

## SSD PERDAS 4.0 - SISTEMA DE SUPORTE À DECISÃO EM GESTÃO DE PERDAS 4.0

### **Caroline Pereira Carvalho**

Engenheira civil, especialista em Gestão de Projetos e atualmente Gerente da Divisão de Controle de Perdas da Unidade de Negócio Leste - ML, trabalha na Companhia de Saneamento Básico do Estado de São Paulo – Sabesp há 10 anos e foi a *Product Owner* do Integra 4.0, projeto de transformação digital da ML que originou o SSD Perdas 4.0.

### **Alexandre Souza Dias**

Tecnólogo em Hidráulica, possui 20 anos de atuação no setor pela Sabesp, compondo atualmente a Divisão de Controle de Perdas. Participou no mapeamento, desenvolvimento e validação do Integra 4.0 e SSD Perdas 4.0, tanto como integrante do *squad* de perdas reais quanto como *Scrum Master* durante o processo de expansão da plataforma para toda a Companhia.

### **Celso Roberto Alves da Silva**

Engenheiro Civil com Especialização em Saúde Pública e Saneamento Básico, com mais de 30 anos de atuação no setor pela Sabesp. Há 6 anos se dedica no desenvolvimento de soluções para controle de perdas de água dos sistemas de distribuição de água.

### **Luis Alberto Rodrigues**

Engenheiro civil, trabalha na Companhia de Saneamento Básico do Estado de São Paulo – Sabesp há 31 anos. Já atuou como gerente nas divisões de Cadastro Técnico e Adução, e atualmente integra a equipe de Controle de Perdas da Unidade de Negócio Leste - ML

### **Mario Masakatu Tomita Junior**

Engenheiro Civil e especialista em Saneamento Ambiental, trabalha na Companhia de Saneamento Básico do Estado de São Paulo há 5 anos e atualmente integra a equipe da Divisão de Controle de Perdas, supervisionando a macromedição, monitoramento e projeção dos Indicadores de Desempenho do processo de Perdas. Foi integrante do *squad* de Perdas Reais do Integra 4.0 e SSD Perdas 4.0 apoiando o mapeamento e *quality assurance do projeto*.

**Endereço:** Rua Najatu, 72 – Penha – São Paulo – SP – CEP: 03613-050 – Brasil – Tel: +55 (11) 2681-3977 – Fax: +55 (11) 2681-3755 – e-mail: cpcarvalho@sabesp.com.br

## **RESUMO**

A gestão de perdas na Unidade de Negócio era realizada através da obtenção de dados de vários sistemas distintos e geração manual de relatórios. Desta forma, os analistas precisavam dedicar mais tempo e esforço para gerar as informações, restando menos tempo para análises e tomada de decisão.

Assim, desenvolveu-se o SSD - Sistema de Suporte à Decisão em Gestão de Perdas, que integra os dados dos sistemas corporativos em uma única plataforma, com apresentação em dashboards.

O sistema coleta automaticamente e trata os dados dos sistemas legados, com aplicação de regras de negócio, gera as informações necessárias para a gestão dos indicadores, bem como suas correlações, utilizando inteligência artificial. As correlações de dados são feitas entre as variáveis e indicadores, possibilitando a predição de variáveis e a identificação de eventuais desvios no padrão das informações

Os dashboards permitem visualizar a evolução dos indicadores e suas variáveis de controle, suas projeções e as variações de ligações e de volumes de água disponibilizada, consumida e medida, fundamentais para a Gestão de Perdas, atualizadas diariamente.

A plataforma permite o acompanhamento da operação dos sistemas de distribuição, podendo identificar eventos atípicos.

Como resultado, o SSD possibilitou maior agilidade e precisão na Gestão de Perdas.

**PALAVRAS-CHAVE:** Indústria 4.0, Redução de perdas de água, Transformação digital

## CONTEÚDO DO TRABALHO

### INTRODUÇÃO

A Unidade de Negócio atua na distribuição de água e na coleta, transporte e afastamento de esgoto de mais de 3 milhões de habitantes. Parte importante de suas atividades é a gestão de combate às perdas de água, que utiliza diversos sistemas corporativos para auxiliar no monitoramento e controle do abastecimento de água. Esses sistemas informatizados exigiam coleta, tratamento dos dados e elaboração de relatórios de forma manual para viabilizar as análises, o que resultava em um grande esforço e investimento de tempo de seus analistas na construção das informações.

