Increase AM Yield: Milestone Report

GitHub Project Link

Instructions

Think of a milestone report as an interim report that you may be asked to share with your client to keep them updated on your findings. It's also an opportunity for you to take stock of how far you've come, what you've found, and practice your data storytelling skills. This is similar to an early draft of the Final Capstone Project 1 Report.

The milestone report compiles all the reports that you've been writing throughout the course. Hopefully, you've been keeping your findings organized and documenting in a systematic manner. You should not need to do any new data analysis for this report.

Steps

1. Write a your Capstone Project 1 Milestone Report and include the following:
    a. 5-6 page Google Doc
    b. **Problem Statement**: Why it's a useful question to answer and for whom
        i. Proposal
    c. **Dataset**: Description of the dataset, how you obtained, cleaned, and wrangled it
        i. Data Wrangling Report
    d. **Findings**: Initial findings from exploratory data analysis
        i. Data Story and Inferential Statistics Reports
        ii. Summary
        iii. Visuals and Statistics to support Findings
2. Update your presentation slides
3. Update your GitHub repository with the capstone project 1 code, milestone report, document, and slides
4. Use the link below to share your report with your mentor for feedback, and update as needed
5. Convert to PDF and add to your repository. Share with your peer community.

Submission

Write a 5-6 page report on the steps and findings of your project so far. Upload this report to your GitHub and submit a link.

Rubric

Learning Objective

- Utilize predictive models appropriate to your story.
- Learn how to apply skills in data collection and wrangling, data storytelling and inferential statistics to a project utilizing real world data.

Criteria

- Completion
    - 5-6 pages google document (commentable by mentor)
- Process & Understanding
    - The submission demonstrates an understanding of how to describe a dataset and detail how it was obtained, cleaned, and wrangled.
    - The submission demonstrates successful application of exploratory data analysis (visualization and inferential statistics), for example histograms, scatter plots and hypothesis testing appropriately.
    - The submission demonstrates successful application of data storytelling techniques appropriate to their target audience, clearly articulating hypotheses and inferences.
- Presentation
    - Document is comprehensive, detailed and actionable
    - Presentation is clear, readable and polished
- Excellence
    - The submission includes innovative ways to visualize the data, revealing surprising insights. The report is shared with the Springboard community.

Report

Outline

Problem Statement

- **Reference**: Proposal
- Why it's a useful question to answer and for whom

Objective

The objective of the project is to increase the yield of Additive Manufacturing (AM) by correlating data sets throughout the build. This increase in yield/efficiency will increase the profits of the 3D printing service bureaus and/or decrease the costs to consumers of the parts.

Current Situation

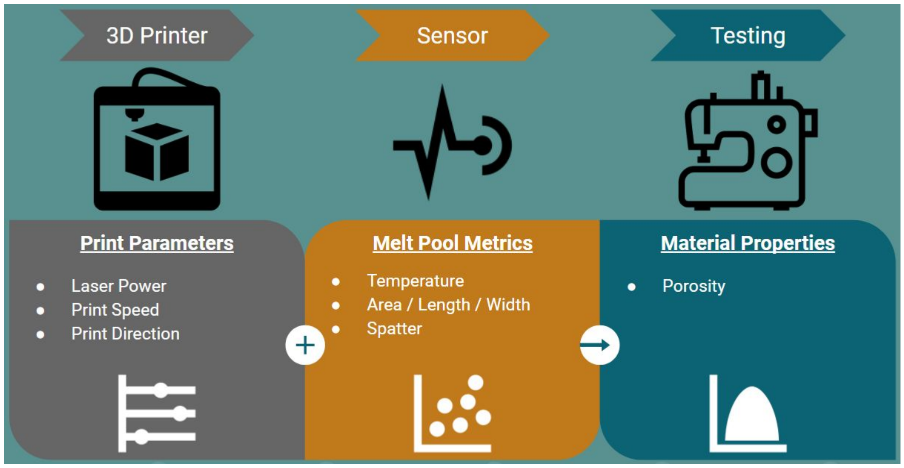
The current state of AM technology is that there is no digital thread throughout the entire process. This leads to inconsistent results in the final part and no way of tracking what input affects what output as experiments are conducted. The only way to understand this complex unknown process further is to digitally connect each sub-process to the next one, piece by piece and begin to understand the physics of the whole manufacturing process through empirical data. The industry is trying to predict how the input parameters affect the outcome of the final part.

