Data Aggregation Tool (VBA)

UC Davis: Biochemical Engineer
GA Tech: Analytics

About - Me

- Education
  - Biochemical Engineering (2015)
  - Masters in Analytics (2020)

- Work Experience

Virology Research (1.5 years)

Process Scientist I (2.75 years)

Process Engineer II (1.0 years)

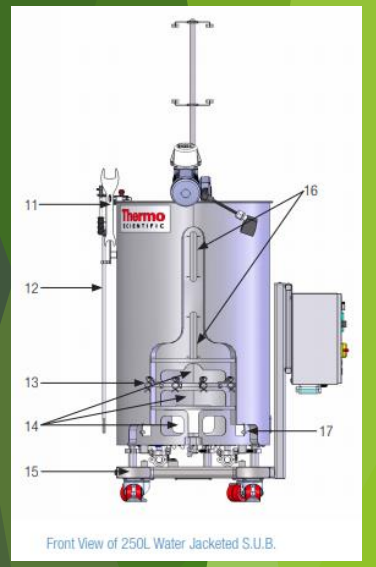
Consultant (2 mo)

About – Software/Hardware & Programming Experience

- Software

- Programming

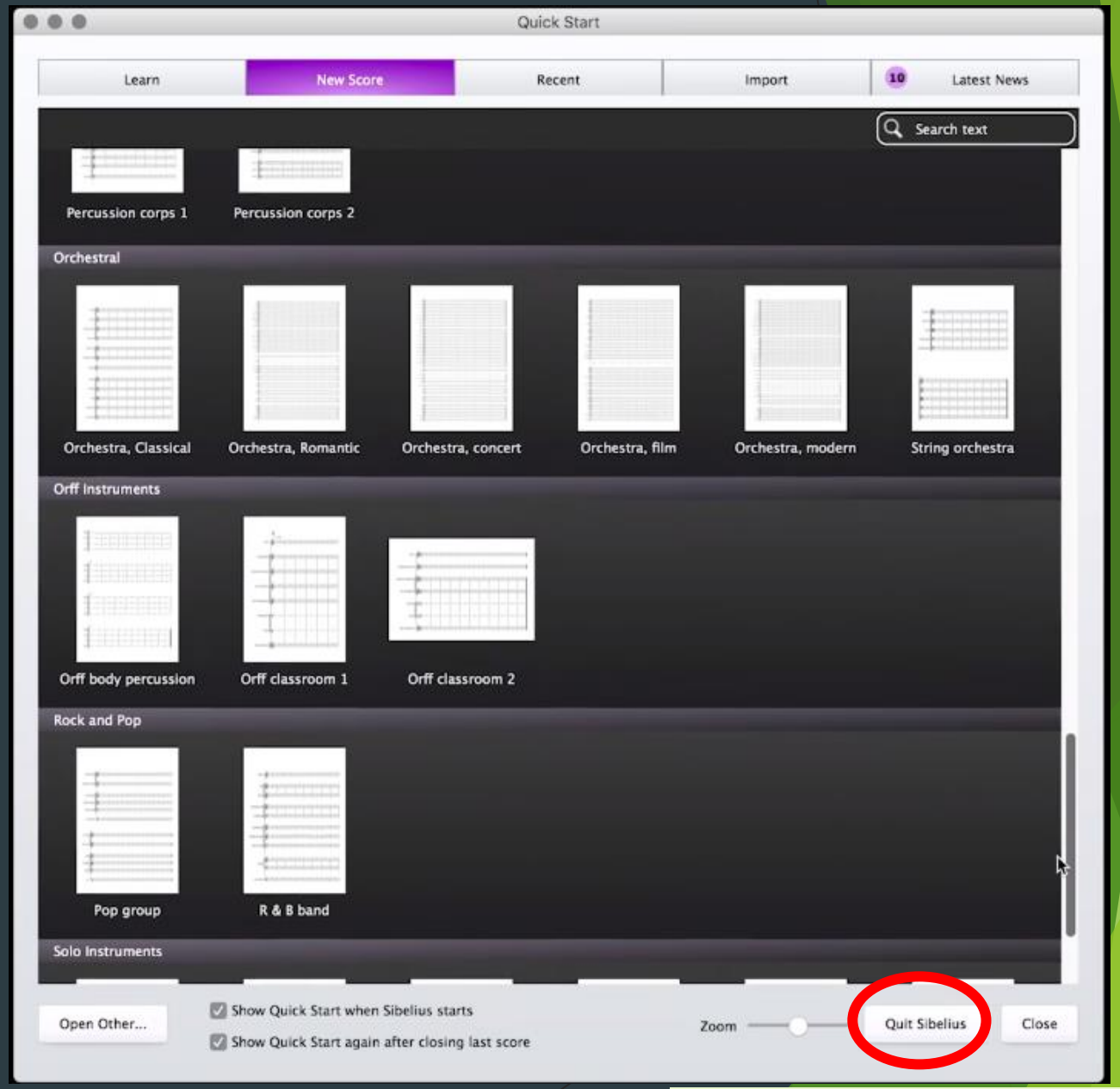
- Equipment (5-layer flasks, Ambr15, Waves, Stainless Steel reactors, disposable reactors)



Initial Problem

- "What is 'Pi'?"
  - A Data Historian system by OSISoft which can track multiple tags from bioreactors simultaneously
- Problem?
  - People knew of Pi, but many could not access the info efficiently
  - Many found it tedious to access the information they desired
- Solution?
  - Design an efficient program

Example of how NOT to design your UI (A music software)

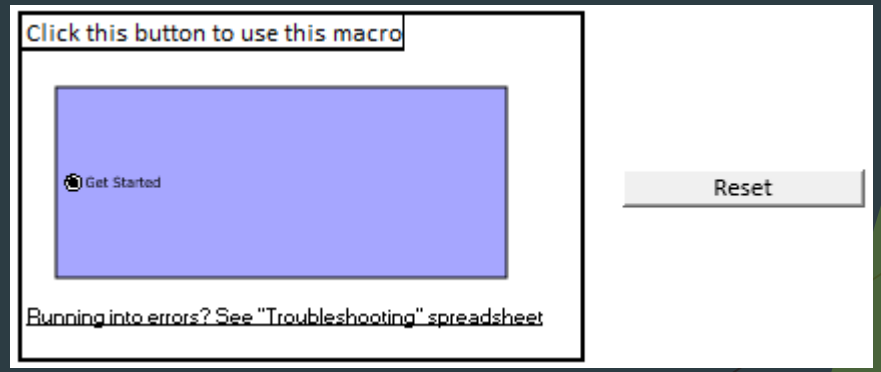


Quit?! Where is confirm??