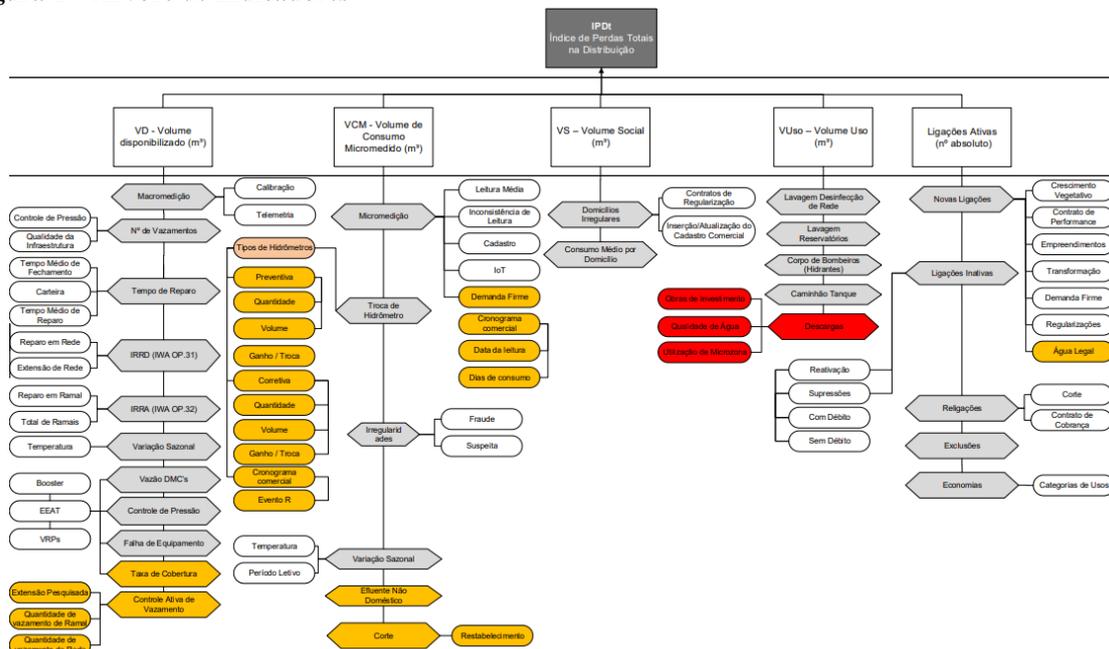
### OBJETIVO

O objetivo do Sistema de Suporte à Decisão em Gestão de Perdas é a aplicação dos conceitos da Indústria 4.0 e a busca da transformação digital através da integração dos dados dos diversos sistemas legados em uma única plataforma, coletando automaticamente os dados e gerando as informações para a gestão dos indicadores, bem como suas correlações através do uso de inteligência artificial viabilizando projeções, planejamento, tomada de decisão e possíveis mudanças de rumo de forma mais ágil e precisa.

### METODOLOGIA UTILIZADA

Inicialmente foram mapeados os indicadores relativos ao processo de perdas, suas variáveis de controle e todos os dados que influenciam nessas variáveis, conforme Figura 1 a seguir:

**Figura 1 – Árvore de Indicadores**



Fonte: Elaborado pelo autor – Acervo pessoal

Em seguida, foram identificadas as fontes dos dados necessários para a composição das variáveis de controle dos indicadores e demais fatores intrínsecos ao processo (Figura 2), bem como suas ineficiências e intervenções necessárias para viabilização do projeto. Com as fontes mapeadas foram elaboradas as regras de negócio para indicar como estes dados deveriam ser tratados e relacionados visando atender os critérios corporativos existentes (Figura 3).