Plan

There are three main sets of data:

- **PP** = Print Parameters (Before the Build)
- **MPM** = Melt Pool Metrics (During the Build)
- **MP** = Material Properties (After the Build)

By linking together the data sets, a correlation can be drawn between the input and the output. This allows the user to select the ideal Print Parameters that best generate a fully dense part (Porosity = 0%).


3D Printer
Sensor
Testing
Print Parameters
Laser Power
Print Speed
Print Direction
Melt Pool Metrics
Temperature
Area / Length / Width
Spatter
Material Properties
Porosity

Dataset

- **Reference**: Data Wrangling Report
- Description of the dataset, how it's obtained, cleaned, and wrangled

Data Description

The **Print Parameters** are set up by the user prior to printing. These include Laser Power, Scan Speed, Layer Height, etc. For this experiment a Condition is a set of Print Parameters and are set to be constant throughout each build.

The **Melt Pool Metrics** are collected through a sensor during the printing process. These include such things as Intensities, Temperatures, and Dimensions. They are used to monitor the process. In a separate project, irregularities could be used to find defects inside the builds that would lead to cracks and failures.

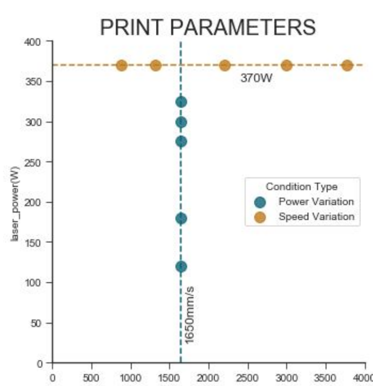
The **Material Properties** are tested for and calculated after the printing process. In this case, Porosity was the only selected property to be studied.

Data Gathering

The **Print Parameters** were grouped into a matrix by the user. For half the conditions (1-5), the Laser Power the varied while the Scan Speed was kept constant. For the second half of the conditions (6-10), the Scan Speed was varied while the Laser Power was kept constant.

Category	Laser Power (W)	Scan Speed (mm/s)	Hatch Spacing (mm)
Nominal	300	1650	0.09
Power Variation	325	1650	0.09
	300		
	275		
	180		
	120		
Speed Variation	370	3780	0.14
		3000	
		2200	
		1320	
		880	

	laser_power(W)	scan_speed(mm/s)	hatch_spacing(mm)	cond_type
condition				
1	325	1650	0.09	PV
2	300	1650	0.09	PV
3	275	1650	0.09	PV
4	180	1650	0.09	PV
5	120	1650	0.09	PV
6	370	3780	0.14	SV
7	370	3000	0.14	SV
8	370	2200	0.14	SV
9	370	1320	0.14	SV
10	370	880	0.14	SV



The **Melt Pool Metrics** were collected and calculated through a sensor that was placed inside the printer. The output from the sensors is divided into two tables (viz and threshold), one of each for each run. It monitors the build throughout the entire process and generates metrics to help understand the physics further by supplementing simulation teams and make more consistent builds.

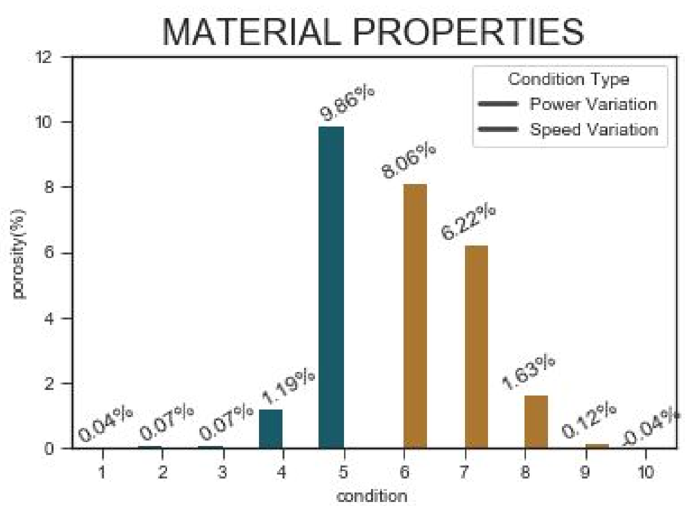
Metric	Units
Frame	-
Time	μs
Threshold	-
Peak X	pixels
Peak Y	pixels
Peak Temperature	℃
Average Temperature	℃
Area	$micron^2$
Length	micron
Width	micron
Spatter Counts	-
Spatter Area	pixels
X - Profile	℃
Y - Profile	℃