Data Querying Program Goals

- Project goals:
  - Obtain Pi Data efficiently
  - Eliminate operator need-to-know of OSIsoft PI Data Historian or underlying VBA
- Requirements:
  - Macro should be easy to utilize
  - Data output should be easily ported over to Spotfire for visual analysis
  - Code must be written such that facility fit upgrades (such as addition of a new reactors or new Pi tags) can be updated without programming knowledge

- Solution:
  - I chose to program in VBA
    - "Build your program in a language that is optimal to the end user"
    - Everyone knows how to use Excel
    - Userforms can be designed to lead customer interaction through the software
    - Everyone knows how to click buttons!



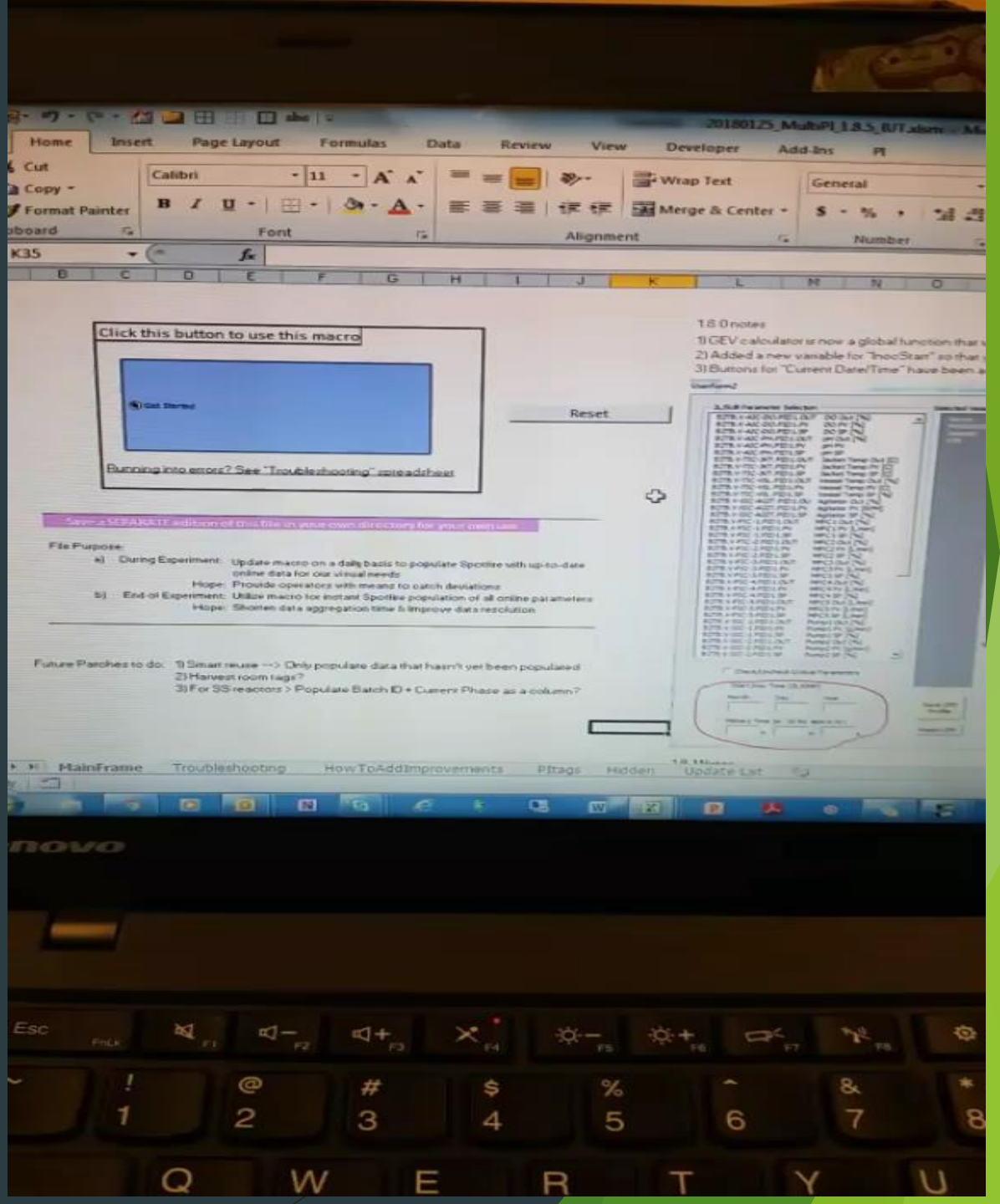
The result - MultiPi

- A program that was easy to use
- So easy in fact…
  - people still don't know what Pi is at BI

- Program Usage (In a nutshell):
  - 1) Select bioreactors
  - 2) Define bioreactors
  - 3) Select Pi Tags
  - 4) Specify START / END
  - 5) Specify STEP interval
  - 6) Accept Parameters before computation starts
  - 7) Save the file output
  - 8) Open the file in Spotfire Template