**Figura 2 – Mapeamento de Fontes de Dados**

| NOME DA TABELA                   | Descrição do conteúdo                               | CAMADA BI:        | NOME DO RESPONSÁVEL:              |
|----------------------------------|---|-------------------|-----------------------------------|
| CONSOLIDADO_SETOR_ABASTECIMENTO  | Consolidado do Setor de Abastecimento               | DATALAKE (RAW)    | XXXXXXX                           |
| Localização da fonte:            | Formato de armazenamento da fonte:                  | Número de Colunas | Frequência de Atualização do dado |
| XXXXXXXXX                        | (.txt)  | 8                 | Diário                            |
| Localização do destino:          | Formato de armazenamento do destino:                | Número de Colunas | Frequência de Atualização do dado |
| /user/hive/raw/consolidado_setor | PARQUET   | 10                | Diário                            |
| Atributos                        |   |                   |                                   |
| CONSOLIDADO_SETOR_ABASTECIMENTO  |   |                   |                                   |
| Campos                           | Descrição   | Tipo              | Posicao coluna                    |
| NUMERO_PONTO_MEDICAO             | Ponto de medição                                    | STRING            | 1                                 |
| DATA_HORA                        | Data de leitura (1/1/2015 00:00:00)                 | TIMESTAMP         | 2                                 |
| VAZAO_M3                         | Vazão diário m3/dia                                 | STRING            | 3                                 |
| VAZAO_L_S                        | Vazão média em l/s                                  | STRING            | 4                                 |
| LEITURA_MEDIDOR                  | Volume medido no totalizador                        | STRING            | 5                                 |
| PORCENTAGEM                      | Variação da vazao diária em relação ao dia anterior | STRING            | 6                                 |
| CODIGO_ESTACAO                   | Código do SCOA-PI                                   | STRING            | 7                                 |
| NOME_PONTO_MEDICAO               | Nome do Ponto de medição.                           | STRING            | 8                                 |
| FILENAME                         | Arquivo de origem                                   | STRING            | 9                                 |
| DT_INGESTAO                      | Data da ingestão                                    | TIMESTAMP         | 10                                |

Fonte: Elaborado pelo autor – Acervo pessoal

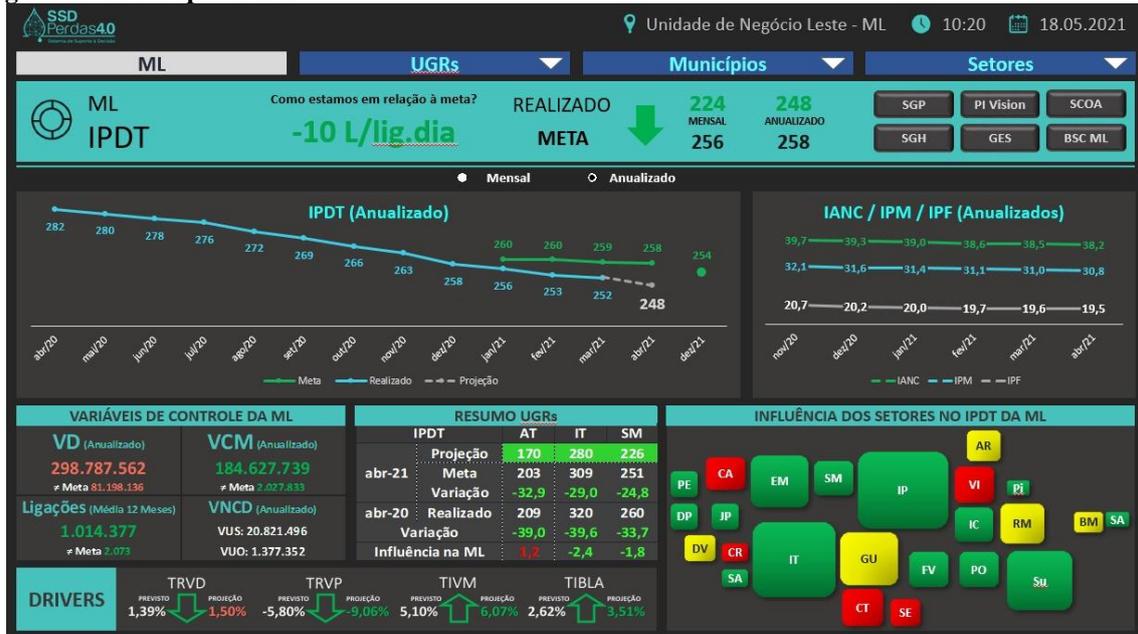
**Figura 3 – Mapeamento de Regras de Negócio**

