

SELLING POWER SYSTEM FLEXIBILITY: ANCILLARY SERVICE AND REAL-TIME ENERGY MARKET CHALLENGES FOR STORAGE

Brendan Kirby
Consultant
Palm City, Florida 34990
kirbybj@ieee.org

ABSTRACT

Power system operators obtain the flexibility required to reliably balance aggregate generation and load through ancillary service and five-minute energy markets. Market prices are based on the marginal opportunity costs of the generators. This market design works well for generators but inherently fails for storage and demand response, denying these new technologies a fair opportunity to compete and denying the power system access to potentially lower cost reliability resources. Market design or regulatory changes may be required for storage and demand response to be viable ancillary service providers.

INTRODUCTION

Providing ancillary services to the power system is potentially very attractive for storage. The power to stored energy ratio is much higher than for daily, weekly, or seasonal energy arbitrage. Ancillary service prices are highest for the fastest, most accurate, but shortest duration and lowest net energy services (regulation followed by spinning reserve). Five-minute sub-hourly energy markets provide another interesting opportunity for storage because of the high within-hour price volatility and the inherently short response duration.

Independent System Operators (ISOs) and Regional Transmission Organizations (RTOs), which serve two thirds of the national load, operate hourly markets for ancillary services and sub-hourly energy markets that clear every five minutes. These markets provide the system operator with the flexibility required to keep the power system balanced and operating reliably and they pay the resources that are able to respond. Ancillary service and sub-hourly energy markets work very well, both for the system operators that require response and for the generators that provide that response. Unfortunately the current market structures were designed around the characteristics of generators and they inherently fail for energy storage and demand response technologies. This is unfortunate for the storage owners because it denies them a potential income source. It is also unfortunate for system operators and loads because it denies the power system access to flexible, fast, and accurate reliability resources that potentially could reduce system costs. Changes to ancillary service market rules may be appropriate.

This paper first discusses each of the ancillary services and notes which are currently paid for through markets. It then discusses the ancillary service market structure, how prices are based on generation opportunity costs, and why this does not work for storage. Alternative market structures are discussed which could give the power system access to the lowest cost ancillary service providers and appropriately compensate storage. Finally, the Appendix provides a table of annual average ancillary service prices from six ISO/RTO markets for the past twelve years.

ANCILLARY SERVICES AND FIVE-MINUTE ENERGY MARKETS: HOW POWER SYSTEMS PAY FOR FLEXIBILITY

Power systems must balance generation and load instantaneously and continuously. They employ a series of ancillary services to accomplish this balancing. They also use sub-hourly generation dispatch to obtain additional balancing. Two thirds of the U.S. load is located in ISO/RTO areas where system operators obtain this required balancing flexibility through ancillary service and five-minute energy markets.

Paid and Unpaid Flexibility

Some flexibility is explicitly paid for. Hourly markets exist for four ancillary services and sub-hourly energy. Payment is being considered for two additional ancillary services. Without explicit mechanisms to monitor and pay for specific services it is unlikely that commercial entities (generators, demand response and storage) will provide them.

Inertia. Inertia is the inherent response a synchronous generator (or motor) provides to the power system when there is a major disturbance. Response is based on the rotational mass of the generator and it slows the power system frequency decline, giving time for governors and AGC to respond. It is not a paid ancillary service and may not become one. Appropriately designed inverter connected storage could provide synthetic inertia. Some wind generation manufacturers already offer synthetic inertia as an option.

Governor/Frequency Response. Governor response or frequency response is the autonomous controlled response provided through generator governor action that increases

output as frequency declines and decreases output when frequency rises (blue curve in Figure 1). It is not yet a paid ancillary service but it may become one, largely because generators have no incentives to provide governor response without compensation. Storage can be designed to provide governor response.