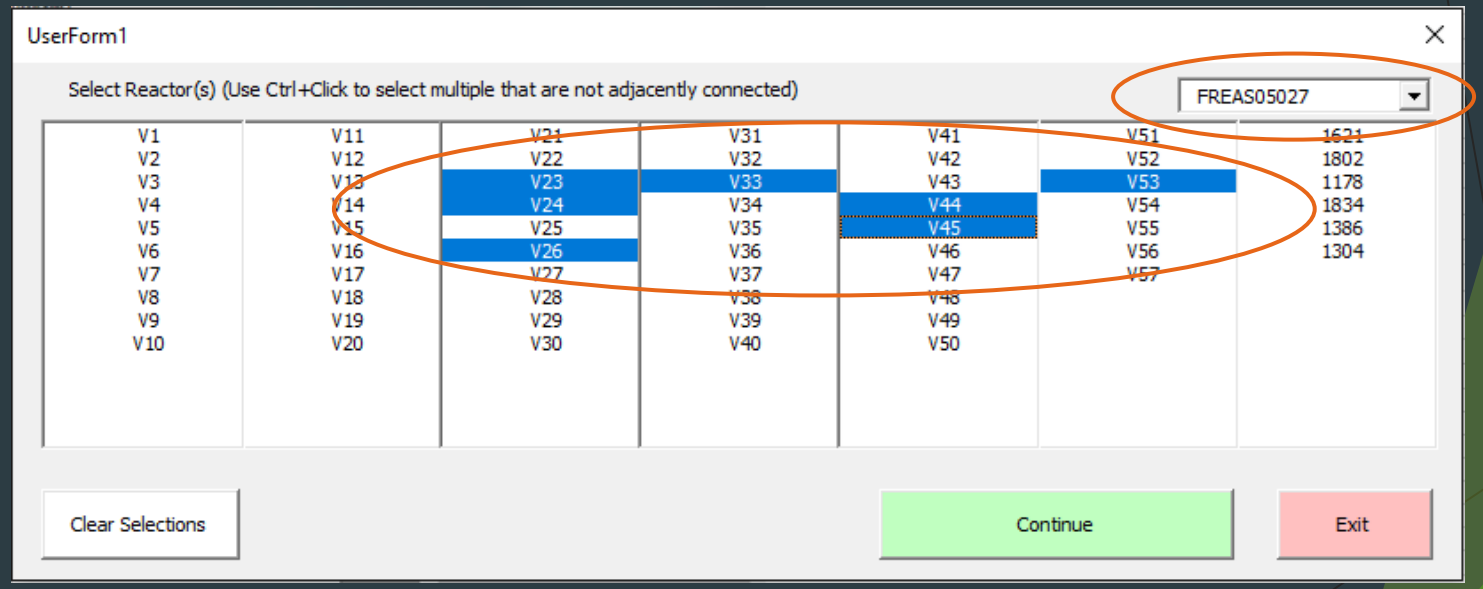
- Detailed program usage on following slides

Video



MultiPi Interface (1/3)

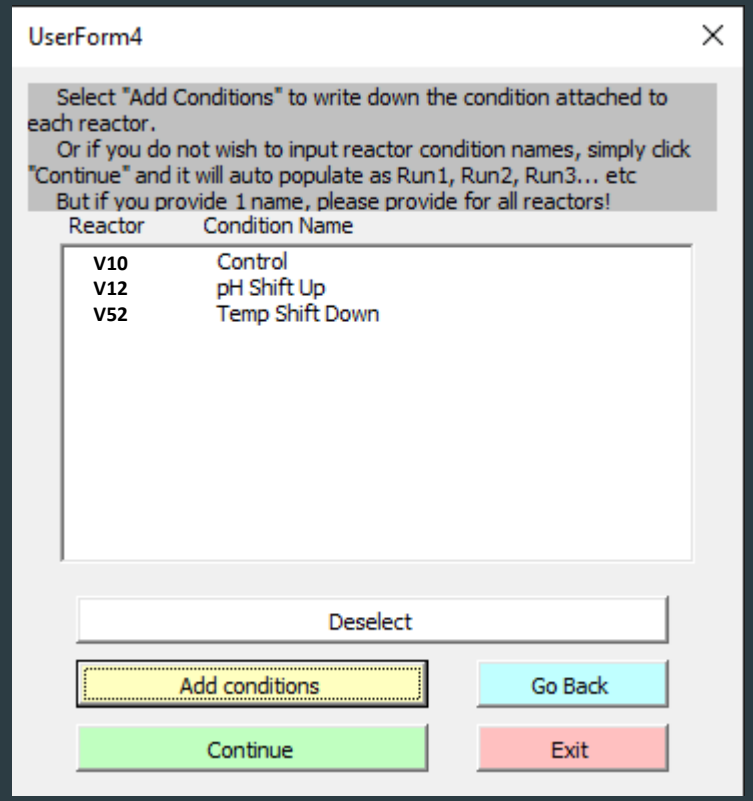
- Select Server
- Select Bioreactor(s)



- Color scheme:
  - Green = Forward
  - Red = Quit
  - Blue = Go back 1 step (Not shown for 1st step!)

MultiPi Interface (2/3)

