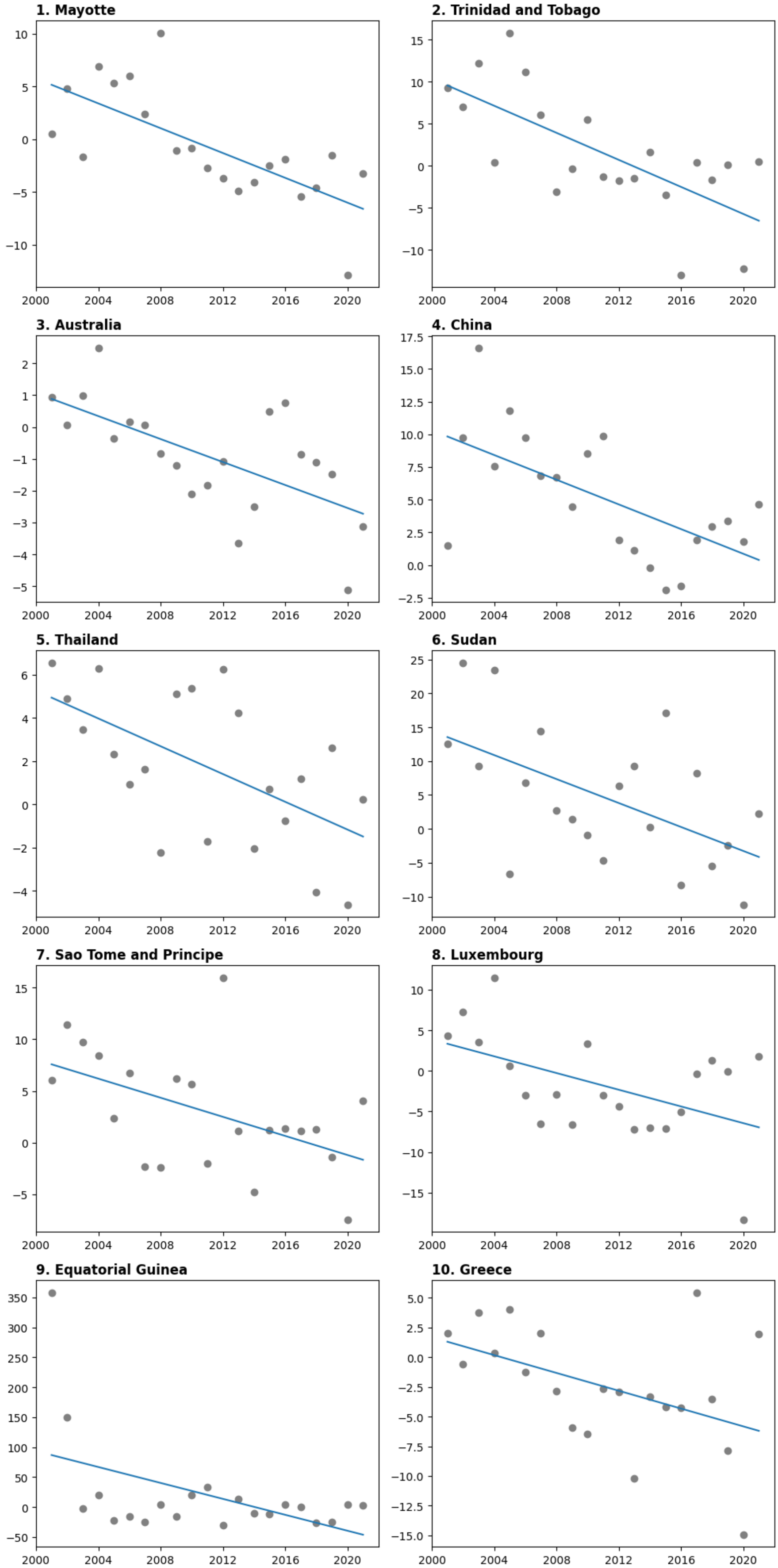
The table below shows the results of the linear regressions on the relative percentage change (RPC) over the years 2001-2021 for the 10 countries that had the largest negative trend in RPC over time. This negative trend is indicative that there have been efforts to reduce the increases in $CO_2$ output and/or enlarge the decreases in $CO_2$ output. These negative trends are also graphically presented in the figure below.