Spinning Reserve. Spinning reserve is generation, storage, or responsive load capacity that is available to respond to power system contingencies (red curve in Figure 1). The resource must begin responding immediately and be fully responsive within ten minutes. Response can be an hour or longer but in practice system operators prefer to release spinning reserves quickly so that the reserves are available to respond to a subsequent event. ISOs run hourly markets for spinning reserve and selected resources are paid for standing ready weather response is required that hour or not.

Non-spinning Reserve. Non-spinning reserve is a paid ancillary service that is very similar to spinning reserve with the same ten minute response requirement. Response need not begin immediately and non-spinning reserve can be supplied by fast-start generation.

Supplemental Operating Reserve. Some ISO/RTOs also procure supplemental operating reserve which is similar to non-spinning reserve but fully responsive in 30 minutes (turquoise curve in Figure 1).

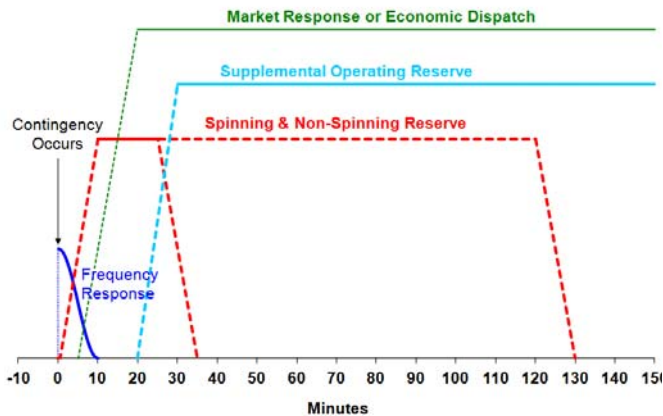


FIGURE 1 CONTINGENCIES ARE RESPONDED TO WITH A SERIES OF ANCILLARY SERVICES.

Regulation. ISO/RTOs procure regulation through hourly markets. It is the use of on-line generation, responsive load and storage that is equipped with automatic generation control (AGC) and that can change output quickly (MW/min) to track the moment-to-moment fluctuations in aggregate customer loads and to correct for the unintended fluctuations in aggregate generation (red curve in Figure 2). Regulation helps to maintain interconnection frequency, manage differences between actual and scheduled power flows between balancing areas, and match

generation to load within the balancing area. AGC commands are typically sent about every four seconds.

Regulation performance (as opposed to regulation capacity which is already measured and paid for) is beginning to be measured and paid for. PJM currently has the most advanced regulation metrics with scores for correlation, precision, and performance. Faster and more accurate response is paid more.

Following. Following is the slower counterpart to regulation which tracks the sub-hourly trend in customer loads and to correct for the unintended fluctuations in generation (blue curve in Figure 2). Following is not yet a paid ancillary service but both the CAISO and the MISO are considering establishing markets for following. Historically five-minute energy markets have been able to provide sufficient following response at essentially no cost to the ISO/RTO so that an additional ancillary service was not needed. With increased wind and solar penetration there is a concern that the five-minute energy markets may not have sufficient depth to compensate for large wind and solar ramps and a dedicated reserve will be required.

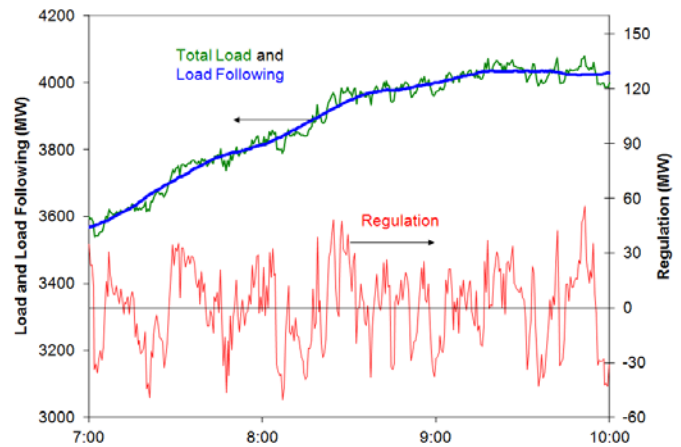


FIGURE 2 REGULATION RESPONDS TO MINUTE-TO-MINUTE RANDOM VARIABILITY WHILE FOLLOWING RESPONDS TO LONGER TRENDS.