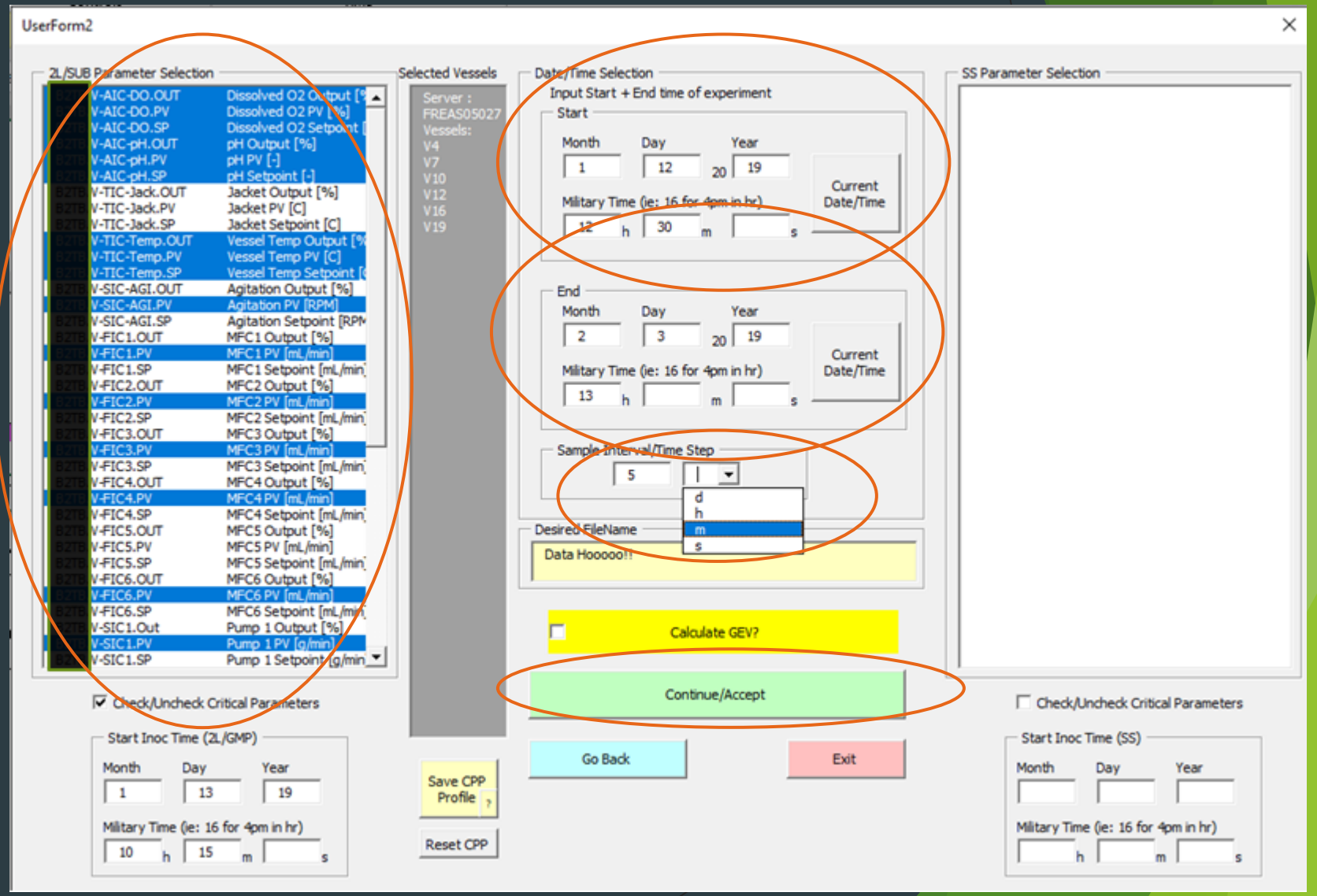
- Define your bioreactor(s) as you wish
  - Ex: Control
  - Ex2: pH Shift Up
  - Ex3: Temp Shift Down



MultiPi Interface (3/3)

- Select Pi Tag(s)
- Input Start
- Input End
- Input Step-size
- (Everything else is optional!)

- Note:
  - Grey box details bioreactors
  - Bottom-left for “Runtime”



MultiPi Output

- Reactor by Reactor output
- Colors utilized to assist in human parsing of bioreactor to bioreactor data
- Digestible by Spotfire (data tags available on row 2) or JMP (Simply delete the data tag row on row 2)