**Top 10 countries increasingly reducing CO2 output**

| Country | R-value | P-value | Slope | Intercept |
|---|---|---|---|---|
| Mayotte | -0.703 | 0.00037 | -0.588 | 1181.7 |
| Trinidad and Tobago | -0.695 | 0.00047 | -0.804 | 1617.858 |
| Australia | -0.636 | 0.00195 | -0.18 | 362.006 |
| China | -0.613 | 0.00315 | -0.473 | 955.439 |
| Thailand | -0.581 | 0.00579 | -0.321 | 647.245 |
| Sudan | -0.551 | 0.00956 | -0.884 | 1783.295 |
| Sao Tome and Principe | -0.505 | 0.01952 | -0.462 | 932.376 |
| Luxembourg | -0.501 | 0.02062 | -0.515 | 1033.547 |
| Equatorial Guinea | -0.479 | 0.02808 | -6.654 | 13402.211 |
| Greece | -0.466 | 0.03305 | -0.375 | 751.186 |

**Top 10 countries increasingly reducing CO2 output**

Relative percentage change (2001-2021) in CO2 output



## Renewable energy costs

***Introduction***

Moving away from fossil fuel and toward renewable energy sources is an important step in both preventing an energy crisis as fossil energy sources are further depleted, and reducing co2 output and the negative consequences associated. While renewable energy sources seem to be the better choice, only a moralistic few will actively choose these options when the costs are higher than the costs for more traditional options. Most will opt for the cheapest option without regard for long term consequences, even when they have the means. Additionally, many households and (small) businesses do not actually have the means to consciously choose more environmentally friendly options when costs are higher. In such, it is important that renewable energy will be widely available at a competitive, if not lower, cost. Here, the changes in energy costs for several renewable energy sources are examined to predict which source holds the most promise to become the most affordable.