	run	frame	time(s)	exp_time(ms)	int_s_p(counts)	int_l_p(counts)	int_s_avg_3(counts)	int_l_avg_3(counts)
0	16	51	0.064	0.024957	2788	1533	2010.1	1073.4
1	16	52	0.065	0.024957	2969	1579	1989.0	1177.9
2	16	53	0.066	0.024957	3185	1719	2259.8	1218.9
3	16	54	0.068	0.024957	3173	1783	2288.6	1375.7
4	16	55	0.069	0.024957	3149	1901	2478.2	1219.0

	run	frame	t1_temp_avg(C)	t1_length(pixels)	t1_width(pixels)	t1_orient(degrees)	t1_area(pixels)	t1_sat_num(-)
0	16	51	1834.1	79.429	19.615	-84.970	931	5
1	16	52	1856.9	72.962	21.994	-86.215	1036	18
2	16	53	1842.3	58.080	14.408	88.929	525	13
3	16	54	1853.2	39.585	12.303	-88.745	324	16
4	16	55	1804.5	32.979	11.086	89.202	230	7

The **Material Properties**, in this case Porosity, is calculated by a testing machine. Each sample is given an average Porosity value across the entire build.

	porosity(%)
condition	
1	0.039
2	0.068
3	0.067
4	1.189
5	9.860
6	8.060
7	6.225
8	1.630
9	0.115
10	-0.044



Data Cleaning

All data tables were imported as individual CSV files and then converted into Pandas DataFrames. This was each table could be dealt with separately until end where all the data could be combined into a single table. Units for each column were set to the correct variable types (int, float, category, boolean, etc.)

The **Print Parameters** were cleaned in some fairly simple ways. Constants were extracted out of the table into number variables to simplify the table. Other variables including scale factor and conversion factors were set up as well. There was also a Run Log used as reference to give more detail about the conditions.

The **Melt Pool Metrics** were combined from the viz and threshold tables. Then by utilizing the other tables, the index could be converted from a simple run to a condition/run/layer/frame. That way, each row had a unique index and could be combined with all the tables correctly. Intensities were normalized by converting counts to counts/ms to account for the different exposure times used by the sensor. Dimensions were converted from pixel units to micron units.

				time(s)	scan_direction(xy)	int_s_p(counts/ms)	int_l_p(counts/ms)	int_s_avg_3(counts/ms)
condition	run	layer	frame					
1	16	1	51	0.064	y	111712	61425	80542
			52	0.065	y	118964	63268	79697
			53	0.066	y	127619	68878	90547
			54	0.068	y	127138	71442	91701
			55	0.069	y	126177	76171	99298

The **Material Properties** were not cleaned or adjusted.

Data Wrangling

Additional **Print Parameters** were calculated utilizing a combination of preselected parameters called Energy Densities (LED, GED, VED). These variables help combine multiple parameters into a single column (Laser Power, Scan Speed, Hatch Spacing, and Layer Height)

- **LED** = Linear Laser Energy Density = LP/SS
- **GED** = Global Energy Density = LP/(SS*HS)
- **VED** = Volumetric Laser Energy Density = LP/(SS*HS*LH)

	laser_power(W)	scan_speed(mm/s)	hatch_spacing(mm)	cond_type	led	ged	ved
condition							
1	325	1650	0.09	PV	0.20	2.19	72.95
2	300	1650	0.09	PV	0.18	2.02	67.34
3	275	1650	0.09	PV	0.17	1.85	61.73
4	180	1650	0.09	PV	0.11	1.21	40.40
5	120	1650	0.09	PV	0.07	0.81	26.94
6	370	3780	0.14	SV	0.10	0.70	23.31
7	370	3000	0.14	SV	0.12	0.88	29.37
8	370	2200	0.14	SV	0.17	1.20	40.04
9	370	1320	0.14	SV	0.28	2.00	66.74
10	370	880	0.14	SV	0.42	3.00	100.11