| Nome da Tabela          | ALFA  | Nome do Cliente  | TRVP (Taxa de Redução do Volume Perdido)  |  |                           | Corretor de Dados | xxx      |
|-------------------------|---|--|---|--|---------------------------|-------------------|----------|
| Fonte de dados          | Endereço da fonte   | Tipo de Fonte (Nativa e ou Intermediária)  | Tipo de Arquivo (txt, xls, csv, BD, etc.) | Frequência de Atualização do dado  | Periodicidade de consulta |                   |          |
| VD Consolidado MICO     | Servidor Integra  | Intermediária  | BD  | D-1  | Diária                    |                   |          |
| VD Dado Bruto PI Vision | Servidor Integra  | Nativa   | BD  | hora-1   | horária                   |                   |          |
| VCM (Fontes Diversas)   | Servidor Integra  | Nativa   | BD  | D-1  | Diária                    |                   |          |
| VUS (Dados Declarados)  | Servidor Integra  | Nativa   | BD  | mensal   | Mensal                    |                   |          |
| Regras de Negócio       |   |  |   |  |                           |                   |          |
| Variável de Controle    | Título da Regra   | Formulas de cálculo  |   | Dados necessários  | Nome do Driver            | Granularidade     | Atualiza |
| TRVP                    | Variação percentual do volume perdido anualizado em relação ao mesmo mês do ano anterior. | <p><b>Objetivo:</b> Acompanhar a variação do volume perdido anualizado.</p> <p>1 - O sistema deverá capturar os dados dos volumes disponibilizados (VD) e atualizados referentes à área consultada no dashboard. No caso da projeção em VD será proveniente da tabela VD Dado Bruto PI Vision. Caso na consulta esteja selecionado um mês fechado o VD será proveniente da tabela VD Consolidado MICO.</p> <p>2 - O sistema deverá capturar os dados dos volumes consumidos medidos (VCM) atualizados referentes à área consultada no dashboard.</p> <p>3 - O sistema deverá capturar os dados dos volumes usos sociais (VUS) atualizados referentes à área consultada no dashboard.</p> <p>4 - O sistema deverá repetir os passos 1, 2 e 3 para capturar os volumes referentes ao mesmo mês em estudo no ano anterior.</p> <p>5 - O sistema deverá subtrair de VD seus respectivos VCM e VUS dentro do mês em análise, obtendo desta forma o volume perdido (VP), da seguinte forma:</p> $VP = VD - VCM - VUS$ <p>6 - O sistema deverá repetir o passo 5 para os volumes referentes ao mesmo mês do ano anterior.</p> <p>7 - Para cálculo da TRVP o sistema deverá obedecer a seguinte fórmula:</p> $TRVP = \left( \frac{VP \text{ atualizado mes ano} - VP \text{ atualizado mes ano anterior}}{VP \text{ atualizado mes ano anterior}} \right) * 100$ |   | Referente à área que estiver selecionada no dashboard o sistema deverá utilizar os seguintes dados:<br>- VD atualizado do mês atual.<br>- VCM atualizado do mesmo mês no ano anterior.<br>- VUS atualizado do mês atual.<br>- VUS atualizado do mesmo mês no ano anterior.<br>- VUS atualizado do mesmo mês no ano anterior. | VD<br>VCM<br>VUS          | mensal            | diário   |

Fonte: Elaborado pelo autor – Acervo pessoal

Após o mapeamento das regras de negócio foram construídos os protótipos dos dashboards. Estes protótipos continham as informações que deveriam ser exibidas, bem como sua disposição (layout) para que pudessem ser visualizadas em formato de gráficos, tabelas e diagramas. Para cada indicador ou variável foi desenvolvido um protótipo (Figura 4).

**Figura 4 – Protótipo do IPDT**



Fonte: Elaborado pelo autor – Acervo pessoal

Para o desenvolvimento do sistema foram utilizadas ferramentas da metodologia ágil, através da criação de squads, times multidisciplinares que realizavam pequenas entregas de valor, denominados mvp (minimum viable product), ou seja, o menor produto viável desenvolvidos e validados em curtos intervalos de tempo, as chamadas Sprints, preceitos do framework Scrum. A aplicação deste modelo de trabalho no desenvolvimento visou o pleno atendimento às expectativas das partes interessadas e a identificação e correção das ineficiências em um tempo hábil, mitigando esforços direcionados ao retrabalho e otimizando os investimentos do projeto.

Em virtude da quantidade massiva de dados gerados por equipamentos de campo e clientes, o desenvolvimento só foi possível através da utilização de uma plataforma Big Data, responsável pela coleta, armazenamento e disponibilização destes dados em tempo real, através da aplicação automática de tratamentos e regras pertinentes identificados no processo de mapeamento, garantindo assim a integridade e consistência das informações disponibilizadas aos usuários.

A integração aos dados dos sistemas legados foi fundamental, pois apenas com a automação na obtenção dos dados gerados a partir dos equipamentos telemetrizados (IoT) como medidores de vazão, pressão e nível de reservatório, ou ainda os dados gerados manualmente como volumes produzidos nos sistemas isolados ou volumes de usos, tornou-se viável a aplicação de um modelo de Inteligência Empresarial (Business Intelligence – BI).