Five Minute Energy Markets. Ancillary services are procured as capacity – the dedicated capability, measured in MW, to respond for a specified duration, usually an hour. Supply of energy is accounted for separately and is almost incidental. Sub-hourly energy markets explicitly pay for energy delivered during the specific interval – measured in MWH during each five minute period – and do not explicitly pay for capacity.

Each of these products provides the system operator with control over the production or consumption of real power so that s/he can fulfill the balancing area’s obligation to match aggregate generation and aggregate load instantaneously and

continuously. Control over production or consumption is equally effective for balancing the power system, it is the control that matters. The resources themselves must be flexible in order to offer control capability to the power system operator. Flexibility is explicitly specified in each of the ancillary services (X MW of capacity range with Y MW/min ramp rate). Flexibility is implicit with sub-hourly energy markets since it is required if a resource is to be profitable by supplying energy during high price intervals and avoiding production or even consuming energy during low or negatively priced intervals. Five minute energy prices tend to be volatile (Table 1, discussed below) providing a strong incentive, but not an explicit requirement, for flexibility.

ANCILLARY SERVICE MARKET STRUCTURE

Ancillary service markets work very well for conventional generators. They appropriately assess the opportunity cost a generator will incur if it withholds capacity from the energy market in order to supply an ancillary service. A generator with a fuel plus variable overhead and maintenance cost of \$30/MWH, for example, would forgo \$10/MW-hr¹ in energy sale profits if it provided spinning reserve rather than selling its full production if the energy market were clearing at \$40/MWH that interval. \$10/MW-hr is then the appropriate spinning reserve bid price for that generator for that market interval. If the spinning reserve market clears above \$10/MW-hr for that interval then the generator will profit by selling spinning reserve. If the spinning reserve market clears below \$10/MW-hr the generator will prefer to sell energy.

Structuring ancillary service markets based on energy market opportunity costs allows the energy and ancillary service markets to be co-optimized, maximizing generator profits while simultaneously minimizing total system cost. It also provides economic incentive for generators to offer ancillary services at actual cost, knowing that they will profit whenever the market clears above their cost. Further, the generators only need to include their actual costs (fuel and variable operations and maintenance) in their ancillary service bids; the system operator's co-optimization software takes care of calculating the energy market opportunity costs.

Real-time five-minute energy markets are similarly effective in providing system operators with flexible balancing capability. They coordinate well with day-ahead and hour-ahead energy markets giving generators additional opportunities to profit by offering flexibility to the power system.

Actual ancillary service and energy market implementations are much more complex, of course, with detailed rules to account for energy costs within ancillary services, multiple market clearing intervals (day-ahead, hour-ahead, etc.), startup and commitment cost, etc. It is, however, this central feature of basing ancillary service prices on energy

market opportunity costs which makes the market structure so successful in optimizing conventional generators but which creates a fundamental problem for some storage (and demand response) technologies.

Concerns for Storage

The current opportunity-cost based ancillary service market structure works so well precisely because the conventional generators have an alternative primary use: selling energy. Conventional generators are almost always built to primarily supply energy. They are, hopefully, also designed to provide ancillary services but that is not their primary purpose. The market participants are happy to supply ancillary services but they always have the alternative of supplying energy. This is true even for generators that may obtain a significant portion of their income from the provision of ancillary services [1]. Consequently, it is not necessary, or desirable, for ancillary service markets to consider resource capital costs in the ancillary service commitment or dispatch. Ancillary service markets are fundamentally structured to allocate available resources with alternative values in real-time. Sunk costs are deliberately, and appropriately, ignored.