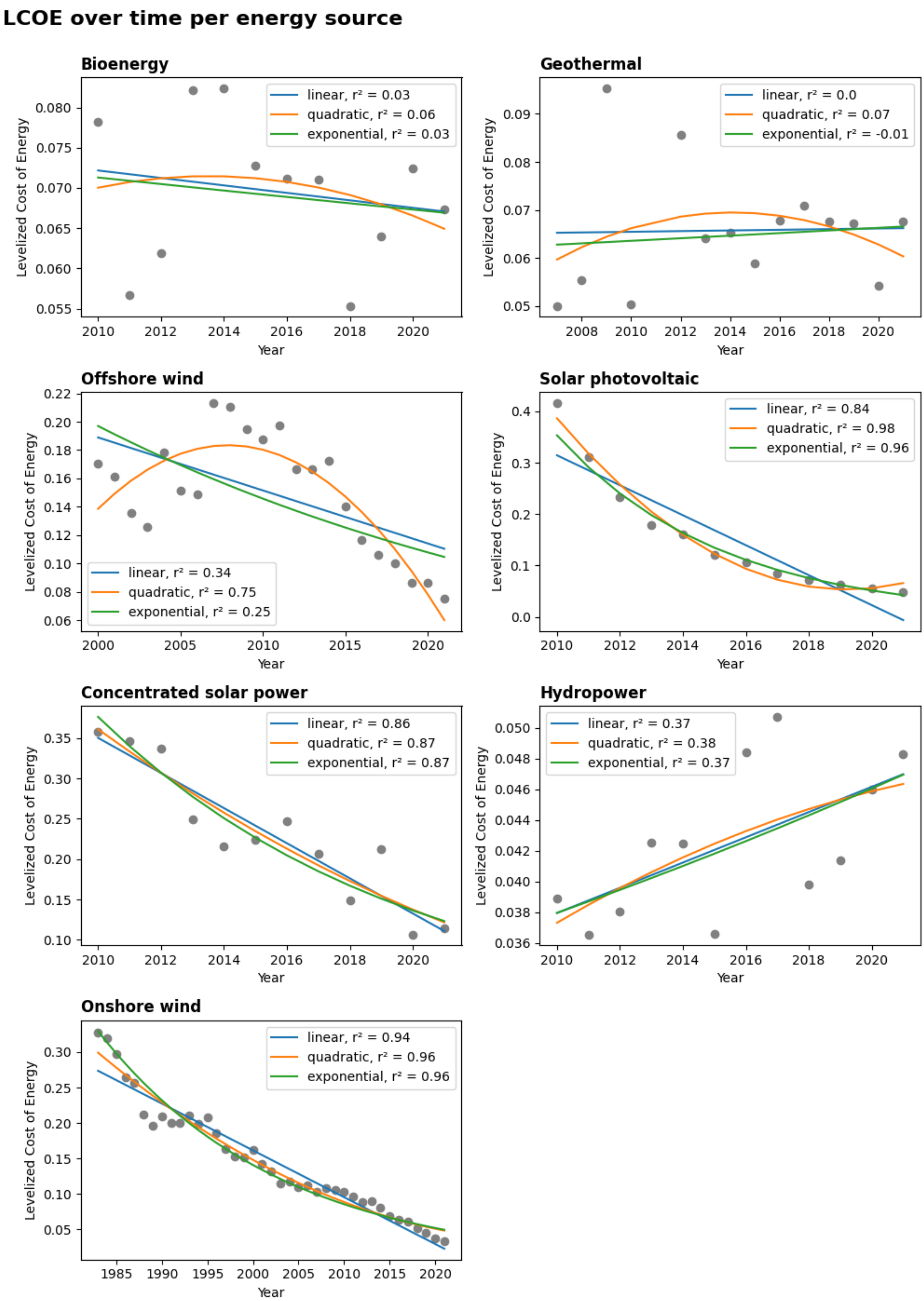
***Data***

Energy costs are reflected in a metric called levelized cost of energy (LCOE). The LCOE takes into account the sum of lifetime costs of the energy system and the lifetime energy production of the energy system. The costs are divided by the production to provide the cost per unit of energy. LCOE data was available for 7 renewable energy sources, with differing amounts of years for which data was available, as can be seen in the table below.

**Number of entries per energy source**

| Energy source | Earliest entry | Most recent entry | Number of entries |
|---|---|---|---|
| Bioenergy | 2010 | 2021 | 12 |
| Geothermal | 2007 | 2021 | 14 |
| Offshore wind | 2000 | 2021 | 22 |
| Solar photovoltaic | 2010 | 2021 | 12 |
| Concentrated solar power | 2010 | 2021 | 12 |
| Hydropower | 2010 | 2021 | 12 |
| Onshore wind | 1983 | 2021 | 39 |

To be able to predict future LCOE, the trend which the LCOE has followed over time has to be established. An initial look at the LCOE over time showed that not for all sources of renewable energy a linear trend could be clearly seen in the data. As such, in addition to a linear model, a second degree polynomial (quadratic) and an exponential model were also fitted to the data for each energy source. Based on visual inspection of how well the models fit the data and their respective r-squared values (as can be seen in the figure below), an informed decision was made on which model would be used for each energy source to predict future LCOE.

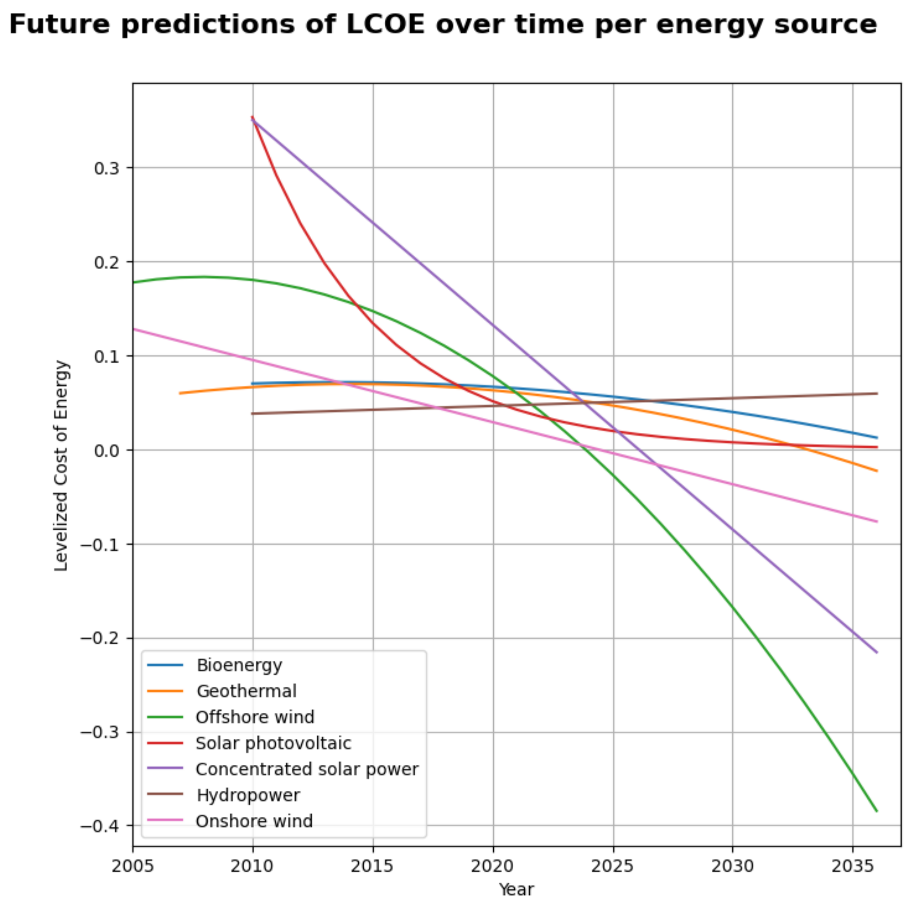



**Bioenergy.** No clear pattern can be seen in the data and while none of the models explain large amounts of variance, the quadratic model has a somewhat higher r² and as such will be used to predict future costs.