**Figura 5 - Equipamentos de medição inteligente**



Fonte: Elaborado pelo autor – Acervo pessoal

No SSD Perdas também foram implementados algoritmos de aprendizagem de máquina e inteligência artificial na previsão de variáveis como volumes de água disponibilizada, consumida e incremento da base de clientes, que utilizou além de suas próprias séries históricas, elementos externos e sazonais como a pandemia e temperatura, de modo aprimorar as projeções e aplicações de controle estatístico nos indicadores de perdas. Portanto, os dados são correlacionados para possibilitar a detecção de possíveis anomalias no sistema de distribuição, previsões de variáveis e a efetividade das ações planejadas e realizadas

A partir dos protótipos de dashboards elaborados e a utilização de uma ferramenta de BI, foram desenvolvidos painéis, onde as informações devidamente processadas são apresentadas, sejam elas variáveis de controle ou indicadores de desempenho, em suas diversas granularidades e periodicidades disponíveis. Seu acompanhamento pode ser realizado através de conjuntos de telas presentes em salas construídas para este propósito, mas também em computadores, notebooks e celulares corporativos pelos usuários devidamente habilitados.

**Figura 6 – Salas de Monitoramento**



Fonte: Elaborado pelo autor – Acervo pessoal

## **ANÁLISE E DISCUSSÃO DOS RESULTADOS OBTIDOS**

O SSD Perdas permite a disponibilidade, confiabilidade, integridade e autenticidade dos dados, proporcionando ganho de escala com a eliminação de atividades sobrepostas e redução de atividades manuais na geração de dados e informações, otimizando a utilização de recursos financeiros e humanos da companhia. Outro resultado relevante é a obtenção de indicadores e suas relações de causa e efeito em tempo real. Isto permite a previsibilidade de resultados, análise e tomada de decisão orientada por dados (data driven) e visão integrada dos processos principais do negócio.

A aplicação do modelo de gestão propiciado pelo SSD Perdas é pioneira no saneamento brasileiro, visto que é aderente ao modelo da Indústria 4.0 e contribui para a transformação digital do setor, na busca da excelência no modelo de negócio a fim de aumentar a competitividade, a qualidade na prestação de serviços, a preservação dos recursos naturais e atender os compromissos estabelecidos no Marco Legal do Saneamento.

Mediante a arquitetura empregada no desenvolvimento do SSD Perdas, cujos dados dos diversos sistemas legados abrangem toda a companhia, o sistema está em expansão para atender às demais unidades de negócio, proporcionando uma padronização das informações que balizam o combate às perdas de água.

## **CONCLUSÕES/RECOMENDAÇÕES**

A implementação do Sistema de Suporte à Decisão em Gestão de Perdas trouxe grandes benefícios tanto para a Unidade de Negócio quanto para a companhia, promovendo mais uma etapa no processo de transformação digital, possibilitando a redução de tempo e esforço na obtenção de dados e geração de relatórios, bem como a otimização de recursos humanos e financeiros. Além disso, o sistema proporciona uma análise mais completa e eficiente dos indicadores e as correlações entre os dados, além da tomada de decisão de forma mais ágil e precisa.

Todavia, por tratar-se de um sistema de Big Data, é imprescindível que haja na ponta a instrumentação necessária para criação destes dados, como medidores de vazão e pressão. Portanto, o nível de maturidade digital da companhia foi fundamental no sucesso do SSD Perdas. No entanto, ainda que a expansão vise atender toda a companhia, algumas localidades ainda precisam ampliar sua base de elementos geradores de dados para atingir um patamar desejável. Para que o modelo seja replicado em outras empresas do saneamento, é importante observar seus graus de maturidade digital e os aprimoramentos necessários para sua viabilização.

#### **REFERÊNCIAS BIBLIOGRÁFICAS:**

- 1 ALEGRE, H. et al. Performance Indicators for Water Supply Services, 2nd Edition. IWA Publishing. IWA–International Water Association, 2006.
- 2 TSUTIYA, M. T. et al. Abastecimento de Água, 2ª Edição, São Paulo, Departamento de Engenharia Hidráulica e Sanitária da Escola Politécnica da Universidade de São Paulo, 2005.
- 3 SILVA Jr., E. G. e CABRAL, R. C. Indicador de Vulnerabilidade da Infraestrutura – Uma Proposta para o Diagnóstico e Tomada de Decisões no Combate às Perdas Reais. Revista DAE nº 55. São Paulo, 2015.