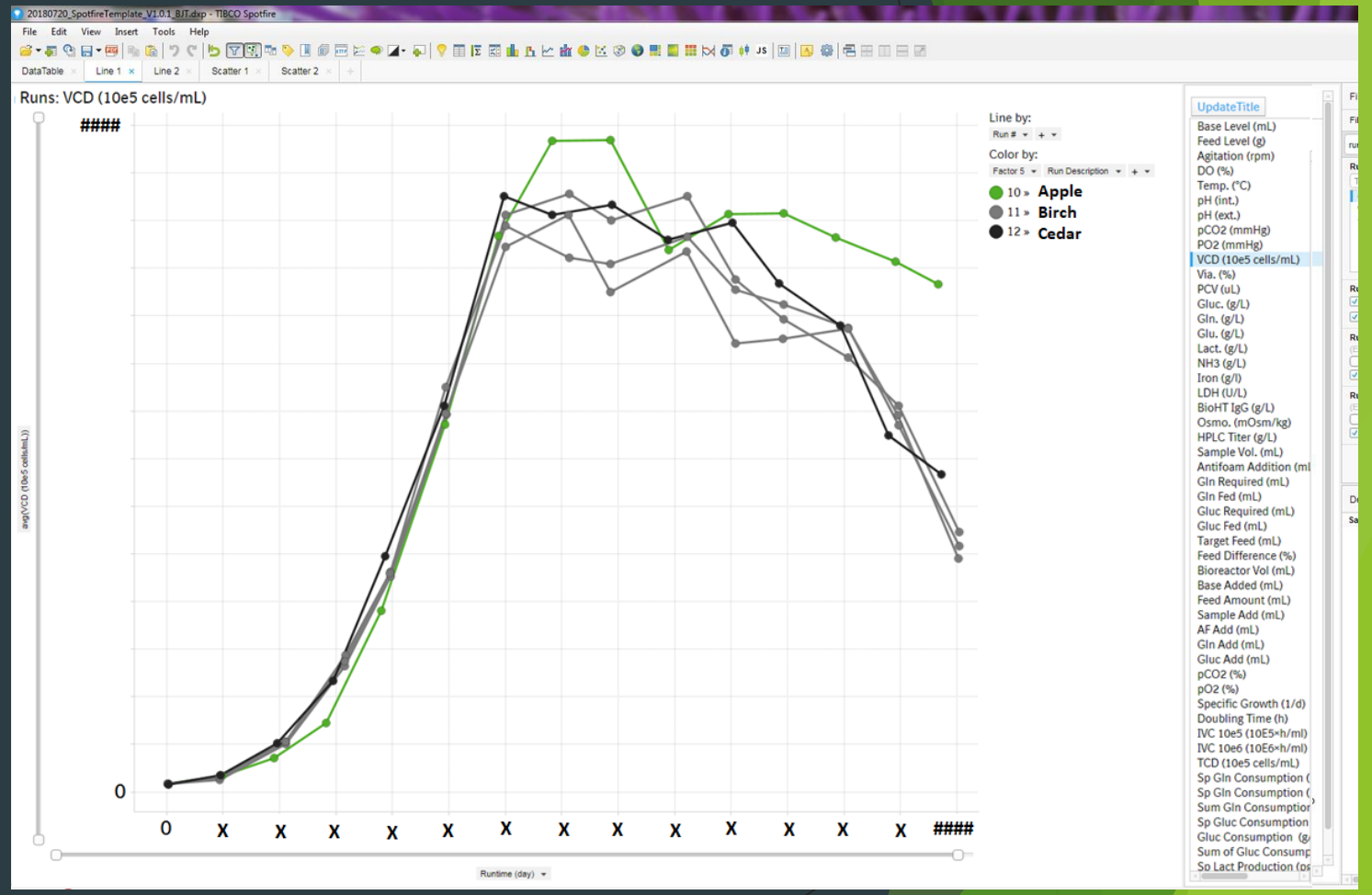
- *Data shown are randomized numbers

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	Vessel	Condition	Date/Time	Dissolved	Dissolved	Dissolved	pH Output	pH PV	pH SP	Vessel Tei	Vessel Tei	Vessel Tei	Agitator P	MFC1 PV [	MFC2 PV [	MFC3 PV [	MFC4 PV [	Pump1 PV	Pump2 PV	Scale 1 [g]	Runtime (	Run Descriptic
2	String	String	DateTime	Real	Real	Real	Real	Real	Real	Real	Real	Real	Real	Real	Real	Real	Real	Real	Real	Real	Real	String
3	V25	170955	12/05/2017 10:00:00	304	255	995	243	993	770	851	43	811	889	237	463	539	780	71	575	558	0	-+-
4	V25	170955	12/05/2017 10:05:00	171	302	448	450	986	571	395	422	786	538	911	769	697	880	4	726	509	0.00347	-+-
5	V25	170955	12/05/2017 10:10:00	926	505	167	46	609	434	908	432	936	450	727	834	407	637	416	100	439	0.00694	-+-
6	V25	170955	12/05/2017 10:15:00	682	622	38	643	793	808	548	591	516	534	983	411	555	810	603	558	252	0.01042	-+-
7	V25	170955	12/05/2017 10:20:00	422	412	568	255	521	49	606	99	536	666	248	849	657	740	860	758	251	0.01389	-+-
8	V25	170955	12/05/2017 10:25:00	845	908	523	281	712	364	777	931	345	796	973	303	884	829	85	910	953	0.01736	-+-
9	V25	170955	12/05/2017 10:30:00	348	866	247	696	921	328	969	186	670	838	617	827	980	576	495	842	104	0.02083	-+-
4900	V32	170963	12/05/2017 10:00:00	77	549	437	416	627	342	516	823	598	832	832	959	332	435	289	223	898	0	Control 2
4901	V32	170963	12/05/2017 10:05:00	898	235	645	66	290	857	73	681	33	594	615	840	239	742	851	751	826	0.00347	Control 2
4902	V32	170963	12/05/2017 10:10:00	69	370	658	540	813	704	5	306	474	98	989	836	184	353	864	598	271	0.00694	Control 2
4903	V32	170963	12/05/2017 10:15:00	926	595	807	24	835	515	705	412	70	900	467	675	798	477	408	191	191	0.01042	Control 2
4904	V32	170963	12/05/2017 10:20:00	395	183	702	569	898	697	442	302	126	34	652	46	460	829	152	960	124	0.01389	Control 2
4905	V32	170963	12/05/2017 10:25:00	434	876	8	412	226	367	128	282	931	733	142	148	812	622	226	778	873	0.01736	Control 2
9797	V38	170959	12/05/2017 10:00:00	37	288	693	372	988	588	498	59	613	185	165	553	614	708	653	858	234	0	Control 2
9798	V38	170959	12/05/2017 10:05:00	929	414	326	775	648	561	224	136	878	240	583	92	907	491	875	38	277	0.00347	Control 2
9799	V38	170959	12/05/2017 10:10:00	39	891	422	21	690	643	967	967	76	772	241	491	506	882	278	531	121	0.00694	Control 2
9800	V38	170959	12/05/2017 10:15:00	136	261	578	160	523	143	167	123	851	420	521	236	374	48	268	43	370	0.01042	Control 2
9801	V38	170959	12/05/2017 10:20:00	356	345	6	95	611	741	910	730	875	735	766	895	844	613	234	766	562	0.01389	Control 2
9802	V38	170959	12/05/2017 10:25:00	832	308	756	347	488	6	302	229	799	993	875	48	73	648	828	500	879	0.01736	Control 2
9803	V38	170959	12/05/2017 10:30:00	900	176	744	362	941	189	798	920	138	192	547	41	413	336	568	363	632	0.02083	Control 2
9804	V38	170959	12/05/2017 10:35:00	589	992	21	49	135	993	513	310	971	607	960	168	627	692	413	634	872	0.02431	Control 2
14694	V48	170957	12/05/2017 10:00:00	196	741	328	111	34	537	539	727	456	643	914	898	578	168	591	218	251	0	WCB / MCB 2