An additional **Melt Pool Metric** was created, Length to Width Ratio, to discover if the aspect ratio of the melt pool played a part in understanding the process. Also, for simplicity, the metrics were categorized into groups

- **General** - Frame, Time, etc.
- **Intensities** - Long/Short Wavelength with Peak, 3, and 5 Pixel Masks
- **Peak Temperatures** - Reference Hybrid Temperature and Peak Temperature
- **Thresholds** - t(1-4) Average Temperature, Area, Length, Width, etc.

The **Material Properties** were not wrangled.

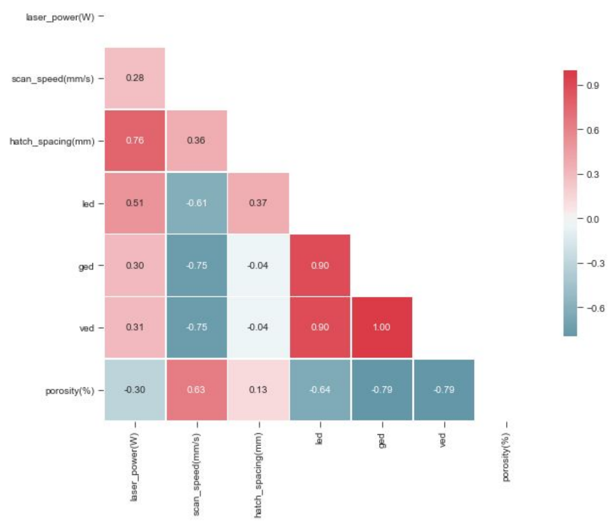
Findings

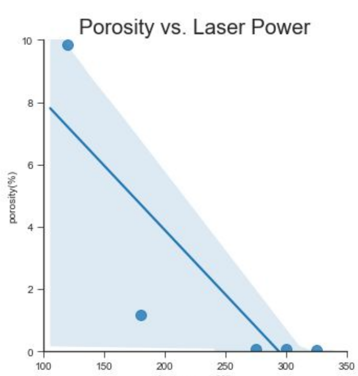
- **Reference**: Data Story and Inferential Statistics Reports
- Exploratory Data Analysis, Plots, Stats, and Summary

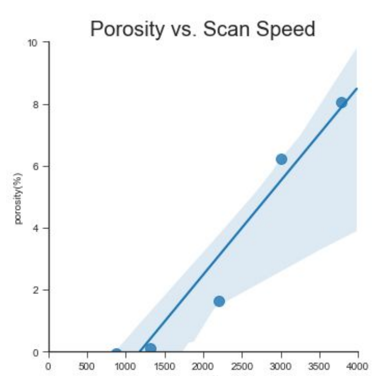
Exploratory Data Analysis

Print Parameters vs. Material Properties

The two main Print Parameters (Laser Power and Scan Speed) generally correlate with Porosity (-0.30 & 0.63) respectively. As Laser Power increases, Porosity tends to decrease, but as Scan Speed increases, Porosity tends to increase. But when Feature Engineering is used to create the Energy Densities like GED and VED, the correlation values increase greatly to -0.79.