While the ancillary service market structure works very well for generators it will fail for dedicated storage and demand response technologies. These technologies can have very low operating cost but significant capital cost. Even if the total cost (capital pulse operating costs) for a dedicated ancillary service storage technology is well below the marginal opportunity cost of the conventional generators the storage technology will economically fail and society will be denied the storage technology benefits.

To see why the current ancillary service market structure fails for storage let's look at a hypothetical example. Assume a California storage project capable of delivering ± 1 MW of regulation (and only regulation). To keep the example simple and illustrate the concept also assume that there are no cycling costs or losses and that the power system use of regulation is energy neutral. Assume too that this storage project is economically viable with an annual income of \$50,000/year.

How should such a project participate in the ancillary service market? It should bid a \$0/MW-hr cost for up and down regulation since it has no operating costs. The California Independent System Operator (CAISO) co-optimization software will add no opportunity cost to the operating cost since the storage project is a dedicated regulation facility and has no alternative opportunities in the energy market. The project's total bid is \$0.00/MW-hr and it will be selected for regulation every hour. On average it would have received \$7.81/MW-hr of up and down regulation in 2013 (\$4.56/MW-hr for up and \$3.25/MW-hr for down on average) for a total

¹ Note the terminology "MW-hr" is used to designate a unit of 1 MW of capacity for 1 hour and is not the same as a unit of energy (MWH)

annual income of \$68,416, well above the \$50,000 required to pay the capital cost.²

So far everything is fine. The market software correctly selects the storage project each hour and it receives appropriate compensation based on the market’s alternative of using generation. The problem arises when additional storage is built. In our example the ±1MW storage project is too small to impact the market clearing price. But CAISO only required ±332MW of regulation on average during 2013 (338MW of up and 325MW of down, on average) and a maximum of ±500MW. Installing ±300MW to ±500MW of regulating storage might be technically feasible but it would drive the market clearing price to zero and storage would not be economically viable. This happens because the ancillary service market is structured around the short-term marginal cost and does not consider capital cost.

Sub-hourly energy markets are similarly problematic for storage. Table 1 compares average hourly day-ahead and real-time (5 minute) energy prices for 2013 from three regions. There is relatively little difference between the annual average day-ahead and annual average real-time energy prices for each of the regions but there is significant within-hour price spread. In CAISO, for example, the five minute energy price averaged just \$2.33/MWH less than the day-ahead price. In contrast the highest 5 minute interval price exceeded the lowest five minute price each hour by \$34.05/MWH on average. The system operator obtains significant flexibility from the real-time energy markets and at relatively low cost: average five minute energy prices are often lower than day-ahead hourly energy prices. Generators provide the flexibility and respond to the five minute prices because it is the last chance to sell energy. Generator profits come from the energy prices rather than from direct payment for response. Still, there are costs to generators for providing this response as evidenced by the large within-hour price spread. It takes a significant price signal to motivate generator movement.

TABLE 1. 2013 AVERAGE ANNUAL \$/MWH ENERGY PRICES [2,3,4]

	Day-Ahead	5-Minute	
		Average	Within-Hour-Range
CAISO	\$43.40	\$41.07	\$34.05
NYISO	\$37.35	\$36.33	\$30.01
MISO	\$31.27	\$31.22	\$28.76

² This example uses data from the CAISO. Though CAISO had the lowest average regulation price for 2013 CAISO publishes more information on ancillary service quantities and sources which enable a more complete example though other locations would likely be more profitable.

Some storage technologies can respond faster and more accurately than generation and could respond to the five-minute fluctuations thus reducing conventional generation ramping costs. The within-hour price spread could be reduced to that required to cover storage cycling losses, perhaps 20% or about \$8/MWH maximum as opposed to the \$34/MWH average seen today. As storage collapses the within-hour price spread, however, it eliminates the payments that could pay for storage. The generators benefit through reduced response costs but storage is unsustainable with this market design.