14695	V48	170957	12/05/2017 10:05:00	761	730	631	191	211	653	891	851	57	878	156	61	6	953	200	890	146	0.00347	WCB / MCB 2
14696	V48	170957	12/05/2017 10:10:00	872	182	624	573	480	360	366	1	392	160	752	822	837	90	940	174	70	0.00694	WCB / MCB 2
14697	V48	170957	12/05/2017 10:15:00	135	116	360	737	737	549	99	932	13	863	431	374	780	897	856	362	565	0.01042	WCB / MCB 2
14698	V48	170957	12/05/2017 10:20:00	442	656	978	564	74	422	171	160	115	805	30	109	127	902	683	74	562	0.01389	WCB / MCB 2
14699	V48	170957	12/05/2017 10:25:00	230	8	804	736	488	944	368	955	538	330	400	419	22	58	520	844	951	0.01736	WCB / MCB 2
14700	V48	170957	12/05/2017 10:30:00	626	391	407	559	987	349	809	4	184	302	2	168	505	23	159	665	675	0.02083	WCB / MCB 2
14701	V48	170957	12/05/2017 10:35:00	384	144	445	88	853	661	863	922	184	234	72	522	909	668	464	519	23	0.02431	WCB / MCB 2
19591	V49	170956	12/05/2017 10:00:00	729	528	666	287	103	248	888	116	547	919	122	350	716	826	966	507	437	0	WCB / MCB 2
19592	V49	170956	12/05/2017 10:05:00	998	61	399	5	589	344	739	896	248	468	697	337	787	287	415	28	870	0.00347	WCB / MCB 2
19593	V49	170956	12/05/2017 10:10:00	950	954	299	675	181	187	495	780	502	326	95	41	275	243	822	800	241	0.00694	WCB / MCB 2
19594	V49	170956	12/05/2017 10:15:00	854	908	976	867	892	755	57	773	580	5	574	556	834	173	177	893	49	0.01042	WCB / MCB 2
19595	V49	170956	12/05/2017 10:20:00	732	3	297	393	279	229	492	919	865	761	550	633	680	2.87E+02	26	273	744	0.01389	WCB / MCB 2
19596	V49	170956	12/05/2017 10:25:00	909	95	43	756	120	894	834	544	192	797	938	702	859	351	916	962	745	0.01736	WCB / MCB 2
19597	V49	170956	12/05/2017 10:30:00	739	88	463	113	68	980	35	139	501	407	164	278	988	499	460	790	305	0.02083	WCB / MCB 2
19598	V49	170956	12/05/2017 10:35:00	242	748	370	147	485	775	237	313	45	828	772	324	632	560	619	350	441	0.02431	WCB / MCB 2
19599	V49	170956	12/05/2017 10:40:00	177	206	469	450	602	906	389	815	406	335	98	218	928	222	388	134	56	0.02778	WCB / MCB 2
24488	V21	170861	11/03/2017 10:00:00	284	964	284	887	761	290	912	875	812	801	519	599	267	409	98	982	958	0	--+-
24489	V21	170861	11/03/2017 10:05:00	378	546	136	2	728	552	958	875	528	417	401	730	666	413	979	48	136	0.00347	--+-
24490	V21	170861	11/03/2017 10:10:00	342	995	167	254	723	786	807	767	119	988	642	570	582	773	543	319	47	0.00694	--+-
24491	V21	170861	11/03/2017 10:15:00	478	814	380	594	929	591	612	256	210	695	786	671	574	944	273	130	473	0.01042	--+-
24492	V21	170861	11/03/2017 10:20:00	37	223	19	316	85	232	474	965	661	913	397	471	327	642	611	399	372	0.01389	--+-
24493	V21	170861	11/03/2017 10:25:00	736	816	616	877	348	206	37	299	821	848	765	705	455	895	532	421	646	0.01736	--+-
24494	V21	170861	11/03/2017 10:30:00	185	342	881	540	91	47	695	601	859	996	401	461	547	566	514	106	69	0.02083	--+-
24495	V21	170861	11/03/2017 10:35:00	48	237	763	346	799	828	973	445	864	977	935	389	364	893	119	157	628	0.02431	--+-
29384	V21	170861	11/20/2017 09:00:00	323	805	725	743	734	721	826	1000	334	606	773	391	489	262	773	391	827	16.95833	--+-