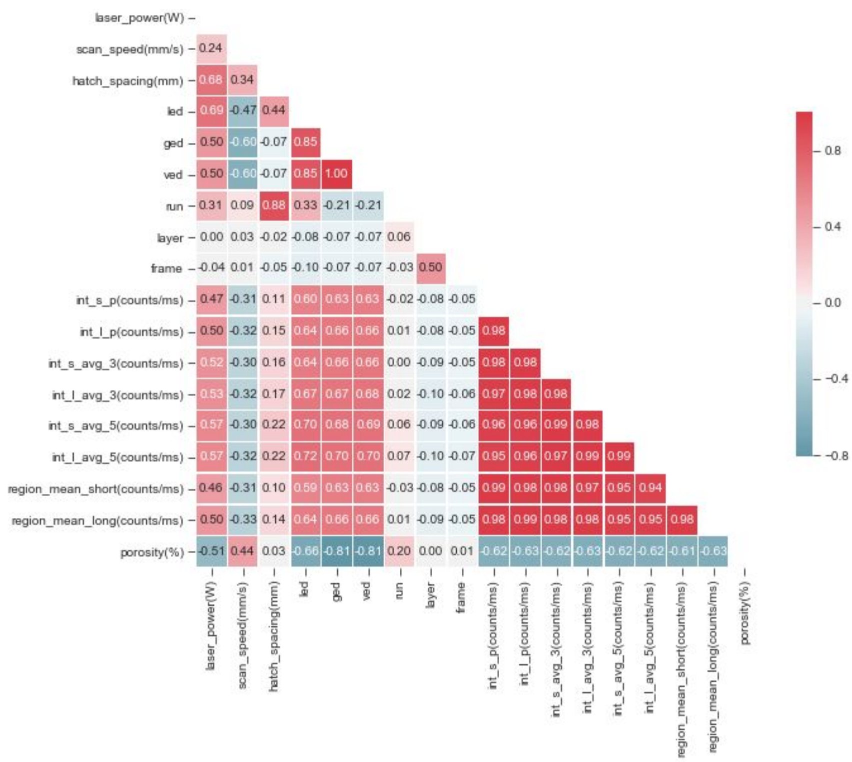


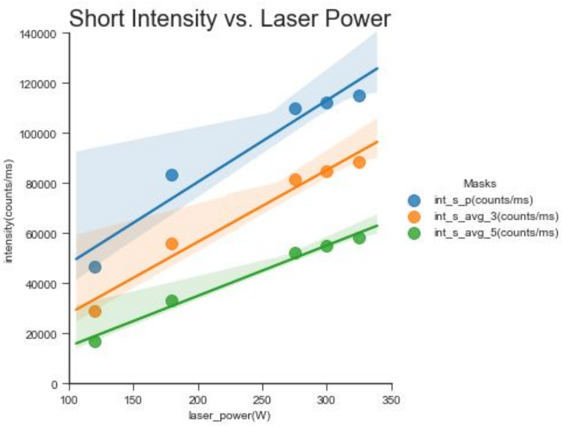
Print Parameters vs. Melt Pool Metrics

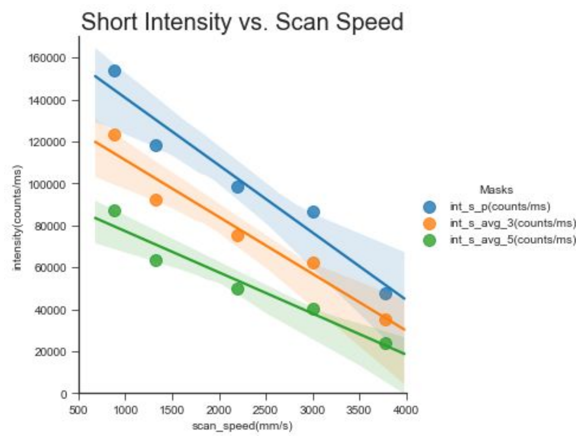
The Long Intensity Averaged over 5x5 pixels correlates fairly well with LED (0.72), while the rest of the intensities and energy densities have a correlation value between (0.59-0.70).

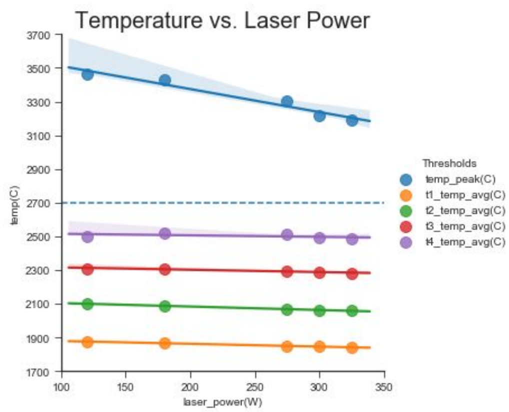
When Laser Power is increased, Intensity increases, but when Scan Speed is increased, Intensity decreased. This trend follows conventional wisdom that the more energy output by the laser, the brighter the melt pool while the faster the laser is scanned, the more mass the laser has to melt and therefore the reflected light/energy is less.