Obviously both market structure results are bad for storage but they are also a bad result for society. A lower cost technology is blocked from the market simply because the market is structured around the opportunity costs of incumbent resources that have alternative (energy production) uses.

Some argue that the ancillary service market structure is fine and storage simply needs to include its capital cost in the market bid. This does not work. In our example the storage project has an annual capital requirement of \$50,000/MW-yr or \$5.71/MW-hr on average, well below the annual regulation market clearing price of \$7.81/MW-hr for up plus down regulation. But if storage bids \$5.71/MW-hr it will be selected 5168 hours rather than 8760 and receive only \$29,509 for the year, not enough to cover the mortgage. This is not just the tough realities of competitive markets, it is a failure of this market structure because it compares one technology’s capital cost with another technology’s marginal operating opportunity cost.

Both storage and society loose because of the basic structure of the ancillary service markets.

POSSIBLE MARKET STRUCTURE SOLUTIONS

Above we saw that simply bidding the average capital requirement as the hourly storage price would not work. Generation would be selected when its marginal opportunity cost was below the average capital cost of storage. This is economically inefficient because storage could supply regulation at a lower cost than generation during those hours as well, so a more expensive resource would be selected over 40% of the time.

Forcing Storage to Guess. A market solution that would work in principal would be for storage to bid in at just below the cost of generation each hour. Storage would always be selected and in our example it would cover its capital cost. More importantly this would provide a power system solution that was at least slightly lower cost than generation. In a sense it is very similar to the initial example of a ±1MW storage project operating as a price taker in the regulation market. The difference is that in the preceding example the ±1MW storage project simply bids \$0.00/MW-hr and lets the system operator’s co-optimization software calculate the appropriate compensation. In this proposed solution, however, the large storage project (or collection of multiple storage projects) must try to guess the opportunity costs of all of the generators each

hour. This is not practical and not a burden placed on any other technology.

Limiting Market Share. The Electric Reliability Council of Texas (ERCOT) solves a similar problem by limiting demand response to providing no more than half of the 2800 MW Responsive Reserve Service (RRS or spinning reserve) requirement. Demand response is similar to our storage example in that the opportunity cost is very difficult to calculate and is consequently ignored by the co-optimizer.³ Were demand response to clear the RRS market and set the price to zero there would be no incentive for demand response to participate. Limiting demand response to no more than half of the RRS requirement assures that generation, which supplies the other half of the requirement, always sets the market clearing price. ERCOT (and the ERCOT loads) get the benefits of using a lower cost technology to supply half of the RRS requirements and the new technology remains viable. The disadvantage is that the new technology is blocked from half of the market and the incumbent generation technology is guaranteed a half market share even if it is more expensive.⁴ Both the new technology and society suffer.

Storage as a Regulated Asset. If storage is lower cost than generation for provision of one or more ancillary services it may be appropriate to treat storage as a regulated asset like transmission. An analysis would be required to demonstrate that storage was likely to remain lower cost than generation for the life of the storage project. Regulatory approval would also be required. Storage operations would then be placed under the control of the system operator and the amount of that ancillary service procured through the competitive market would be reduced or eliminated.

Long Term System Operator Contracts. An alternative to treating storage as a regulated asset might involve the system operator determining that storage made long term economic sense. The system operator might then enter into a long-term ancillary service provision contract with one or more storage projects. Clearly this would have a major impact on the ancillary service markets (possibly eliminating some) and would require detailed analysis, modification of the ancillary service market design, stakeholder approval, and likely regulatory approval.

Self Provision. Load serving entities (LSEs) often have the option to self provide ancillary services. They do this by designating ancillary service resources and giving control of the resource over to the system operator. About 16% of day-ahead regulation, 6% of day-ahead spin, and 26% of day-ahead non-

spinning reserve were self provided to CAISO in 2013. LSEs might determine that storage was less expensive than paying for their allocated share of the CAISO ancillary service market costs. The LSEs could either invest in storage themselves or they could purchase ancillary services from storage projects. One potential disadvantage is that the ancillary service obligations rise and fall as the LSE's energy market share changes making it difficult for LSEs to make long term commitments to build storage facilities. LSEs might buy and sell storage capability in much the same way they meet their changing energy requirements.