29384 rows * 17 tags = **499,528 data points took ~6min**

Spotfire Integration

- Template created in Spotfire
  - Reads all column names in uploaded file into Listbox registry
- Automatically plots selected entry from ListBox
  - Example shown is for “Runfile” project plotting VCD



MultiPi Version Control

- Updates would be made to:
  - Add new features
  - Improve upon old features
  - Fix bugs
- All changes were documented
- Code is highly commented

```
'==============================Massive for-loop for calculation==============
Counter = 0 'Counter variable for remembering which row to put the data into
ColumnCounter = 0 'Counter variable for remembering which column we are in
Resetcounter = 0 'Counter variable to count how many times through the loop the code has gone (aka how many tags there were)
VesselCounter = 0 'Counter variable to count how many times vessels have been selected

'Populate the table with some column headers ahead of time
Range("A1").Value = "Vessel"
Range("B1").Value = "Condition"
Range("C1").Value = "Date/Time"

If CC2L_Selected = True Then
    For i = 0 To 57 '0 to # 2L vessels
        If Sheets("Hidden").Range("A1").Offset(i, 0) = 1 Then 'If a vessel has been selected, do the following
            Range(Range("A2").Offset(Counter, 0), Range("A2").Offset(Counter + numpts, 0)).FormulaArray = "V" + CStr(i + 1) 'Populate the first column with Vessel ID
            Range(Range("B2").Offset(Counter, 0), Range("B2").Offset(Counter + numpts, 0)).FormulaArray = Sheets("Hidden").Range("D1").Offset(i, 0).Value 'Populate condition for that reactor beside it
            For j = 0 To ListBox21.ListCount - 1 'Index for tags
                If ListBox21.Selected(j) = True Then 'If a PI tag has been selected, do the following
                    'Obsolete line' Sheets("Hidden").Range("B1").Offset(j, 0).Value = 1 'Save the tag for future use
                    Range("D1").Offset(0, ColumnCounter).Value = Sheets("PItags").Range("C2").Offset(j, 0).Value 'Grab the descriptor name for the respective tag and put it as the column header
                    Tag = Sheets("PItags").Range("B2").Offset(j, 0).Value 'Save the name just so its easier on the eyes for the next line

                    'First part with the large computation is just calculating a % of what has been completed over the total amount of things that need to be completed
                    Application.StatusBar = "Part 1/4 | Progress: " + CStr(Round(100 * ((VesselCounter) * Total2LTags + ColumnCounter + 1) / (Total2LReactors * Total2LTags), 1)) + "% | Currently popula

                        If Sheets("Hidden").Range("G6").Value = 0 Then 'FREAS05027
                            TagName = Left(Tag, 6) + CStr(Sheets("Hidden").Range("A1").Offset(i, 0).Row) + Mid(Tag, 7, Len(Tag)) 'Complete tag name
                        ElseIf Sheets("Hidden").Range("G6").Value = 1 Then 'FREAS05019
                            TagName = Left(Tag, 1) + CStr("3") + Mid(Tag, 3, 4) + CStr(Sheets("Hidden").Range("A1").Offset(i, 0).Row) + Mid(Tag, 7, Len(Tag))
                        End If
        '---Result is a 2D matrix with columns for (Time,Value) of the specified parameters... but we only want "Value" and to not duplicate "Time" over and over
                    If ColumnCounter = 0 Then 'If this is the first time around the loop
                        'We want to populate time along with the results
                        result = Application.Run("Pisampdat", CStr(TagName), CDate(StartPoint), CDate(EndPoint), CStr(TimeStep) + CStr(TimeStepUnit), 1, CStr(Server))
                        'The "1" index in the Application.run command tells it to calculate a time column as well
                        Range(Range("C2").Offset(Counter, 0), Range("C2").Offset(Counter + numpts, 1)).FormulaArray = result 'Grab both columns of "Result" and populate the next 2 columns
                    Else
                        result = Application.Run("Pisampdat", CStr(TagName), CDate(StartPoint), CDate(EndPoint), CStr(TimeStep) + CStr(TimeStepUnit), 0, CStr(Server))
                        'Because we don't need a time column to be repeated... we now use a "0" index
                        Range(Range("B2").Offset(Counter, ColumnCounter + 2), Range("B2").Offset(Counter + numpts, ColumnCounter + 2)).FormulaArray = result
                    End If
                    Resetcounter = Resetcounter + 1
                    ColumnCounter = ColumnCounter + 1 'Increase the index
                Else 'If the tag is not selected
                    Sheets("Hidden").Range("B1").Offset(j, 0).Value = 0
                End If
            Next j
            Range(Range("B2").Offset(Counter, 0), Range("B2").Offset(Counter, Resetcounter)).ClearContents
            Counter = Counter + numpts + 1
            VesselCounter = VesselCounter + 1
        End If
        ColumnCounter = 0 'Reset the index for ColumnCounter as we're starting a new vessel
    Next i
End If
```

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		Project Start: 06/27/16				Created by BJT for PS Cell Culture									
2		Version 1.0 release: 06/29/16				Designed to be user friendly for those who don't know PI, customizable to fully utilize the capabilities of PI, and functional enough to plot in Spotfire									
3															

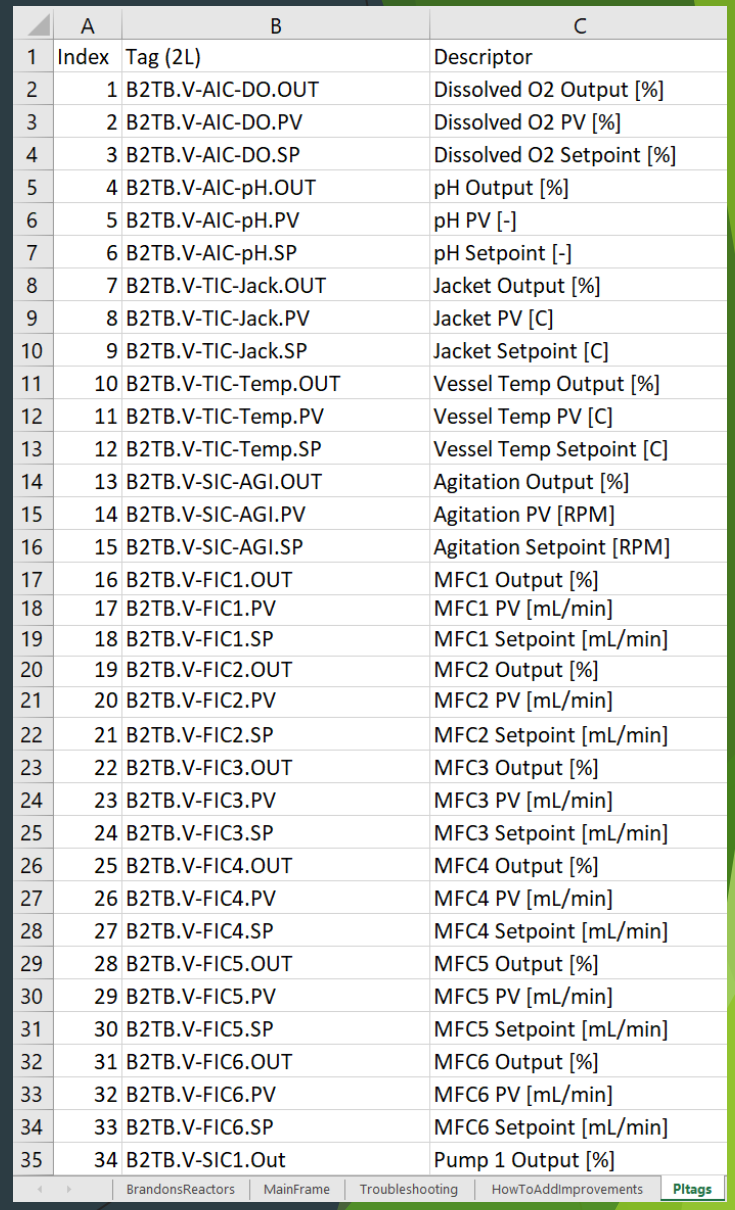
Version #	Issue resolved	Date + Initial
1.1.1	Added COLOR and enabled feature of not having a Desired File Name specified... will not break program anymore	06/29/16 BJT
1.2	Allowed >20 reactor input (previous would crash color palette) and introduced saved parameters into "Hidden" sheet as well as a "Reset" button to reset parameters	06/29/16 BJT
1.3	Added "Clear" to listbox and combobox entries to prevent duplication of items populating (bug fix). Also introduced text filter to prevent erroring when utilizing special characters like ?*{}<>: in a file name	07/05/16 BJT
1.3	Formatted the "Reset" button to be smaller and separated from the "Get Started" button as well as giving context for the "Reset" button to clarify its purpose as a convenince button and not an "update file" button.	07/05/16 BJT
1.3.1	Made this file "Read-Only" so that the "Reset" button is no longer necessary	07/11/16 BJT
1.4	Resolved the bug where 1 time-point would always be missing	07/18/16 BJT
	Removed "Read-only" format as that then makes every subsequent save-as version of this file a "read-only" file as well	07/20/16 BJT
	Added color to the various buttons in the userforms	07/20/16 BJT
		07/20/16 BJT
1.4.1	Removed from the tag list a tag called "GEV Calculator" because this was my alternative method to calculating GEV which was not	08/04/16 BJT
1.4.1	Added "type row" headers that Spotfire requires in order to read the data correctly	08/04/16 BJT
1.5		09/09/16 BJT
1.5.1		09/14/16 BJT
1.6		09/15/16 BJT
1.6.1		09/29/16 BJT
1.7.0		02/02/17 BJT
1.7.0		05/30/17 BJT
1.8.0		07/19/17 BJT
1.8.1		7/27/2017 BJT
1.8.1		7/28/2017 BJT
1.8.1		8/1/2017 BJT
1.8.2		8/17/2017 BJT
1.8.3		JT 10/12/2017
1.8.4		12/12/2017 BJT
1.8.5		1/25/2018 BJT