**Geothermal.** No clear pattern can be seen in the data and while none of the models explain large amounts of variance, the quadratic model has a somewhat higher $r^2$ and as such will be used to predict future costs.
**Offshore wind.** The data can be seen to follow a quadratic pattern, which is supported by the higher $r^2$ relative to the other two models. The quadratic model will be used to predict future costs.
**Solar photovoltaic.** The data follows a curved trend, which can be captured in both a quadratic and exponential model. While the quadratic model has a slightly higher $r^2$, visual inspection shows that the exponential model approaches the data more closely in more recent years and as such will be used to predict future costs.
**Concentrated solar power.** Visual inspection shows that the data seems to follow a relatively linear pattern and all models explain similar amounts of variance. In such, the most parsimonious model, being the linear model, will be used to predict future costs.
**Hydropower.** The data does not show a clear pattern and all three models explain similar and moderate amounts of variance. The most parsimonious model, being the linear model, will be used to predict future costs.
**Onshore wind.** The data follows an approximate linear trend. While the quadratic and exponential models, which account for the curve visible in the data in earlier years, have a slightly higher $r^2$, visual inspection shows that the linear model better captures the linear trend in more recent years and as such will be used to predict future costs.

***Results***

Using the models that seemed to have the most approximate fit to the available data, the LCOE for 2022-2035 were predicted for each energy source. As can be seen in the graph below, the LCOE for most energy sources is predicted to decrease in the coming years. It is important to take into consideration that some of the models are based on a more limited amount of data entries (e.g. bioenergy, $n = 12$), and in such, their predictions are likely to be less reliable than predictions made with models based on more data entries (e.g. onshore wind, $n = 39$). Similarly, models that have a poor fit (e.g. geothermal, $r^2 = 0.07$) are likely to yield less reliable predictions than models with a much better fit (e.g. solar photovoltaic, $r^2 = 0.96$). The most promising energy sources in regard to a lower future LCOE seem to be offshore wind, concentrated solar power and onshore wind. The model for offshore wind is based on a decent amount of data entries and has a relatively good model fit ($n = 22$, $r^2 = 0.75$). The model for concentrated solar power ($n = 12$, $r^2 = 0.86$) is based on less data entries, but has a somewhat better model fit. Lastly, the model for onshore wind ($n = 39$, $r^2 = 0.94$) is based on the most data entries and has the best model fit as well. Out of the three most promising energy sources, the predictions for onshore wind are likely to be the most reliable. Nonetheless, the predictions for all these three energy sources do give an approximate indication of what to expect of their LCOE in the future and how they are likely to be the energy sources with the lowest cost in times to come. While a lower LCOE is not a direct measure of a lower consumer price, it is very likely that a lower LCOE is also reflected in a lower consumer price and in such, likely that these three energy  will have the best price in the future.




## Conclusions

***Predictors of CO2 output***
Unsurprisingly, all three fossil energy sources and the usage thereof per capita are predictors for CO2 emissions per capita. All three have a highly significant relationship with CO2 emissions, with the relationship between gas usage and CO2 emissions being notably large. Regarding the socio-economic topics, there seemed to be relationships between schooling and urbanisation, and CO2 emissions, but these were largely driven by their connections to GDP per capita, which in itself did have a significant and strong relationship with CO2 emissions. Based on the analyses conducted in this report, it could be concluded that oil, gas, and coal usage per capita, and GDP per capita are predictors of CO2 output.

***Decreasing CO2 output***
After investigating the relative percentage change in CO2 emission over the time period of 2001-2021, it could be concluded that Denmark, Finland and Zimbabwe are the countries that have been most consistently decreasing their CO2 output over the years. It could also be concluded that some countries have been reducing the increases in CO2 output and/or enlarging the decreases in CO2 output, possibly reflecting a more active approach for reducing CO2 output. This trend was most notable for Mayotte, Trinidad and Tobago and

Australia. So, the six mentioned countries are the countries that have either consistently or actively made strides in decreasing $CO_2$ output.

***Renewable energy costs***

After investigating the trends in LCOE for renewable energy sources seen in the last years, predictions were made for the future costs. While almost all renewable energy sources are predicted to become somewhat lower in costs, the most notable are offshore wind, concentrated solar power and onshore wind, which are predicted to be the most likely to keep decreasing to much lower costs relative to the other sources. Hopefully, this will also become reflected in consumer prices and will make it so that offshore wind, concentrated solar power and onshore wind have the best price for the consumer in the (near) future.