CONCLUSIONS

Power systems require flexible resources to maintain the generation/load balance instantaneously and continuously. They obtain the required flexibility by procuring ancillary services and through sub-hourly energy scheduling. In ISO/RTO areas these services are obtained, and paid for, through ancillary service and five-minute energy markets. Storage technologies with fast and accurate control can be ideal suppliers of the required flexible response. The limited energy of some storage technologies can match the limited response duration required for ancillary services and five-minute energy markets making storage a very good technical match for power system reliability requirements. Unfortunately the ancillary service market structure for all of the ISOs and RTOs is built around generator opportunity costs. While this works well for generators it does not work for energy-limited storage which may not have an energy-market based opportunity cost. The result is that cost effective storage will collapse the ancillary service market prices, driving itself out of business and returning the power system to exclusive reliance on more expensive generation based resources.

Similarly, fast and accurate storage could respond to five-minute energy prices and eliminate the current high within-hour energy price volatility. This would eliminate the need for generators to incur losses and increased maintenance costs as they maneuvered to balance the power system. Unfortunately the storage response would collapse the within-hour energy price spread and eliminate the payment for storage response.

Forcing storage to conform to the current market design is impractical. Attempting to limit storages' own market participation to assure that it never set the market clearing price would require collusion among all storage participants and detailed knowledge of all competing generators. Submitting hourly bids that were just below generator bids, rather than being based on storages' own costs, would require similar in depth knowledge of the competing generators real-time costs. This is not require of any other market participants and is counter to the desirable market structure where participants are encouraged to bid their own marginal costs.

Three market structure solutions appear to be practical. All are based on an analysis being performed that demonstrates that

³ The opportunity cost for demand response is not related to the production cost of energy but is related to the opportunity costs the load faces in the functions it is currently performing. This varies from load to load and is extremely difficult to calculate. The practical solution has been for the power system to ignore demand response opportunity costs.

⁴ There can be other technical reasons for limiting the market share for demand response or any other technology but they are not relevant for this discussion of market structure.

a storage technology is the lowest cost solution over some reasonable investment interval.

LSEs could determine that storage was a lower cost option than relying on the ISO market for ancillary service provision. The LSE would invest in storage and turn control over to the system operator. In practice this would require that the LSE commit to obtaining a specific amount of storage for the life of the storage project or the life of the storage contract, which may be problematic for an LSE with a changing obligation.

The system operator itself could determine that storage was a lower cost solution than use of its own real-time markets. Storage response might then be obtained through long term contracts. Stakeholder and regulatory approval would likely be required.

Regulators themselves could determine that storage was a lower cost solution and in the best interest of ratepayers. Storage would become a regulated asset, like transmission, and be paid for through the transmission tariff.

In any case, ancillary service and sub-hourly energy market reform is likely required for storage (and demand response) to be economically viable even if they are the preferred technical solution.

REFERENCES

- [1] Kirby, B., 2013, "The Value of Flexible Generation", PowerGen 2013, November
- [2] CAISO OASIS web site, <http://oasis.caiso.com/mrioasis>
- [3] NYISO OASIS web site, http://www.nyiso.com/public/markets_operations/market_data/pricing_data/index.jsp
- [4] MISO OASIS web site, <https://www.misoenergy.org/MarketsOperations/RealTimeMarketData/Pages/RealTimeMarketData.aspx>
- [5] ERCOT OASIS web site, <http://ercot.com/mktinfo/dam/>
- [6] ISO-NE OASIS web site, http://www.iso-ne.com/markets/hst_rpts/hstRpts.do?category=Hourly#anchor2

ANNEX A

ANNUAL AVERAGE ANCILLARY SERVICE PRICES [2,3,4,5,6]