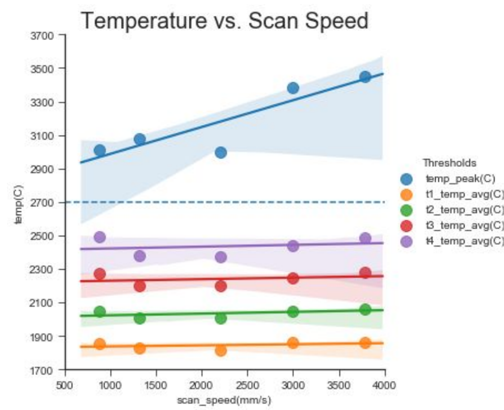
When Laser Power is increased, the Peak Temperature decreases, while when the Scan Speed is increased, the Peak Temperature increases. At first, more Laser Power may seem to predict a higher Peak Temperature, but there is a theory that the additional energy is actually buried further down inside the part, not laterally on the surface of the powder bed, which then can't be detected by the sensor. The increase in Peak Temperature due to an increase in Scan Speed hasn't been explained yet. Further testing must be down to try to explain this phenomenon.





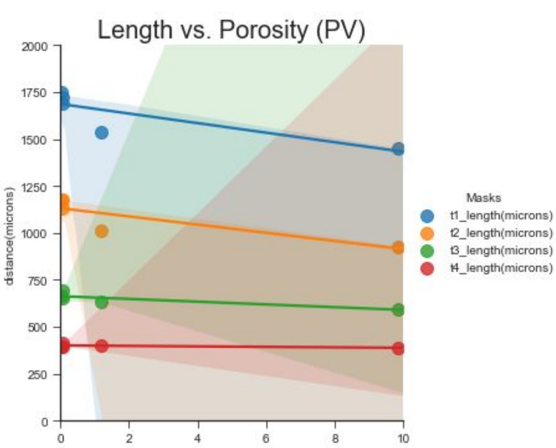


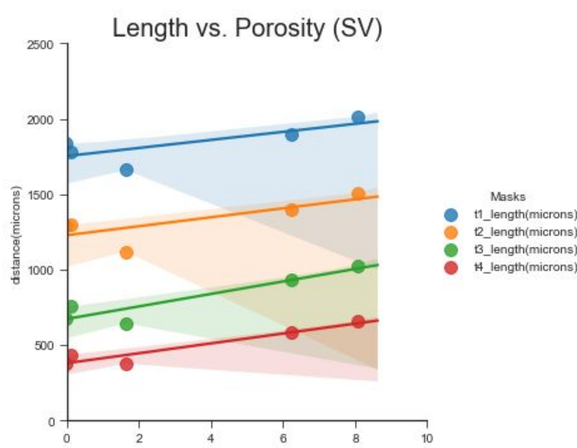




Melt Pool Metrics vs. Material Properties

According to the Correlation Plot above, all the intensity columns inversely correlate roughly the same with Porosity, between -0.61 and -0.63 (shown in the bottom right of the plot). The Length of the Melt Pool when varying Laser Power has a real hard time correlating with Porosity, while the Length correlates much better with Porosity when varying the Scan Speed. This can be further tested by expanding the Print Parameter matrices for a future experiment.





Summary

Predicting Porosity with Print Parameters

The most highly correlated Parameters are GED and VED, with a correlation value of (-0.79).

Predicting Porosity with Melt Pool Metrics

Long Peak Intensity Pixel value [int_I_p(counts/ms)] is the best indicator of Porosity (-0.63).

Predicting Melt Pool Metrics with Print Parameters

Long Intensity Averaged by 5 [int_I_avg_5(counts/ms)] correlates fairly well with LED (0.72).

Conclusion

By combining the assumptions, plots, and findings, the following table of variables should be expanded upon. Increase the size of the matrix of Print Parameters and analyze the correlation with Intensities and Porosity.

Print Parameters	**Melt Pool Metrics**	**Material Properties**
Laser Power	Intensities	Porosity
Scan Speed	-	-
LED, GED, VED	-	-