MultiPi Goal Check

- Project goals:
  - Obtain Pi Data efficiently
  - Eliminate operator need-to-know of OSIsoft PI Data Historian or underlying VBA
  - Macro should be easy to utilize
  - Data output should be easily ported over to Spotfire for visual analysis
  - Code must be written such that facility fit upgrades (such as addition of a new reactors or new Pi tags) can be updated without programming knowledge

MultiPi Goal Check

- "**can be updated without programming knowledge**"
- If PI servers updated to a new address, tags got renamed, or new tags were added
  - All of the coding referenced excel sheets for non-coder modification
- If expansion happens and more than 57 reactors exist?
  - … Honestly I forgot to put this outside of the code.
  - I'm still waiting for the day they call me about this



	A	B	C
1	Index	Tag (2L)	Descriptor
2	1	B2TB.V-AIC-DO.OUT	Dissolved O2 Output [%]
3	2	B2TB.V-AIC-DO.PV	Dissolved O2 PV [%]
4	3	B2TB.V-AIC-DO.SP	Dissolved O2 Setpoint [%]
5	4	B2TB.V-AIC-pH.OUT	pH Output [%]
6	5	B2TB.V-AIC-pH.PV	pH PV [-]
7	6	B2TB.V-AIC-pH.SP	pH Setpoint [-]
8	7	B2TB.V-TIC-Jack.OUT	Jacket Output [%]
9	8	B2TB.V-TIC-Jack.PV	Jacket PV [C]
10	9	B2TB.V-TIC-Jack.SP	Jacket Setpoint [C]
11	10	B2TB.V-TIC-Temp.OUT	Vessel Temp Output [%]
12	11	B2TB.V-TIC-Temp.PV	Vessel Temp PV [C]
13	12	B2TB.V-TIC-Temp.SP	Vessel Temp Setpoint [C]
14	13	B2TB.V-SIC-AGI.OUT	Agitation Output [%]
15	14	B2TB.V-SIC-AGI.PV	Agitation PV [RPM]
16	15	B2TB.V-SIC-AGI.SP	Agitation Setpoint [RPM]
17	16	B2TB.V-FIC1.OUT	MFC1 Output [%]
18	17	B2TB.V-FIC1.PV	MFC1 PV [mL/min]
19	18	B2TB.V-FIC1.SP	MFC1 Setpoint [mL/min]
20	19	B2TB.V-FIC2.OUT	MFC2 Output [%]
21	20	B2TB.V-FIC2.PV	MFC2 PV [mL/min]
22	21	B2TB.V-FIC2.SP	MFC2 Setpoint [mL/min]
23	22	B2TB.V-FIC3.OUT	MFC3 Output [%]
24	23	B2TB.V-FIC3.PV	MFC3 PV [mL/min]
25	24	B2TB.V-FIC3.SP	MFC3 Setpoint [mL/min]
26	25	B2TB.V-FIC4.OUT	MFC4 Output [%]
27	26	B2TB.V-FIC4.PV	MFC4 PV [mL/min]
28	27	B2TB.V-FIC4.SP	MFC4 Setpoint [mL/min]
29	28	B2TB.V-FIC5.OUT	MFC5 Output [%]
30	29	B2TB.V-FIC5.PV	MFC5 PV [mL/min]
31	30	B2TB.V-FIC5.SP	MFC5 Setpoint [mL/min]
32	31	B2TB.V-FIC6.OUT	MFC6 Output [%]
33	32	B2TB.V-FIC6.PV	MFC6 PV [mL/min]
34	33	B2TB.V-FIC6.SP	MFC6 Setpoint [mL/min]
35	34	B2TB.V-SIC1.Out	Pump 1 Output [%]

BrandonsReactors | MainFrame | Troubleshooting | HowToAddImprovements | **PItags**

MultiPi in Summary

- Program became tool to query data most people never knew how to access in the first place
  - Both non-GMP and GMP registries were built-in
- People in other departments are utilizing my program
  - Process Tech Transfer Specialist getting data to a client
  - Clinical Ops comparing batch to batch performance
- MultiPi grabs 20 reactors for 12 tags across 14 days in ~15minutes
  - >500% efficiency increase to how most people were grabbing the data into excel
- Great care was taken for the customer:
  - Each GUI was drawn out and drafted before finalizing in VBA
  - Many features were built-in for end-user comfort (Ex: Pre-selection of Pi-tags in a single checkbox click)
  - Updates could be done on the excelbook ▫ No VBA required!