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
	Annual Average and Maximum \$/MW-hr											
California												
Regulation	26.9	35.5	28.7	35.2	38.5	26.1	33.4	12.6	10.6	16.1	10.0	7.8
up+down	111	164	166	188	399	421	618	500	124	120	134	121
Spin	4.3	6.4	7.9	9.9	8.4	4.5	6.0	3.9	4.1	7.2	3.3	2.7
	250	92	125	110	225	400	400	416	66	48	130	120
Non-Spin	1.8	3.6	4.7	3.2	2.5	2.8	1.3	1.4	0.6	1.0	0.9	0.2
	92	92	129	125	110	400	399	416	66	35	130	115
Replacement	0.90	2.9	2.5	1.9	1.5	2.0	1.4					
	80	55	90	36	70	175	244					
ERCOT												
Regulation		16.9	22.6	38.6	25.2	21.4	43.1	17.0	18.1	31.3	9.2	13.5
up+down		177	156	1451	351	322	534	528	517	2744	1333	3013
Responsive		7.3	8.3	16.6	14.6	12.6	27.2	10.0	9.1	22.9	9.1	9.8
		150	51	731	351	100	2000	185	125	2606	1456	3000
Non-Spin		3.2	1.9	6.1	4.2	3.0	4.4	2.3	4.3	11.8	6.7	3.5
		249	400	510	125	180	2000	175	296	1500	1461	3000
MISO												
Regulation								12.3	12.2	10.8	7.8	9.1
								52	102	102	145	49
Spin								4.0	4.0	2.8	2.3	3.3
								39	34	29	132	43
Non – Synchronous								0.3	1.5	1.2	1.4	1.8
								25	9	27	132	43
New York East (DA)												
Regulation	18.6	28.3	22.6	39.6	55.7	56.3	59.5	37.2	28.8	11.8	10.4	10.1
	99	195	99	250	250	300	300	500	250	95	96	150
Spin	3.0	4.3	2.4	7.6	8.4	6.8	10.1	5.1	6.2	7.4	6.0	8.6
	150	55	44	64	171	53	68	39	63	81	213	196
Non Spin	1.5	1.0	0.3	1.5	2.3	2.7	3.1	2.5	2.3	3.9	3.8	4.2
	45	3	3	64	171	12	59	10	13	75	213	196
30 Minute	1.2	1.0	0.3	0.4	0.6	0.9	1.1	0.5	0.1	0.1	0.3	0.5
	45	3	3	4	31	9	4	5	6	6	50	20
New York West (DA)												
Regulation	18.6	28.3	22.6	39.6	55.7	56.3	59.5	37.2	28.8	11.8	10.4	10.1
	99	195	99	250	250	300	300	500	250	95	96	150
Spin	2.8	4.2	2.4	4.9	6.0	5.4	6.2	4.2	4.4	3.4	3.1	4.8
	150	55	44	50	45	53	60	25	56	45	50	113
Non Spin	1.4	1.0	0.3	0.6	0.9	1.6	1.7	1.7	0.9	0.1	1.1	1.0
	45	3	3	13	38	12	10	9	10	6	50	27
30 Minute	1.2	1.0	0.3	0.4	0.6	0.9	1.1	0.5	0.1	0.1	0.3	0.5
	45	3	3	4	31	9	4	5	6	6	50	20
New England												
Regulation (+”mileage”)			54.6	30.2	22.3	12.7	13.8	9.3	7.1	7.2	6.7	11.7
			344	561	100	100	100	100	82	95	70	692
Spin					0.3	0.4	1.7	0.7	1.8	1.0	1.7	3.0
					72	179	716	121	638	418	313	1026
10 Minute					0.1	0.3	1.2	0.5	1.2	0.4	1.0	2.5
					50	154	716	114	638	418	288	1026
30 Minute					0.0	0.1	0.1	0.1	0.4	0.3	1.0	2.3
					16	100	76	75	112	136	279	500